

MGI

Master Degree Program in
Information Management

A BI Framework to Analyse Wildfires

Understanding the Occurrences and Behaviours of Wildfires in
Portugal Through the Lens of BI

Henrique Cabral de Noronha e Meneses Pires

Master Thesis

presented as partial requirement for obtaining the Master Degree in Information Management

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

A BI Framework to Analyse Wildfires

Understanding the Occurrences and Behaviours of Wildfires in Portugal Through the Lens of
BI

by

Henrique Cabral de Noronha e Meneses Pires

Master Thesis presented as partial requirement for obtaining the Master's degree in
Information Management, with a specialization in Knowledge Management and Business
Intelligence

Supervised by

Bruno Jardim, PhD, NOVA Information Management School

July, 2024

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledged the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Lisbon, 15/07/2024

ABSTRACT

In the context of rising wildfire challenges in Portugal, this study explores the challenges current analytical models have from scientific studies and the lack of indicators integrated in online dashboards when analysing wildfire behaviour and occurrences. Traditional statistical methods and machine learning models often lack in providing a narrative and contextualization for wildfire data, limiting the dissemination of information and user engagement. By leveraging a descriptive approach, this study creates a modern dashboard, empowering users to explore historical fire data from 2001 to 2022. Exploratory features benefit users to explore and assess fire risk in specific areas in Portugal using a wide range of factors. The focus of the study is identifying key performance indicators (KPIs) to explain wildfire occurrences and behaviours, improving the narrative and contextualisation, leveraging a clearer understanding of insights. It contributes to the knowledge within descriptive applications in wildfire analytics indicating and combining critical indicators necessary to understand wildfire behaviours and occurrences in an interactive dashboard.

KEYWORDS

Business Intelligence; Data Visualization; Dashboard; Descriptive Methods; Fire Analytics;
Power BI

Sustainable Development Goals (SDG):



TABLE OF CONTENTS

Statement of Integrity	i
Abstract	ii
List of Figures.....	v
List of Tables.....	vi
List of Abbreviations and Acronyms.....	vii
1. Introduction.....	1
1.1. Research Context.....	1
1.2. Motivation	1
1.3. Objectives and Methodology	2
2. Related Work.....	4
2.1. Evolution of Fire Analytics	4
2.2. Traditional Methods	4
2.3. Descriptive Methods	5
2.4. Diagnostic Methods.....	7
2.5. Machine Learning and Predictive Analytics	8
2.6. Prescriptive Methods	10
2.7. Wildfire Dashboards.....	12
2.8. Gap Analysis.....	15
3. Methodology	18
3.1. Original Data source	19
3.1.1. Gathering Business and Data Requirements.....	19
3.1.2. Collaborative Dimensional Modelling Workshops.....	21
3.2. Four-Step Dimensional Design Process	24
3.2.1. Identify the Business Process	24
3.2.2. Identify the Grain	24
3.2.3. Identify the Dimensions	25
3.2.4. Identify the Fact	25
3.3. Selected indicators from studies explored in related studies.....	25
3.3.1. Wildfire Danger Index (FWI).....	26
3.3.2. Statistical indicators	26
3.3.3. Causes of Wildfires indicators	27
4. Results & Discussion	28
4.1. Dashboard Model.....	28
4.1.1. Calculated Measures	29

4.1.2. Parameters	30
4.2. Dashboard Results	30
4.2.1. Overview	31
4.2.2. Map	33
4.2.3. Geographic	34
4.2.4. Comparisons	36
4.2.5. Temporal	38
4.2.6. Seasons	40
4.2.7. Details	41
4.3. Discussion	43
4.3.1. Comparative study analysis of Indicators	43
Conclusions and Future Works	45
Bibliographical References	46

LIST OF FIGURES

Figure 2.1 – Main page of the Fogos.pt dashboard (Pina, 2018)	12
Figure 2.2 – Fire page of the Global Forest Watch Dashboard (The World Resources Institute, 2009)	13
Figure 2.3 – Main page of the Fire Management Dashboard (Koele, 2023)	14
Figure 2.4 – Dashboard analysing wildfire occurrence from 2013-2015 (Dias, 2018)	15
Figure 3.1 – Diagram shows steps taken in the studies methodology (Kimball et al., 2013)	18
Figure 4.1 – Dimensional model created referring to Star-Schema	28
Figure 4.2 – Field parameters used for filtering by calculated measures helping users to quickly compare different perspectives of data and insights	30
Figure 4.3 – First page of dashboard	31
Figure 4.4 – second page of dashboard	33
Figure 4.5 – third page of dashboard	34
Figure 4.6 – Fourth page of dashboard	36
Figure 4.7 – Fifth page of dashboard	38
Figure 4.8 – sixth page of dashboard	40
Figure 4.9 – Seventh page of dashboard	41

