EAI Endorsed Transactions
on Smart Cities
Research Article **EAI.EU**

Data Fusion and Visualization towards City Disaster Management: Lisbon Case Study

Luís B. Elvas[1*], Sandra P. Gonçalves[1], João C. Ferreira[1,3] and Ana Madureira[2,3]

[1]Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisboa, Portugal
luis.elvas@iscte.pt; sandra_goncalves@iscte-iul.pt ; jcafa@iscte-iul.pt
[2]Instituto Superior de Engenharia do Porto – P.PORTO, ISRC, Porto, Portugal
amd@isep.ipp.pt
[3]Inov INESC Inovação—Instituto de Novas Tecnologias, 1000-029 Lisbon, Portugal

Abstract

INTRODUCTION: Due to the high level of unpredictability and the complexity of the information requirements, disaster management operations are information demanding. Emergency response planners should organize response operations efficiently and assign rescue teams to particular catastrophe areas with a high possibility of surviving. Making decisions becomes more difficult when the information provided is heterogeneous, out of date, and often fragmented.
OBJECTIVES In this research work a data fusion of different information sources and a data visualization process was applied to provide a big picture about the disruptive events in a city. This high-level knowledge is important for emergency management authorities This holistic process for managing, processing, and analysing the seven Vs (Volume, Velocity, Variety, Variability, Veracity, Visualization, and Value) in order to generate actionable insights for disaster management.
METHODS: A CRISP-DM methodology over smart city-data was applied. The fusion approach was introduced to merge different data sources.
RESULTS: A set of visual tools in dashboards were produced to support the city municipality management process. Visualization of big picture based on different data available is the proposed work.
CONCLUSION: Through this research, it was verified that there are temporal and spatial patterns of occurrences that affected the city of Lisbon, with some types of occurrences having a higher incidence in certain periods of the year, such as floods and collapses that occur when there are high levels of precipitation. On the other hand, it was verified that the downtown area of the city is the most affected area.

Keywords: Disaster Management, Data mining, Smart City, CRISP-DM.

Received on 12 March 2022, accepted on 31 May 2022, published on 06 June 2022

doi: 10.4108/eetsc.v6i18.1374

1. Introduction

Natural and man-made disasters have become more common across the world, with devastating repercussions reflected in the loss of human life and material/facilities destruction [1]. In reality, 3 751 natural catastrophes such as earthquakes, tsunamis, and floods have been observed globally in the previous ten years, causing $1 658 billion in damages and affecting more than 2 billion people [2]. As a result, disaster management measures must be implemented in order to reduce the risks.

Catastrophe management is a comprehensive process with the core aims of avoiding, reducing, responding to, and recovering from disaster impacts in the system. Disaster response requires a variety of groups, including governmental, public, and private organizations, as well as several tiers of authority, due to the complexity of major situations [3]. The engagement of several entities in disaster management procedures emphasizes the need of cooperation and coordination systems, since these agencies

*Corresponding author. Email: luis.elvas@iscte.pt

must communicate, coordinate, and work with one another in order to be effective in a catastrophe situation.

Some issues, such as a lack of situational awareness or the difficulties in deploying technical solutions for disaster response due to their high costs, may make communication amongst stakeholders' problematics [4]. Cities must provide better services and infrastructure to their citizens since population density and the frequency of catastrophes have increased in recent years.

In this environment, the Smart City (SC) concept emerges as the appropriate answer for overcoming the problems posed by globalization and urbanization [5]. Cities aiming to achieve SC status employ digital and networked technology to solve a variety of issues, including enhancing service quality, becoming more sustainable, boosting the local economy, improving quality of life, and ensuring the safety and security of their residents [6].

Electronic devices and network infrastructures are combined in a SC to gain high-quality services, and when cities acquire the most up-to-date network infrastructure, smart devices, and sensors, a large quantity of data is collected, referred to as Big Data (BD).

This data may contain a considerable quantity of contextual, geographical, or temporal information [7]. In catastrophe scenarios, BD plays a critical role in disaster management procedures because it is feasible to use data mining (DM) and analytical tools to examine trends and forecast disasters, allowing the creation of appropriate disaster management plans based on historical data [6]. In this sense, the use of BD technologies aids agents in decision-making by allowing them to recognize possible risks and, as a result, establish suitable plans to deal with catastrophic circumstances, so increasing the SC's resilience [2].

The goal of this study is to use a data-driven strategy to extract knowledge regarding catastrophes in the context of a SC to improve the city's management. In order to achieve a descriptive and predictive analysis of data given by the Lisbon City Hall, which includes information on occurrences that have happened in the city.

This research will be conducted using multiple data sources, where from the data collected of the firefighter's incidences between 2011 and 2018, we will merge datasets containing the average age of the buildings in each parish, the number of populations, and temperature, allowing us to perform a complete and deep descriptive analysis finding patterns between these variables, the types of incidents and the area where they occur.

The investigation was carried out in two parts in both situations, with the first phase focusing on a general examination of the reported events and the second phase focusing on occurrences that harmed city structures.

2. State of the art

Due to the large number of works that have been produced, data-driven disaster management is a new sector that has been evolving [8].

In this regard, using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) approach [9] and the Systematic Literature Review stages described by Okoli and Schabram [10], a survey and critical appraisal of the literature related to the chosen issue was conducted.

As a result, a systematic search was done on the issue in two electronic databases: Scopus [11] and Google Scholar [12], with the primary goal of identifying and selecting research publications relating to data-driven disaster management research. With this in mind, a question was posed in order to narrow down the work done in this area.

The question is as follows: (("Disaster Management" OR "Incident Management") AND ("smart city" OR "data analysis" OR "data mining" OR “big data”)). Additionally, a ten-year time window was defined (2010-2020), and the research covered areas such as Decision Science, Computer Science, Environmental Science, and Engineering. In terms of document typology, only journal articles, articles, and book chapters were considered. The documents were selected through the abstract and in cases where the information contained in the abstract was not sufficiently complete, the document was consulted in its entirety. The work done in this area covers both natural and man-made disasters.

2.1. Natural disasters

Natural catastrophes are distinguished by the significant influence they have on society, disrupting its regular functioning. In the field of data-driven disaster management, work has been done to develop decision support systems that aid decision-makers in making faster and better-informed judgments based on analytical results. With this in mind, Jeong and Kim [13] undertook a statistical study of electrical mishaps such as fires and system failures that occurred in Korea as a result of climate change. A relationship was established between climate change and electrical-equipment-related incidents in this investigation.

In 2017, another research [14] found a correlation between BD systems and disaster management. To examine hydroclimate data, Big Data Analytics tools were applied to a dataset from Malaysia's National Hydraulic Research Institute. The purpose was to gain knowledge about climate change and use it to prepare for, reduce, respond to, and recover from natural catastrophes. The use of BD technologies enabled the identification of exceptional precipitation and runoff events, as well as the tracking of drought occurrences. Briones-Estébanez and Ebecken [15] used DM approaches to discover and evaluate trends in the incidence of widespread and intense occurrences, such as floods, river overflows, and landslides, in five Ecuadorian cities. Other works have been produced to undertake a quantitative study of the damage caused by natural disasters, in addition to works made to assess catastrophes from a spatial and temporal perspective. Park et al. [16] used a similar technique to evaluate the potential impacts or effectiveness of damage caused by three types of catastrophes in Korea, including typhoons, severe rain, and earthquakes, on water delivery systems.

The work done in the field of data-driven disaster management is diverse, since different methodologies are used to make data available to decision-makers. In the work by Saha, Shekhar, and Sadhukhan [17], the analytical results were presented in a more iterative manner by constructing a dashboard to forecast and identify flood-prone locations in West Bengal, India, utilizing geographic map visualization. Other research [18]–[21] created catastrophe susceptibility maps using a mix of DM and GIS methodologies. The main goal of these studies is to identify and classify places that are prone to natural catastrophes, with the exception that various DM models are employed in different research projects.

2.2 Man-made disasters

In the case of man-made disasters, Smith et al. [22] conducted study on the use of Big Data technology for disaster management. They analysed a dataset about fires that happened in Australia using the statistical program R, as well as its graphical capabilities.

Balahadia et al. [23] used the K-means clustering technique to establish patterns and build clusters of fire incidents based on data from fires that happened in Manila, Philippines. In summary, the purpose was to collect fire event characteristics that might be utilized for risk assessment and risk management in the case of such catastrophes, as well as to aid in the creation of preventative measures.

Asgary et al [24] attempted to evaluate the geographical and temporal patterns of fire-related occurrences in Toronto, Canada, using spatiotemporal approaches. The link between the economic, physical, and environmental features of distinct communities and the overall number of fires that occurred in those neighbourhoods was analysed to extract insights.

A DM technique based on utilizing Bayesian Network to model building fires in urban settings was suggested by Liu et al. [25]. They examined the potential fire risk based on building design characteristics and environmental variables using historical data of fires in a Chinese city between 2014 and 2016. Lee et al. [26] used the Support Vector Machine model to investigate the link between building attributes, inhabitants, and fire incidences in Sydney in another study aimed at analysing fire trends.

Finally, in a study by Wan, Xu, He, and Wang [27], BD technologies were used to investigate the distribution and influence factors of harmful gases in the Chongqing city's urban underground sewage pipe network, as well as the impact of smart city developments on harmful gases in the urban underground sewage pipe network.

In the particular case of Lisbon, we can see that author on study [28], aim to increase catastrophe resilience in a smart city, offering an integrated resilience system that connects interrelated vital infrastructures, increasing the total resilience capability of the city by allowing it to plan, adapt, absorb, respond, and recover from disasters by utilizing the linkages between its numerous essential infrastructures.

Regarding incidents management and data analytics over Road Accidents, authors on studies [29], [30], recurring to the data fusion of several data sources, achieve conclusions that the accidents are due to human factors, occurring mostly on good weather, and where environmental factors may impact their severity, and noticing that most incident occur on the historical part of the city, where the majority of older buildings are present.

In summary, the literature review revealed that the majority of the research in this field was conducted in China, and that the research in this field covers both natural and man-made disasters, with a predominance of flood incident analysis in natural disasters and fire-related incident analysis in man-made disasters.

3. Methodology

This study analysis has its main focus that on performing a spatial-temporal analysis of occurrences collected in Lisbon to extract knowledge about the circumstances in which they occur. The dataset from the Fire Brigade Regiment was subjected to the Cross-Industry Standard Process for Data Mining (CRISP-DM) [31] technique to extract insights on disasters that impact the city of Lisbon with a focus on buildings.

The CRISP-DM methodology-based analysis approach began with a business knowledge that allowed the project's scope to be contextualized and understood.

In this way, a commercial problem was assessed by looking at several features of the city of Lisbon from multiple views, such as demographic, climatic, and educational perspectives. Following the completion of the business understanding phase, the following phases were data understanding, data preparation, modelling, and assessment.

In order to make the data more valuable and extract more information from it, it was necessary to use data mining techniques such as feature engineering, creating new variables from the ones we have, making the information more valuable, and using data integration and data fusion techniques, where we join several data sources, bringing more richness and knowledge to our dataset and study

The firefighter's dataset, given by Lisbon City Hall, is a CSV file that provides information on the incidents that the firemen have reported. The description of the event, the date of the occurrence, the location of the occurrence (i.e., latitude, longitude, and address), and the human (number of people) and material resources (number of vehicles) assigned to each occurrence are all covered by information.