LIST OF TABLES

Table 2.1 – Summary of indicators identified in several studies explored	16
Table 3.1 – Data set description including variables, data type and description	19
Table 4.1 – Shows the calculated measures created in Power BI	29
Table 4.2 – Displays most critical results for each BQ of the first BN	32
Table 4.3 – Displays most critical results for each BQ of the second BN	34
Table 4.4 – Displays most critical results for each BQ for the third BN	35
Table 4.5 - Displays most critical results for each BQ for the fourth BN	37
Table 4.6 – Displays most critical results for each BQ for the fifth BN	39
Table 4.7 – Displays most critical results for each BQ for the sixth BN	41
Table 4.8 – Displays most critical results for each BQ for the seventh BN	42

LIST OF ABBREVIATIONS AND ACRONYMS

ICNF	Conservação da Natureza e das florestas
ISA	Instituto Superior de Agronomia
GNR	National Republic Guard
GIS	Geographic Information System
TFK	Traditional Fire Knowledge
FWI	Fire Weather Index
ISI	Initial Spread Index
BUI	Buildup Index
DC	Drought Code
DMC	Duff Moisture Code
FFMC	Fine Fuel Moisture Code
MaxEnt	Maximum Entropy
ANN	Artificial Neural Networks
UAV	Unmanned Aerial Vehicle
VAM	Visual Analytics Module
DSM	Decision Support Model
FECI	Fire Event Criticality Index
BI	Business Intelligence
SGIF	Associação para o Desenvolvimento da Aerodinâmica Industrial
BN	Business Need
BQ	Business Question
BI	Business Intelligence
Ha	Hectares

1. INTRODUCTION

1.1. RESEARCH CONTEXT

During the past 20 years, wildfires, both natural and human caused phenomena, have been a massive challenge for the Portuguese communities to control. The damage done to ecosystems is significant, especially considering the resources wasted to manage and minimize wildfire. Combating this, Portugal has relied on field experts, local communities, firefighters, and research conducted by institutions and universities (Beighley et al., 2018) to support organizations such as, Instituto da Conservação da Natureza e das florestas (ICNF), Instituto superior de Agronomia (ISA) and several others in the creation of fire management strategies and policies. However, according to the European Commission (2021) Portugal has had difficulties in grasping the big picture, given a scarcity of skilled labour in providing empirical knowledge for decision-making.

Consequently, in June 2017, Portugal suffered the second largest wildfire recorded in history, meeting the death of 64 people, 24,000 hectares of land burned, and more than 500 houses destroyed (Holroyd, 2022). According to Safe Communities Portugal (2017), authorities were slow given the uncontrollable and unpredictable nature of wildfires, and the mobilization of resources such as the National Republic Guard (GNR) was ineffective. Not even an assessment of fire analysts and specialized meteorologists was conducted, despite the historical regularity of wildfires in Portugal.

Moreover, the use of analytical tools and implementation of research labs in Portugal have grown in recent years, and the need for forest fire prediction, prevention and management strategies has also increased. Most prediction models belong to machine learning and prediction models, including supervised and unsupervised learning algorithms (Rajapaksha et al., 2023). A similar study by Bot & Borges (2022), did a systematic review and identified most techniques also deployed complex models, including deep learning and statistical techniques. However, wildfires are extremely challenging to predict and understand, given the contribution of human factors, climate change, weather conditions and fuels (Sun et al., 2023).

The dissemination of wildfire knowledge and information can be challenging between organisations and local communities (European Commission, 2021) and the time to interpret and understand such insights, when dealing with predictive models and statistical techniques, can significantly impact the response time of decision-making (Bot & Borges, 2022). There is a need to identify emerging hotspots in Portugal, considering climate and land conditions, and human activities (Visner et al., 2021).

1.2. MOTIVATION

In the context of rising wildfire challenges in Portugal, this study has conducted research in the field of descriptive, diagnostic, predictive and prescriptive analytics, to explore methods and identify indicators used in Fire Analytics, proposing a modern dashboard. The study goes beyond the utilization of statistics to provide insights into Portugal's fire history. It identifies key performance indicators (KPI) explaining wildfire occurrences, intensity, and severity. The aim of this tool is to

empower users to explore wildfire occurrences since 2001 to 2022, focusing on user interaction rather than traditional statistical methods, and implementing an interactive storytelling in the dashboard. According to Khun et al. (2023) current models have limited data visualizations for fire incidents, lacking in user-friendly information dissemination and in understanding trends. Additionally, several models implemented do not examine long-term trends and occurrences and have incomplete datasets (Visner et al., 2021; Khun et al. 2023; Meira Castro et al., 2023; Bot and Borges, 2021).