There are 135 200 entries (rows in the CSV file) and 22 characteristics in the dataset, which spans 2011 to 2018 (columns in the CSV file). The columns are all of the type "object," and 13 of them have null values.

During the data preparation, it was discovered that the years 2011 and 2012 have much less data than the others, thus those years were omitted from the study so that all years have representative data. In addition, during this phase, cleaning procedures such as column format conversion were used, as well as the selection of relevant features/characteristics for analysis, with attributes that did not add value to the scope of the study being deleted. Because the null values could not be replaced by the mean or median because they are geographic coordinates, parishes, and descriptions of occurrences, the records with null values were removed.

In this study we have created new attributes from existing attributes, such as the type of street, where from the address of the occurrence, we have created a new variable that gives the information if the accident occurs on an Avenue, on a Street, on a Square, etc. We have also done data fusion between data from multiple source such as INE [32] and IPMA [33]. This external data brings valuable information about the weather on the period of the incidents, and demographic and architectural aspects of the city of Lisbon. For example, we have merge data that gives us the information of the average age of buildings on each parish, as well as the fraction of structures in need of substantial repairs or that are severely deteriorated. Meteorological factors such as average air temperature, relative humidity, average wind speed, and precipitation define the city of Lisbon.

Finally, because there were so many different sorts of occurrences in structures, it was important to categorize them. Categorization helps with visual analysis. The "Occurrence Description" column contains information on the sorts of occurrences, and this property has 25 categories of occurrences established by the firefighters' occurrence management system. The following seven categories were created from the 25 different sorts of occurrences: Infrastructures – Collapse, Infrastructures – Floods, Infrastructures – Landslides, Fire, Accidents (with machinery or elevators), Industrial Technology – Gas Leak, and Industrial Technology – Suspicious Situations are all examples of industrial technology (check smoke or check smells).

The modelling step begins when the data preparation phase is completed. This phase focuses on gathering information that will assist decision-makers in effectively managing the city in the event of a crisis. The first step in the process is to figure out how the data has changed over time. It was feasible to verify that the number of events recorded in the firefighter's occurrence management system decreased from 2013 to 2018, however this decline was not linear since there were fluctuations over the years. In the year 2013, there were 17 176 incidents, 17 607 occurrences in 2014, 16 717 occurrences in 2015, 15 089 occurrences in 2016, 17 582 occurrences in 2017, and 13 368 occurrences in 2018.

Firefighters are called to a variety of situations involving a variety of actions. The types of occurrences were evaluated for a better understanding of the actions conducted by firemen, and it was confirmed that the distribution is not balanced among the nine categories of occurrences reported in the dataset. There is an over-representation of one category, namely Services, which accounts for 45.6 % of all occurrences in the dataset. Road cleaning, opening and closing doors, hospital transport, water supply, and preventative services during shows, sports, and patrols are all included in this category.

Infrastructure and communication route incidents, which include collapses, floods, landslides, falling trees and structures, and falling electric wires, account for 14.7 percent of the total number of occurrences in the dataset. Accidents, which include train accidents, road accidents, and accidents involving equipment (elevators, escalators), account for 10.1 percent of all occurrences.

Activities, with 5.9% of the total occurrences, Industrial-technological with 5.1 percent, Legal conflicts with 0.5 percent, and civil protection incidents with 0.004 percent of the total occurrences are the categories with the least representation in the dataset.

The study focuses on the occurrences that occurred in the buildings of the city of Lisbon to classify them geographically and chronologically after a broad examination of the types of occurrences.

As indicated in Figure 1, the types of incidents that most damage the structures in the city of Lisbon are collapses (1816 records) and floods (1778 records), followed by suspicious situations (including verification of odours and smoke) (1478 records). Failing structures, with a total of 1234 records, accidents with equipment's with 1166 records, fires in buildings with 926 records, and with equipment with 646 records. Other types of incidents, have a representation of 631 records. All have a considerable number of incidents but are less expressive when compared to the previously stated categories.

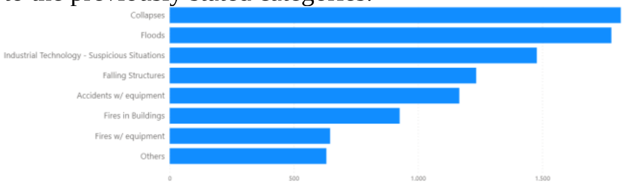

Figure 1. Number of each type of incident

When these events are evaluated over time, i.e., their distribution over years (Figure 2), it is confirmed that some occurrences, such as collapses, suspicious circumstances (checking smoke or odors), and floods accidents, occur in higher proportion over time.

Flooding had a greater prevalence in 2013 and 2014, then declined in the following years.

Focusing the analysis on each occurrence to extract insights about its pattern of occurrence over the course of the year, it is possible to verify that in the case of the infrastructure categories, i.e., collapses and floods. Represented in Figure 3, we can see that accidents related with equipment occur most during the middle of the summer until the beginning of the winter, they are evenly distributed throughout the region of Lisbon, but with greater concentration in the city centre, with regard to the means of intervention, this type of incident requires an average of 6.2 persons and 1 vehicle.

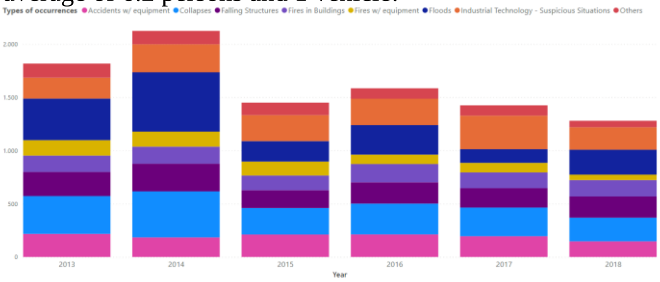

Figure 2. Types of Incidents per Year

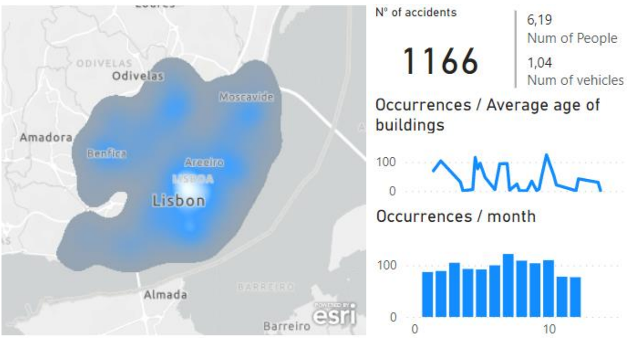

Figure 3. Equipment Incidents

Regarding collapses, on Figure 4, we can see that this occurs more frequently in the autumn and winter months, with maximum values (over 150 records) in the months of October and January. The frequency of recordings of this sort of occurrence declines as the spring and summer months approach, peaking at lower levels in the summer peak. This type of incidents is mainly concentrated on downtown Lisbon and on the city centre, with an average of 7.2 people and 1,4 vehicles per intervention.

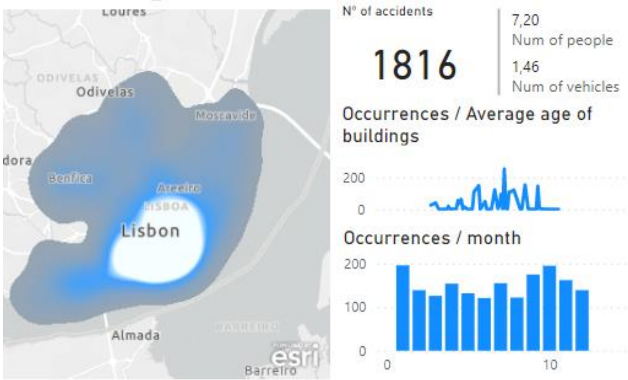

Figure 4. Collapses

Concerning Falling Structures, on Figure 5 it is noticeable the difference between the winter and the summer months, where they are considerable higher when we compare January and October with the others. This type of accidents, occur mostly on downtown Lisbon.

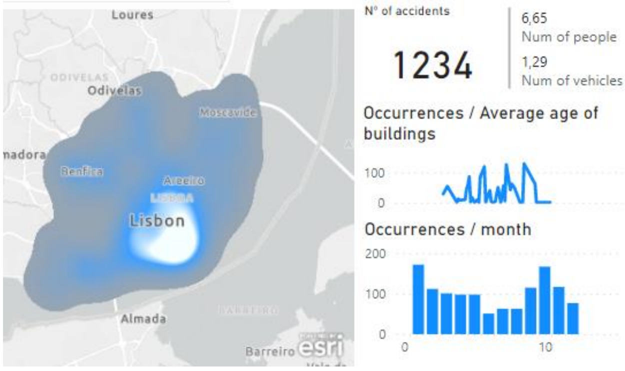

Figure 5. Falling Structures

On Figure 6, we can see the distribution of fires in buildings. This kind of accidents are the ones taking more personnel, with a number of vehicles that quadruples the average of the other accidents, and the triple number of people. The monthly distribution is almost evenly, with a higher number during the cold months, since it his when people turn on the heaters, being among them fireplaces and braziers.

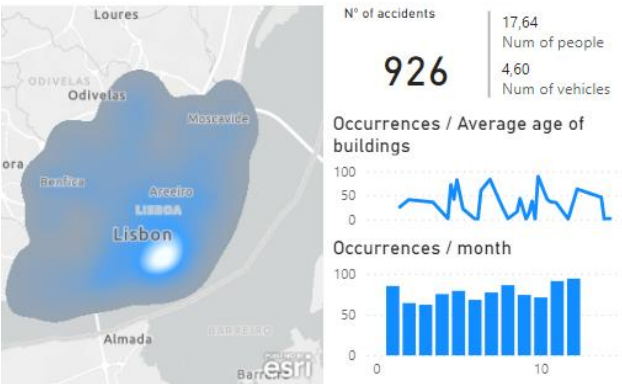

Figure 6. Fires in Buildings

In terms of floods, Figure 7, shows us that the winter months have a greater incidence, with the greatest values in the months of October to January, whereas the summer months have significantly lower values when compared to the winter months.

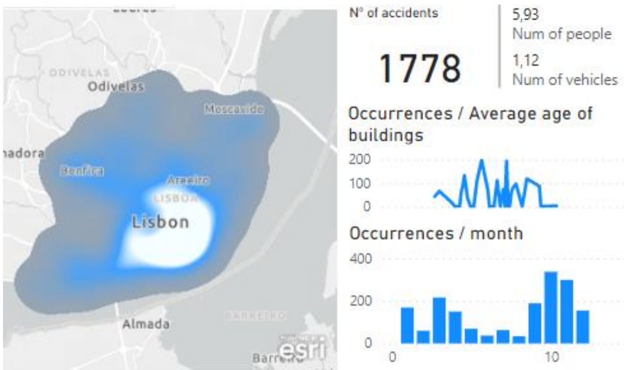

Figure 7. Floods

On Figure 8, we can depict that suspicious situations occur more frequently in the winter months, especially in December, and are comparable to the categories outlined above.