Pham et al. (2020) recognises that despite the widespread of applications, societies are still vulnerable to wildfire incidents. They argue that currently no model exists to accurately capture wildfire behaviours and predict occurrences, given the variation of training data in different regions and the interference of human and meteorological factors. Furthermore, Khun et al. (2023) explores that storytelling can improve the narrative and contextualisation, facilitating the interpretation of insights and delivering quicker decision-making. Therefore, the research gap that this study aims to exploit; current analytical models lack narrative and contextualisation of wildfire occurrences and trends. Visualizations are limited, restricting the dissemination of data, and do not leveraging a clear understanding of insights.

In collaboration with the Field Tech Lab, this study develops a dashboard to analyse historical data and facilitate the interpretation of insights, improving response-time and decision-making. The Field Tech Lab is part is a laboratory part of ADAI University in Coimbra. They develop and implement technological solution, and research in the field of forestry robotics, remote sensing, fire engineering, and autonomous systems (Field Tech Lab, 2024). This laboratory has indicated a need to improve their understanding of wildfires in Portugal. They lack in data necessary to develop a historical analysis and to determine the risk of wildfires in different zones, considering multiple factors that influence their behaviour and occurrence. Despite the importance of combating such horrific events, this thesis does not intend to investigate the causes behind wildfires. Instead, it intends to show the value of information through its availability, and impact that appropriate visualization can have in facilitating the interpretation of insights. Also, supporting decision-making to organization such as ICNF, whether in prevention and planning, operation, or prior stage. The objective is also to build a template, including appropriate indicators that can be reproduced by more research in the future.

1.3. OBJECTIVES AND METHODOLOGY

The objective of this research is to develop a systematic analysis and review on current research works that focus on various fire analytics approaches: descriptive, diagnostic, predictive and prescriptive. Subsequently, after understanding what variables explain the occurrence of wildfires, its severity and intensity, with data provided by ICFN, this study proposes a descriptive approach focussed on developing a dashboard using Power BI. This descriptive approach intends to support research and knowledge in the field of wildfire analytics in Portugal, presenting a more visual and explanatory descriptive approach in communicating insights. Also, contributing to wildfire literature and analytics by exploiting a lack of analytical tools used to provide quick insights and knowledge into the evolution of wildfire occurrences in Portugal (Pham et al., 2020; Khun et al.,2023).

- Review literature within the field of wildfire analytics to understanding what indicators and methods are used to explain wildfire behaviour and occurrences.
- Develop a dashboard for Field Tech Lab to analyse the historic evolution of wildfires in Portugal from the period of 2001 to 2022.
- Create a dashboard template with a combination of indicators that can be reproduced by future works.

To ensure that these objectives listed above in section 1.2 are achieved, the next parts of the thesis: **Related Work** explores and evaluates previous wildfire analytics research and projects that were developed and implemented in the field. Thus, the need to identify critical areas within the fire analytics fields that require more research and practical implementation; **Methodology** defines the Business problem, needs and questions, followed by the model creation in Power BI, parameters and calculated measures used to provide the necessary combination of insights presented in visuals; **Results and Discussion** will report and discuss the results achieved from the Power BI solution; **Conclusion** will summarize the results and conclude the overall effectiveness of the dashboard developed throughout the research; and finally the **Limitations and Recommendations for Future Works** addresses the limitations and possible paths to be taken in this research that could be taken in the future.

2. RELATED WORK

2.1. EVOLUTION OF FIRE ANALYTICS

Fire analytics started with traditional methods where researchers manually observed wildfires as they occurred, and post to their occurrence. These findings were then analysed and shared with institutions and universities to better understand their behaviour (Vázquez-Varela et al., 2022). Scientists conducted comprehensive field surveys and visited areas prone to fires. They observed these landscapes, collected data on types of vegetation, weather conditions, and other relevant factors. Post to wildfire occurrences also involved similar methods. As technology began advancing, more sophisticated methods were used, enabling a better monitorization of larger areas at real-time. Satellite capabilities then emerged, where wildfires could then be detected from space. Finally, the integration of Geographic Information System (GIS) even further transformed research of fire analytics (Visner et al., 2021), granting a spatial mapping of wildfire incidents and improving the understanding of the geographical spread of wildfires.

This chapter of the thesis will explore the historical development of fire management analytics from traditional to advanced methods. The related work will cover the evolution and implementation of various analytical methods over time. More specifically, looking at traditional methods to descriptive analytics in analysing historical data, followed by diagnostic analytics to identify root causes of fire incidents. The related work section will then move on to predictive analytics using models and statistical algorithms to predict and anticipate possible fire incidents. Lastly, prescriptive analytics will be explored, given the implementation of the most advanced algorithms and techniques in receiving actionable insights.

The aim of this exploration address gaps identified in existing research related to wildfire analytics and develops a dashboard template with indicators to improve the understanding of wildfire behaviour and occurrences in Portugal. To achieve this aim, research studies using descriptive, diagnostic, predictive and prescriptive approaches will be reviewed to identify these indicators. Methods used in these approaches are also outlined to also assess the success and failures in the field of fire management analytics.