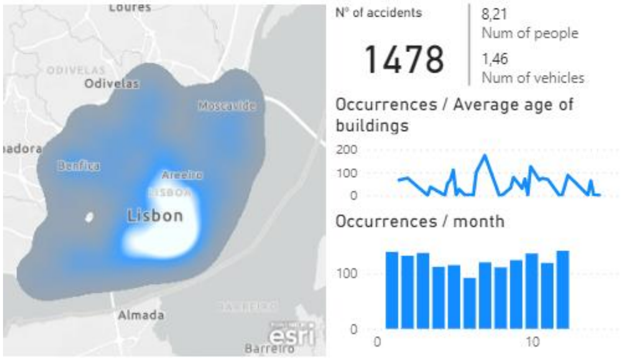

Figure 8. Suspicious Situations

The first analysis revealed that certain types of occurrences have a higher incidence in specific seasons of the year, such as collapses, floods, and suspicious situations (checking for smoke or odors), which have a higher incidence in the winter/spring months. The impact of weather conditions on the incidence of various types of events affecting the city of Lisbon has been confirmed. With this in mind, the impact of precipitation on various forms of occurrence data was examined during four separate periods: when it does not rain, when it rains lightly, when it rains moderately, and when it rains heavily. The development of these four categories allows for the classification of precipitation in terms of quality. An interquartile technique was used for this, and using the interquartile ranges, four datasets with the four precipitation amounts previously indicated could be created.

It was feasible to deduce from the study of occurrences according to the four precipitation levels that there are two categories of occurrences, namely floods and collapses, that grow as precipitation levels rise. In the case of floods, it is noticeable from Figure 9, the increase in incidence based on precipitation levels is remarkable, since the incidence was 9% when there was light precipitation level, 32% when there was moderate precipitation, 46% when there was heavy rain, and 8% when there were abnormal levels of precipitation.

Heatmaps were created for the six categories of occurrences that most influence buildings in the city of Lisbon, shifting the focus to a study of occurrences from a spatial perspective to verify how occurrences are spread throughout the cities of Lisbon. From Figure 10, it is possible to see, how the Precipitation levels impact certain type of occurrences, where we can see that the highest peak is on flood, preceded by Falling Structures and Collapses.
The regional distribution of collapses and flooding is depicted in Figure 11.
According to the heatmaps presented on Figure 11, collapses, which are the type of event that most affects the city of Lisbon, have a higher concentration of points in the city's central zone, implying that collapses primarily affect parishes in the city's central area, such as Arroios, Santo António, So Vicente, Misericórdia, Campolide, Avenidas Novas, Penha de França, and areas of the Historical Center

Floods, like collapses, have a larger concentration in the city's downtown region, with the exception that this type of event also occurs with a significant frequency in the north western portion of the city, notably the parishes of Benfica and So Domingos de Benfica.

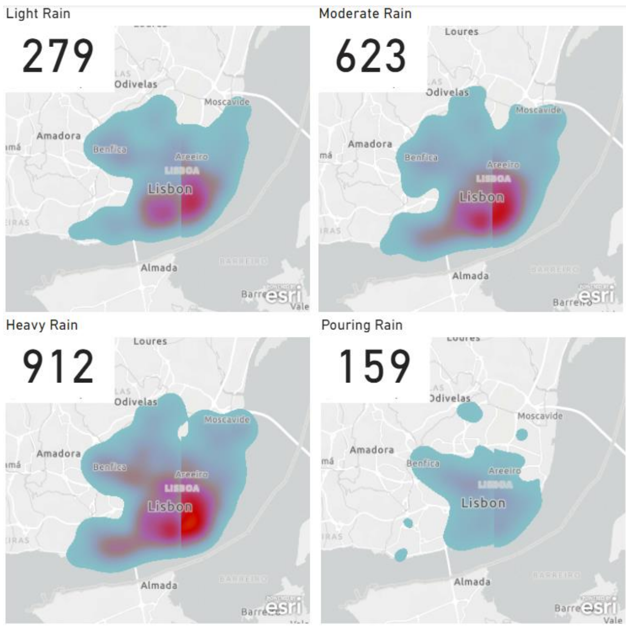

Figure 9. Heatmaps with the 4 types of precipitation and the number of incidents progression

Because there is a concentration of events in a certain location of Lisbon, it was decided to gain a better understanding of the city by examining characteristics such as the condition of conservation of structures and the average age of buildings in different parishes. It is possible to create a relationship between the spatial concentration of occurrences and the condition of the structures through spatial visualization of buildings that are deteriorated or in need of repair, as well as the visualization of parishes where the oldest buildings are situated.

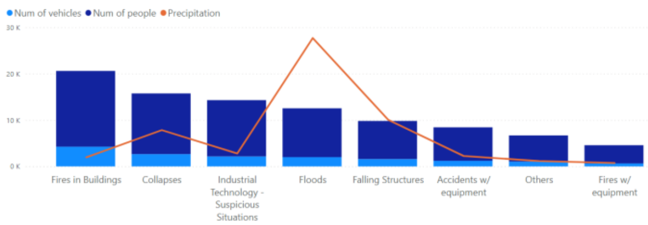

Figure 10. Number of of people and Precipitation by each type of incident

From the conclusions drawn from the two heatmaps shown in Figure 12, it is clear that areas with older buildings and a higher proportion of degraded buildings in need of repair are more vulnerable to events such as collapses, floods, suspicious situations (check smoke or smells), and gas leaks, which occur in greater numbers in these parts of the city.