2.2. TRADITIONAL METHODS

Most studies related to traditional methods of fire analytics discuss Traditional Fire Knowledge (TFK) as the focus in the field of global fire ecology, even nowadays. Essentially, exploring historical and contemporary techniques of traditional fire practices on ecosystems and landscapes worldwide. According to the authors Vázquez-Varela et al. (2022), they acknowledge that traditional fire knowledge is typically location-based, detailed, and reasoned. These techniques involved challenges in sharing knowledge due to its dependence in intrinsic knowledge being difficult to transfer, and based on the culture and location the knowledge methods may differ. In other words, accessing the knowledge is extremely difficult given it remains with indigenous and local people of the communities. However, the adoption of these traditional methods also ensures indigenous wisdom and long-term perspectives. Basically, their strong historical understanding, as indigenous knowledge

often spans for generations, and their interconnectedness to the ecosystem, brings a fundamental contextualisation and experience to understanding fire behaviours and occurrences. Indicators used are number of occurrences during prior years, weather and seasonal pattern, and population density. It is important to note that this study has a substantial limitation, that most modern fire analytical studies can outline; ability to measure the accuracy and validity of findings and results.

Additionally, according to Zou et al. (2019), traditional fire and forestry analytics was based on manual observations and methods, which then further developed with field experts, with expertise in mathematics and statistics. Field surveys were conducted, where manual data and observations were collected, focussing on ecological and meteorological factors. Similarly, to what Vázquez-Varela et al. (2022) acknowledged, manual observations from experts and scientists were great for contextualisation. An expert's observation may result in identify factors that are systematically near impossible for computers or systems to identify, given the lack of contextual awareness and model bias. Model bias is when algorithms are unable to capture true relationships and become too familiar with training data and are unable to predict outcomes (Budhathoki, 2023).

On the other hand, Thompson et al. (2019) recognises that the level of data is significantly increasing and that there is a need for rapid responses, that should lead less dependency on traditional methods. However, this is only possible when the quality of the data is present and data models are well designed and managed. The author criticizes that often, agencies significantly lack in data quality and reliability for informed decision-making is, and the presence of cognitive errors and biased decision-making can result in analytic methods as non-practical. Essentially, the limited scientific evidence and decision support for evaluating data and strategies, leads to a reliance on expert judgment. TFK are still commonly used in local communities, and especially in indigenous tribes. Specifically in locations that lack in technology and in interactivity with external societies (Vázquez-Varela et al., 2022).

2.3. DESCRIPTIVE METHODS

Meira Castro et al. (2020) implemented Spatial Temporal and hierarchical agglomerative clustering algorithms; a descriptive approach necessary to gain historical insights. In their analysis of fire incidents in Portugal, the authors have successfully identified human-caused fires by observing various variables such as, fire locations, size, intensity, vegetation types, fuel, and weather conditions. The scope of their research was focussed on understanding, "What happened" and "what is happening", essentially describing that most fires within the period of 1996 and 2015, are a human cause rather than a natural cause. Hierarchical clustering techniques implemented displayed results in heat maps, graphs, and maps. Also, clustering was seen as particularly useful for revealing accurate groups of variables, identifying patterns and similarities based on fire behaviours and occurrences within the dataset, without prior assumptions.

While Meira Castro et al. (2020) conducted primarily utilized hierarchical clustering, spatial temporal and heat maps for exploratory data analysis, for end-users this study considers an explanatory approach. A potential research gap is identified, suggesting a need for more exploratory methodologies to enhance user-interaction and narrative in future studies, where future research efforts may consider incorporating also attempt to include both explanatory and exploratory to

bridge this lack of narrative and contextualization gap. Several visuals were attempted to improve the exploratory experience, such as tables, box plots, hierarchical clustering graph, heatmap and bar charts. However, visuals were used to display clusters and reveal patterns, meaning the interpreting of these cluster insights may be challenging, and generally requires additional context and subject expertise. In Miera Casto et al. (2020) research paper, the following indicators are considered: total number of fires occurrences, human population density and forest area burned.

A similar study from Visner et al. (2021), developed a dashboard including scatter plots, line graphs, scorecards, and maps to enhance the explanatory element of fire analytics. This study considered bushfire frequency, geographical distribution to explain hotspots, population density and persons correlation test, applied to examine the relationship between frequency and size of bushfire. Adding interactivity features and a basic storytelling narrative improves the overall experience of the reader. Correlations and heatmaps were used to exploit relationships between variables to determine causes of fire and their intensity whilst date filters allow readers to compare results over several years. The authors provided a descriptive context to “what is happening” and “what happened”, however, they acknowledged difficulties with exploring the causes of increased bushfire frequency. Attempting to integrate two analytical types within a dashboard, descriptive and diagnostic, is quite challenging. Especially when focussing on an explanatory approach. Even though exploratory dashboards are designed for better data exploration and analysis, they do not necessarily provide speed, when in need of immediate responses and can lack in clarity in contrast to an explanatory approach.

Both Meira Castro et al. (2020) and Visner et al. (2021) highlight challenges associated with the lack of comprehensive information on the causes of forest fires. Analysing fire behaviour and occurrences are extremely unpredictable, and considering data quality challenges, especially when lacking in resources and technology. Considering our study focusses on developing a tool to enhance the understanding of fire behaviours and occurrences, it is found that there is a significant research gap with prior descriptive studies on providing a tool that enhances the contextualisation of fires. Tools and research need to adopt clear connection between visuals and different sections in the dashboard, and it must persuade and enhance the readers curiosity.

One of the main objectives regarding this project, is to develop a user-friendly tool that can facilitate the decision-making of fire management departments. Most descriptive studies discussed in this section have implemented clustering techniques and time series, and so far, all lacked in providing a clear narrative for readers. Referring to another study from Tutmez et al. (2018), significant efforts were made using clustering techniques to analyse what is happening, however, it also lacks in clarity for non-technical users to explore insights. This study focussed more on wildfire behaviours, identifying the different causes and types of fires, classifying them into human influence and natural cases.

Within the scope of this project, improving the understanding is quite subjective depending on the end user. Several of these studies provide appropriate ideas and concepts on aspects that are important to gain insights, and examples of indicators used to describe specific wildfire patterns. Likewise, the end goal of descriptive analytics can usually be for developing predictive and prescriptive models. Hence, this statement supports why most studies discussed in this section can lack in an explanatory analysis. Essentially most focus on exploratory methods to further develop an

analysis using advanced analytical tools. During the next section, diagnostic approaches will be evaluated regarding how studies used these methods to improve understanding on fire behaviours and occurrences.

2.4. DIAGNOSTIC METHODS

Alhammadi et al. (2022) developed a root analysis as they recognise the need for producing knowledge on wildfire behaviours and occurrences, describing the various factors such as weather conditions and human activities leading to an increase in wildfires. The scientific study aims to contribute to wildfires management, proposing a system that addresses key areas that can have an impact on wildfire behaviours. To do so, the authors explore the effect of weather conditions on wildfires and then provide a management plan, reducing the costs associated with firefighting.

Additionally, a root cause analysis was developed to generate a management plan, providing insights into what affects wildfires behaviours and occurrences in Portugal. In this study, Alhammadi et al. (2022) have attempted to implement several explanatory approaches. Visuals are designed to convey a narrative towards the root analysis, supporting the findings and allowing users to explore underlying factors that contribute to wildfire occurrences. While some studies have showcased the applicability of exploratory methods, this study has attempted to implement both. Indicators used are mostly related to environmental factors: temperature, humidity, wind speed, dead fuel moisture, dryness, and fire occurrences. The aim was to study the main important factors that contribute to wildfire initiation and propagation.

Fire occurrences and behaviours often involve complex interactions among various factors, such as environmental conditions, human activities, and landscape features. Diagnosing the exact causes and relationships in such complex systems can be challenging and may introduce a high level of uncertainty. This study has a limited scope and generalization, only attempting to analyse data from five massive fires in California. Furthermore, the root analysis is primarily directed to climate change and weather factors. The study also lacks in Temporal Analysis, where it only focusses on recent years. For understanding fires behaviours, it is vital to collect data that contains a time lapse of many years. Additionally, the visualizations included by the author require some time to analyse and understand its insights. In contrast, the purpose of our study is to facilitate the readers understanding of context, whilst interacting with each visual. Subsequently, visuals must be simple, easy to comprehend, supporting end-users to immediately make conclusions without spending too much time on one visual.

Another study attempted a diagnostic analysis by implementing linear regression instead, to explain relationships between important factors contributing to long-term wildfire occurrences and patterns. Koutsias et al. (2010) leverages crucial indicators necessary to achieve this studies objective in building the self-service dashboard. Koutsias et al. (2010) focusses on socio-economic, demographic, land-cover, agricultural indicators, to explain patterns and trends related to wildfires. These indicators are focussed on forest area, fire occurrences, demographic changes, and density of employees in agricultural areas. This is relevant as it helps understand important factors to include and explore when creating the dashboard. However, like other studies evaluated in the descriptive section, and considering the prior study Alhammadi et al. (2022), there are significant challenges

regarding data quality. More specifically, authors acknowledge temporal discrepancies, where data collection methods differ over time, affecting the comparability of variables over time.