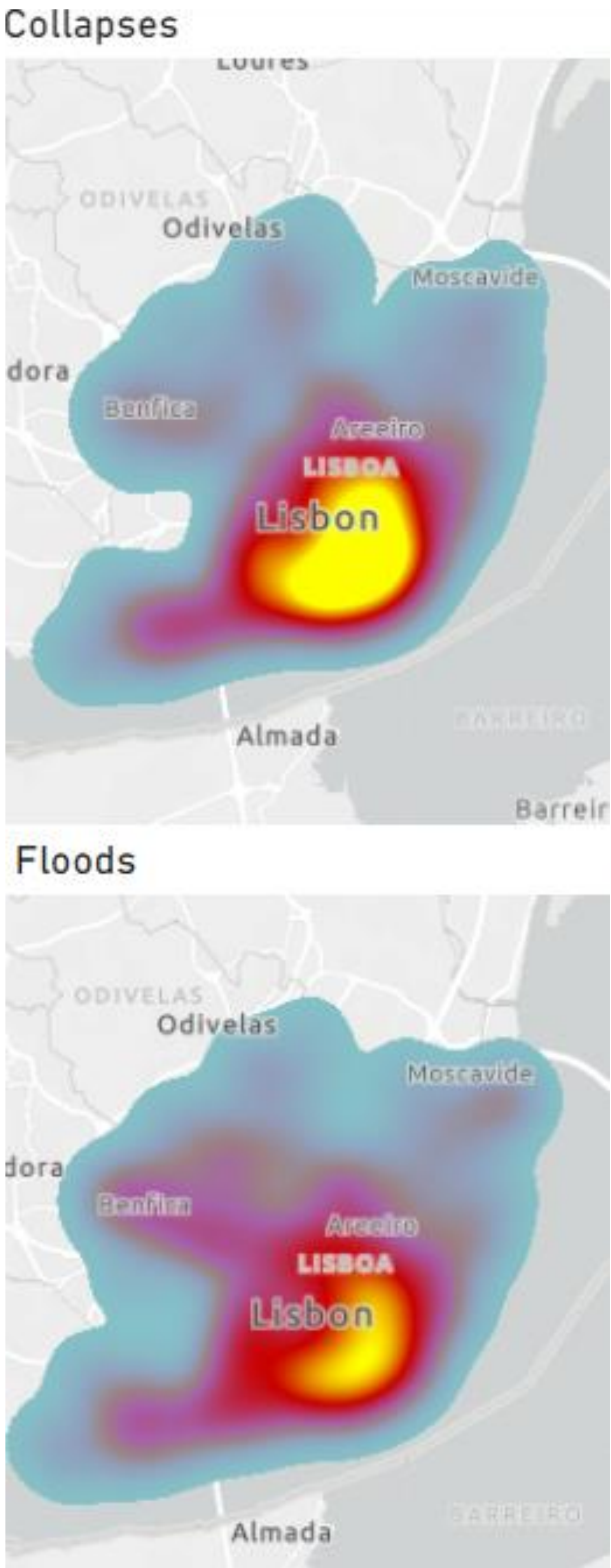

Figure 11. Regional Distribution of Collapses and Floods

On Figure 13 it is also possible to acknowledge that the accidents where the proportion of degraded buildings have more impact are on collapses, followed by Floods, Suspicious Situations (check smoke or smells) and Falling structures

Figure 12. Figure **A** shows the spatial representation of the proportion of buildings that are degraded or in need of major repairs and figure **B** shows the spatial representation of the average age of the buildings per parish

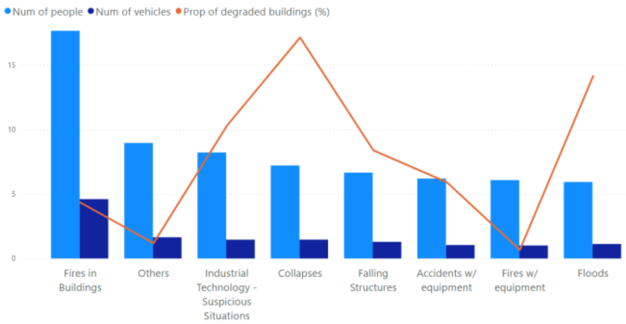

Figure 13. Number of people/vehicles and Proportion of degraded buildings by each type of incident

On Figure 14 we can see almost the same pattern regarding the Age of the building, where we can conclude that the oldest buildings are also the ones in need of major repairs.

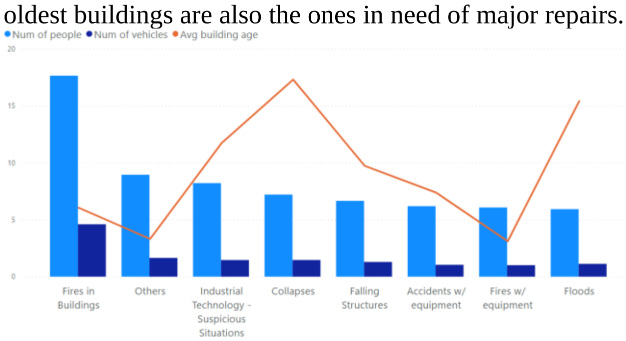

Figure 14. Number of people/vehicles and Proportion of degraded buildings by each type of incident

4. Conclusion

Big Data's importance in disaster management has evolved over time. Nowadays, scientists face one of the most daunting tasks of organizing massive amounts of data collected after catastrophes. Due to the massive volume of data created by disasters, conventional data storage and processing systems are having difficulty meeting the performance, scalability, and availability requirements of big data. We propose a solution for data integration, aggregation, and visualization must be developed effectively while optimizing the decision-making process, as the quality of judgments made by disaster management authorities is dependent on the quality of accessible information

Nowadays, visual analytics dashboards for decision support systems for disaster management are critical, as the frequency of such calamities continues to increase. It is very beneficial to make decisions in a time-sensitive situation using a broad variety of quick data. As a result, it is vital to minimize the overhead associated with data integration and visualization in order to facilitate decision making. Geographical map visualization may be a useful solution in these situations since it enables the extraction, integration, and presentation of disparate data. The purpose of this article is to construct an analytics dashboard for detecting and visualizing risk zones and susceptible locations organized by different accident types in a city. As a result, city management authorities will have more time to prepare and a more detailed strategy for solving disruptive events and relocating resources in a proper manner.