Both Koutsias et al. (2010) and Alhammadi et al. (2022) describe challenges with accuracy of fire occurrence data, and that collecting data from long period of time is extremely challenging when considering several variables. Within the scope of diagnostic approaches in fire analytics, the nature of unpredictability of wildfires and challenges with data quality, makes it difficult to understand why their occurrences happen and what influenced their behaviour. Diagnostic approaches typically require detailed and comprehensive datasets, including information about contributing factors and contextual details. Obtaining such data, especially in real-time or for historical events, can be difficult due to resource constraints, data availability and collection issues.

This following study discusses approaches providing information for decision making by understanding the vulnerabilities and the factors influencing fire occurrences. Badia et al. (2019) conducted several analyses to assessing the magnitude and direction of changes of wildfires and contributing to a comprehensive understanding of landscape dynamics. The author does not specifically state the algorithms that were implemented, however, most results indicate that spatial analysis were the main focal points, given GIS tools were mentioned. These tools were used to collect data on total wildfire occurrences, changes in forest areas, land usage, rate of change, which provides insights into how rapidly land use is transitioning within each study area. Study indicates the importance of number of wildfires, burned areas and percentage of total surface area affected as key indicators to generate insights into wildfire occurrences.

The unpredictability of fire behaviours and occurrences, and the challenges with data quality, makes it difficult to understand the dynamics of wildfires and why certain factors impact their behaviours and occurrences. Badia et al. (2019) is a perfect example of a big limitation within these studies. To get the final picture, readers must explore the entire content, and analyse results from each model developed by authors. However, in most instances under fire emergencies and strategy development, fire management departments do not have time as a resource, and non-technical readers will not follow the narrative of this analysis. Its difficult to disseminate information and there is a lack of clear visuals to display information.

Both Koutsias et al. (2010) and Alhammadi et al. (2022) also display similar limitation as Badia et al. (2019). Diagnostic analytics involves understanding the reasons behind observed patterns or outcomes. This often requires a deeper exploration of data, and analysing the root wildfire behaviours and occurrences can be more complex than simply conducting a descriptive or predictive analysis of what has happened or will happen. Additionally, once the root causes are identified, there may be a need to be re-adjustments algorithms. Typically, this further involves planning and adapting to changes to address the identified issues, which can extend the overall timeline even more.

2.5. MACHINE LEARNING AND PREDICTIVE ANALYTICS

Following most studies and methods applied within the scope of analytics in wildfires, predictive analytics is probably the most explored approach when analysing wildfire behaviours and occurrences. Bot & Borges (2022) made a systematic review of applications in machine learning

techniques for wildfire management on 153 scientific studies and applications between the period of 2019-2022.

They separated fire management into three stages: (i) prevention and preparedness, (ii) detection and response, (iii) restoration and adaptation activities. Within each of these stages different algorithms and methods were deployed, and according to the authors, techniques were then classified based on an area of application, machine learning methods, case study location and performance metrics. Most implemented techniques deployed by studies were random forest, support vector machine, and neural networks, to generate predictions and better interpretability of unexpected events. Bot & Borges (2022) conveyed similar challenges that were also communicated by prior studies. While Bot & Borges acknowledge the potential that predictive analytics and machine learning models can have in forecasting and enhancing management decision support, there is a need to develop faster models and improve the interpretability insights. This author's critique is relevant, given one of the end goals of this thesis is to produce a self-service dashboard to improve the interpretability wildfire behaviours and occurrences in Portugal.

Within the systematic review of Bot & Borges (2022), they identified that most studies considered population density, precipitation, wind speed and direction, humidity, temperature, burned area, topography, vegetation type and density. They also found that most studies used the Canadian Forest Weather Index system. It's the collection of several codes and indexes that produce a final index called Fire Weather Index (FWI), which signifies the potential fire intensity (Fernandes & Palheiro, 2015).

Another study used a different methodology, where Duane et al. (2015) deploys a correlative model approach based on maximum entropy (MaxEnt). MaxEnt is a probability modelling technique that aims to estimate the probability distribution that is consistent with a set of observed data and conditions that the model must satisfy. The study then considers environmental factors for predicting wildfire behaviours and support fire management. The aim was to differentiate fire spread patterns, justifying different severity levels and sizes depending on weather conditions and fuel factors. However, interpretation of results is not exactly linear, and more information is needed to produce spatial and temporal projections. More specifically, as Bot & Borges (2022) indicates, models may be difficult to interpret, and Duane et al. (2015) also further highlights that even if model performance is good, it does not necessarily mean that models have an accurate representation of reality. By examining these studies, important variables that are needed to exploring wildfire behaviours and occurrences can be identified, and how authors choose to communicate or visualize results.