Massive amounts of geographical and temporal data are created in a disruptive event in a city from a variety of sources. Because charts, tables, and static maps have limited exploratory capabilities, it is not feasible to efficiently interpret these large amounts of data. As a result, choices may be postponed. Our approach allows to create a geo-space and time visualization dashboards that can allow management authorities get big picture and prioritizes intervention teams using this knowledge preservation by communicating geospatial data, integrating it with other databases, and creating a dynamic environment that enables quicker decision-making. Geo-visualization enables more interactive maps, such as the ability to explore various levels of the map, zoom in and out, and modify the map's visual look, which is often shown on a computer monitor.

Additionally, risk indices may aid in allocating for example post-flood rescue and relief operations to high-risk zones in terms of shelter placement, central depot establishment, logistics, and evacuation strategy. Map depiction of these essential places using different colour codes may assist in delineating them and ensuring that they get preferential care.

From this study, it is possible to conclude the areas that need more attention, since we can see those events, such as collapses and floods occur mostly in buildings that are very degraded and older. It is also possible to see that these buildings are mainly concentrated in the historic area and downtown of the city, which is also the area where most incidents occur.

The suggested system requires various enhancements as part of future study. To make the dashboard more dynamic and engaging, real-time fluctuations in risk levels within a municipality may be integrated. Additionally, dynamic dashboard visualization may be accomplished using D3.js, JavaScript, CSS, and Bootstrap.

Acknowledgements.

This work was supported by EEA Grants Blue Growth Programme (Call #5). Project PT-INNOVATION-0045 – Fish2Fork.

Abbreviations

DM – Data Mining
AI – Artificial Intelligence
SC – Smart Cities
BD – Big Data
INE – Instituto Nacional de Estatística
IPMA – Instituto Português do Mar e da Atmosfera
GIS - Spatial and Geographical Information System
CRISP-DM - Cross Industry Standard Process for Data Mining

References

[1] J. J. Wellington and P. Ramesh, “Role of Internet of Things in disaster management,” in *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, Mar. 2017, pp. 1–4. doi: 10.1109/ICIIECS.2017.8275928.

[2] C. Yang, G. Su, and J. Chen, “Using big data to enhance crisis response and disaster resilience for a smart city,” in *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)*, Mar. 2017, pp. 504–507. doi: 10.1109/ICBDA.2017.8078684.

[3] A. Zagorecki, D. Johnson, and J. Ristvej, “Data Mining and Machine Learning in the Context of Disaster and Crisis Management,” *Int. J. Emerg. Manag.*, vol. 9, pp. 351–365, Jan. 2013, doi: 10.1504/IJEM.2013.059879.

[4] “(PDF) Crowdsourcing Disaster Response.” https://www.researchgate.net/publication/268448750_Crowdsourcing_Disaster_Response (accessed Aug. 26, 2021).

[5] S. A. Shah, D. Z. Seker, M. M. Rathore, S. Hameed, S. B. Yahia, and D. Draheim, “Towards Disaster Resilient Smart Cities: Can Internet of Things and Big Data Analytics Be the Game Changers?,” *IEEE Access*, vol. 7, pp. 91885–91903, 2019, doi: 10.1109/ACCESS.2019.2928233.

[6] S. Chaudhari, A. Bhagat, N. Tarbani, and M. Pund, “Dynamic Notifications in Smart Cities for Disaster Management,” in *Computational Intelligence in Data Mining*, Singapore, 2019, pp. 177–190. doi: 10.1007/978-981-10-8055-5_17.

[7] S. A. Shah, D. Z. Seker, S. Hameed, and D. Draheim, “The Rising Role of Big Data Analytics and IoT in Disaster Management: Recent Advances, Taxonomy and Prospects,” *IEEE Access*, vol. 7, pp. 54595–54614, 2019, doi: 10.1109/ACCESS.2019.2913340.

[8] T. Li *et al.*, “Data-Driven Techniques in Disaster Information Management,” *ACM Comput. Surv.*, vol. 50, no. 1, p. 1:1-1:45, Mar. 2017, doi: 10.1145/3017678.

[9] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, “Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement,” *Int. J. Surg.*,

vol. 8, no. 5, pp. 336–341, 2010, doi: 10.1016/j.ijsu.2010.02.007.

[10] C. Okoli and K. Schabram, “A Guide to Conducting a Systematic Literature Review of Information Systems Research,” *SSRN Electron. J.*, 2010, doi: 10.2139/ssrn.1954824.

[11] “Scopus - Document search.” https://www.scopus.com/search/form.uri?display=basic&edit.scft=1#basic (accessed Aug. 09, 2021).

[12] “About Google Scholar.” https://scholar.google.com/intl/en/scholar/about.html (accessed Aug. 09, 2021).

[13] M.-C. Jeong and J. Kim, “Prediction and analysis of electrical accidents and risk due to climate change,” *Int. J. Environ. Res. Public. Health*, vol. 16, no. 16, 2019, doi: 10.3390/ijerph16162984.

[14] M. F. Abdullah, M. Ibrahim, and H. Zulkifli, “Big Data Analytics Framework for Natural Disaster Management in Malaysia:,” in *Proceedings of the 2nd International Conference on Internet of Things, Big Data and Security*, Porto, Portugal, 2017, pp. 406–411. doi: 10.5220/0006367204060411.