On the other hand, the scientific study from Abid (2021) recognises that artificial neural networks (ANN) provided the most promising results for both stages (i) prevention & preparedness, and (ii) detection & response (Bot & Borges, 2022). Abid (2021) finds that they are particularly useful for pattern recognitions, exploring historical fire incidents, meteorological and vegetation data. ANNs are also particularly good with complex datasets, especially with non-linear problems, which is fundamental when considering the unpredictability of weather patterns and human activities. However, this method is extremely challenging and complicated to develop, especially for fire management to interpret. While these machine learning applications on predictive analytics works provide a coherent understanding of fire behaviours and occurrences, they are difficult to interpret

by non-technical users. So far, predictive analytics has shown challenges regarding how results are communicated. Future studies may want to consider perhaps integrating in predictive analytics, an explanatory approach and display results on a dashboard. Indicators used are very similar to those identified by Bot & Borges (2022) systematic review. Abid considered temperature, humidity, light intensity, CO, heat release rate, smoke and gas rates, temperature, rain, wind speed for burned area prediction, and FWI.

A similar study by Sevinc et al. (2020) developed Bayesian model that only needs to use new data and it generates predictions on forest fires. However, it does not discuss model performance metrics, nor has the author considered data quality measures. Without insights into the data quality practices, it becomes difficult to determine the reliability of the predictions and the extent to which the model can be trusted for real-world applications. A common point found within Bot & Borges (2022), Duane et al. (2015), Abid (2021) and Sevinc et al. (2020), is that in predictive analytics, models are much more susceptible to inaccurate results and misclassification. Whereas descriptive and diagnostic analytics, both include simpler methodologies to follow, and are not so vulnerable to errors. Sevinc et al. (2020) in their study, found that humidity, temperature, wind speed, distance from settlement indicating the proximity of human populations, burn area and tree species were good indicators to predict wildfire occurrences and better understand their behaviour.

Predictive analytics involves complex models that might lead to misclassifications due to data variability, model assumptions, or unforeseen patterns. Instead, our project intends to provide a dashboard using descriptive analytics approaches for fire departments. This should, where it aids in contextualisation, and in support of local knowledge and ecological experience, understanding weather conditions and human activities may become easier to predict.

2.6. PRESCRIPTIVE METHODS

This following study integrated descriptive, predictive, and prescriptive analytics into wildfire management. Thompson et al. (2019) emphasize the need of creating a balance between expert judgment and data analysis to generate appropriate decision support systems. The study discusses the need to incorporate analytics into wildfire management and generate better information to support risk-informed decision-making. Authors do not specifically present a final tool, but instead use a case study of the 2018 Ferguson Fire and showcasing descriptive, predictive, and prescriptive analytics, and how it was used to support the understanding to fire incidents. This study considered proximity of fire to populated areas, water resources, soil erosion, weather and terrain conditions and fuel type crucial indicators to generate appropriate recommendations for improved decision-making.

Essentially for an effective prescriptive approach, both descriptive and predictive analytics should be successfully aligned. The study does not specifically address the algorithms implemented; however, it reviews how to strategically improve managements understanding of wildfire behaviours and occurrences. The authors particularly highlight that the lack of an effective descriptive analytics on wildfires is a bottleneck when developing predictive and prescriptive solutions. Essentially most liabilities regarding accuracy of results are based on descriptive methods. Additionally, Thompson et al. (2019) states that technology such as automating resource tracking and machine learning

algorithms, can bridge the gap in effective data-driven decision-making. However, this concept is quite utopian, given not all geographical locations have the technologies or know-hows available to bridge such gaps. This study also focuses in one case study, meaning that conclusions could be quite biased. It would have been more appropriate and reliable to compare several wildfire cases within different geographical locations that include diverse level of technologies.

Another study introduces the development and application of a visual analytics tool called Fire Event Analytics to manage wildfires. Georgouloupoulos et al. (2016) explore the need of an early fire detection, prevention, monitoring, prediction, and firefighting tool. They have explored numerous technologies to stream data and get early fire detections such as Unmanned Aerial Vehicles and sensors. Fire Event Analytics consists of two parts: (i) Visual Analytics Module (VAM) and (ii) Decision Support Module (DSM). Using the VAM, users grasp the big picture status of areas susceptible to fires based on current factors. DSM then enhances user perception, facilitating decision-making by including fire growth and simulation functionalities. The system develops a Fire Event Criticality Index (FECI), which facilitates the interpretation of several results from visualizations and indications on the VAM.

The authors argue that current and most technologies face difficulties in operational conditions, low reliability, and accuracy of results, costly, and typically require an iterative development and continuous improvement of models. Considering the gap that Thompson et al. (2019) identified, Georgouloupoulos et al. (2016) has taken a step further in by implementing the tool using the Israel Fire & Rescue Services in the real-world scenario. The tool provides both exploratory and explanatory approaches in communicating information. Considering the study's results, the author has successfully implemented an explanatory tool, easy to interpret. Essentially, end-users do not have to be involved in the data analysis process and can simply grasp the final insights to understand and anticipate fire behaviours and occurrences. Also, they found that wildfire occurrences, weather forecasts, land coverage, fire growth estimations, burning index and FECI, which combines temperature, humidity, wind speed, and CO2 into one index.