[15] K. M. Briones-Estébanez and N. F. F. Ebecken, “Occurrence of emergencies and disaster analysis according to precipitation amount,” *Nat. Hazards*, vol. 85, no. 3, pp. 1437–1459, 2017, doi: 10.1007/s11069-016-2635-z.

[16] J. Park *et al.*, “Ensemble model development for the prediction of a disaster index in water treatment systems,” *Water Switz.*, vol. 12, no. 11, pp. 1–19, 2020, doi: 10.3390/w12113195.

[17] S. Saha, S. Shekhar, S. Sadhukhan, and P. Das, “An analytics dashboard visualization for flood decision support system,” *J. Vis.*, vol. 21, no. 2, pp. 295–307, Apr. 2018, doi: 10.1007/s12650-017-0453-3.

[18] R. C. dos Santos Alvalá *et al.*, “Mapping characteristics of at-risk population to disasters in the context of Brazilian early warning system,” *Int. J. Disaster Risk Reduct.*, vol. 41, p. 101326, Dec. 2019, doi: 10.1016/j.ijdrr.2019.101326.

[19] S. Lee, S. Lee, M.-J. Lee, and H.-S. Jung, “Spatial Assessment of Urban Flood Susceptibility Using Data Mining and Geographic Information System (GIS) Tools,” *Sustainability*, vol. 10, no. 3, Art. no. 3, Mar. 2018, doi: 10.3390/su10030648.

[20] Y. Liu, Z. Li, B. Wei, X. Li, and B. Fu, “Seismic vulnerability assessment at urban scale using data mining and GIScience technology: application to Urumqi (China),” *Geomat. Nat. Hazards Risk*, vol. 10, no. 1, pp. 958–985, Jan. 2019, doi: 10.1080/19475705.2018.1524400.

[21] W. Chen, S. Zhang, R. Li, and H. Shahabi, “Performance evaluation of the GIS-based data mining techniques of best-first decision tree, random forest, and naïve Bayes tree for landslide susceptibility modeling,” *Sci. Total Environ.*, vol. 644, pp. 1006–1018, Dec. 2018, doi: 10.1016/j.scitotenv.2018.06.389.

[22] S. Smith *et al.*, “Adoption of data-driven decision making in fire emergency management,” presented at the 24th European Conference on Information Systems, ECIS 2016, 2016.

[23] F. F. Balahadia, B. G. Dadiz, R. R. Ramirez, M. Luvett, J. P. Lalata, and A. C. Lagman, “Application of Data Mining Approach for Profiling Fire Incidents Reports of Bureau of Fire and Protection,” in *2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)*, Dec. 2019, pp. 713–717. doi: 10.1109/ICCIKE47802.2019.9004420.

[24] A. Asgary, A. Ghaffari, and J. Levy, “Spatial and temporal analyses of structural fire incidents and their causes: A case of Toronto, Canada,” *Fire Saf. J.*, vol. 45, no. 1, pp. 44–57, Jan. 2010, doi: 10.1016/j.firesaf.2009.10.002.

[25] X. Liu, Y. Lu, Z. Xia, F. Li, and T. Zhang, “A Data Mining Method for Potential Fire Hazard Analysis of Urban Buildings based on Bayesian Network,” in *Proceedings of the 2nd International Conference on Intelligent Information Processing*, New York, NY, USA, Jul. 2017, pp. 1–6. doi: 10.1145/3144789.3144811.

[26] E. W. Lee, G. Yeoh, M. Cook, and C. Lewis, “Data Mining on Fire Records of New South Wales, Sydney,” *Procedia Eng.*, vol. 71, pp. 328–332, 2014, doi: 10.1016/j.proeng.2014.04.047.

[27] Z. Wang, J. Xu, X. He, and Y. Wang, “Analysis of spatiotemporal influence patterns of toxic gas monitoring concentrations in an urban drainage network based on IoT and GIS,” *Pattern Recognit. Lett.*, vol. 138, pp. 237–246, 2020, doi: 10.1016/j.patrec.2020.07.022.

[28] L. B. Elvas, B. M. Mataloto, A. L. Martins, and J. C. Ferreira, “Disaster Management in Smart Cities,” *Smart Cities*, vol. 4, no. 2, Art. no. 2, Jun. 2021, doi: 10.3390/smartcities4020042.

[29] J. Mesquitela, L. B. Elvas, J. C. Ferreira, and L. Nunes, “Data Analytics Process over Road Accidents Data—A Case Study of Lisbon City,” *ISPRS Int. J. Geo-Inf.*, vol. 11, no. 2, Art. no. 2, Feb. 2022, doi: 10.3390/ijgi11020143.

[30] L. B. Elvas, C. F. Marreiros, J. M. Dinis, M. C. Pereira, A. L. Martins, and J. C. Ferreira, “Data-Driven Approach for Incident Management in a Smart City,” *Appl. Sci.*, vol. 10, no. 22, Art. no. 22, Jan. 2020, doi: 10.3390/app10228281.

[31] “[PDF] Crisp-dm: towards a standard process modell for data mining | Semantic Scholar.” https://www.semanticscholar.org/paper/Crisp-dm%3A-towards-a-standard-process-modell-for-Wirth-Hipp/48b9293cfd4297f855867ca278f7069abc6a9c24 (accessed Aug. 27, 2021).

[32] “Portal do INE.” https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_inst_legislacao&xlang=pt (accessed May 04, 2021).

[33] “IPMA - Serviços.” https://www.ipma.pt/pt/produtoseservicos/index.jsp?page=dados.xml (accessed May 04, 2021).