A similar study by Alix et al. (2018) explored a forest fire modelling and analytics tool called Incinerate that simulates the fire propagation. This model has integrated several factors including slope, elevation, wind speed and distance, and attempt to produce accurate predictions, simulating realistic fire. The aim was to improve the understanding of fire behaviours by modelling various factors together and produce realistic results. In contrast to Thompson et al. (2019) and Georgouloupoulos et al. (2016) that serve as a decision support tool, Alix et al. (2018) have instead developed a learning tool to help fire management departments better understand what fire factors have significant impacts on fire propagations. This study has two critical limitations. Firstly, the tool is purely exploratory, as end users need to spend time to understand and conclude insights after going through several simulations. More studies should be taken to also consider fire occurrences and not just their behaviour and help anticipate locations that are more susceptible to fires. Secondly, the authors identified that there is a significant gap between results in simulations and real-word fire occurrences. There is a need for more data, so factors that stimulate fires generate a better understanding of simulation results.

Finally, Alix et al. (2018) identified that forest density, land slope and elevation, wind presence, percentage of damaged forest overtime, where indicators necessary to generate accurate simulations of fire propagation on slops.

The final section of the related studies chapter will include practical use cases of descriptive approaches in the field of fire analytics. Currently, gaps have been identified in various scientific studies, but a focus on more practical implementations of online wildfire dashboards will be explored. This helps us to identify real-world applications that were successfully deployed, bridging the gap between theory and application.

2.7. WILDFIRE DASHBOARDS

FOGOs.pt is an online website that provides real-time data and insights into wildfires in Portugal. According to Pina (2018), the main part of the dashboard is the map, where users can select various filters that will change the results displayed on the map. For example, there are weather filters, risk of possible fires by date filters, and the type of fire filters. This dashboard is interactive and can help viewers explore the risk of possible wildfire occurrences, given the use of geospatial data. However, the online platform FOGOs.pt does not provide insights into historical patterns. Results are only provided at real-time on the map, and some comparisons are shown in graphs. Referring to figure 2.1, it shows the main indicators used are to explore the risk of wildfire events are temperature, pressure, wind, rain, and clouds.

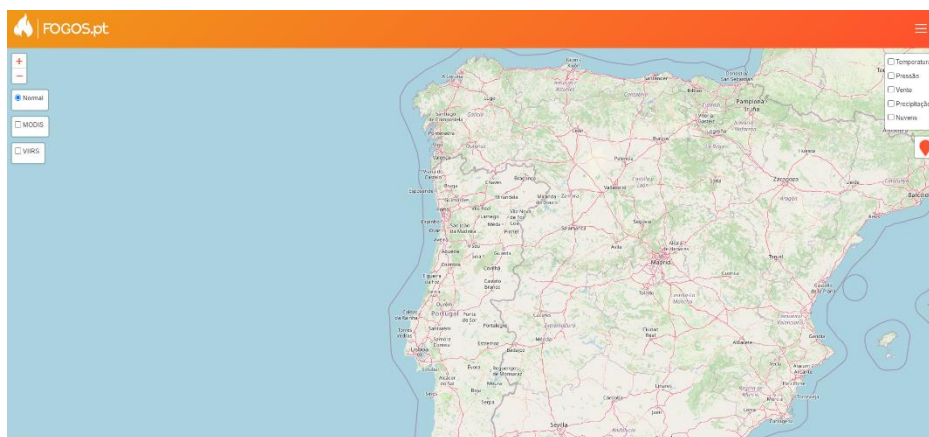


Figure 2.1 – Main page of the Fogos.pt dashboard (Pina, 2018)

In contrast, referring to figure 2.2 shows the dashboard Global Forest Watch deployed by The World Resources Institute (2009), where viewers can select data from various years and make comparisons based on the historical evolution of fires in Portugal. Considering this is an exploratory dashboard; it is interactive with filter and drill-down features. Rather the dashboard creator exercising control over the narrative, and providing predefined insights, users have an open-ended exploration. They have a high degree of control and will require to spend time to draw conclusions.

The dashboard only shows current year data and compares it to prior years. This makes it difficult to conduct a historic analysis and gain insights over the years. The map also only shows a maximum of 3 months of fire data as well and the dashboard is not exclusive to wildfires. It also explores deforestation, climate, land cover, forest change as well, where the primary objective is to determine impacts on deforestation. Considering only the Fire page of the dashboard, it included the following indicators: Loss of trees and number of wildfires. However, indicators from other pages can be cross analysed such as land coverage, deforestation and forest change, fire occurrences, climate changes such as CO2 emissions.

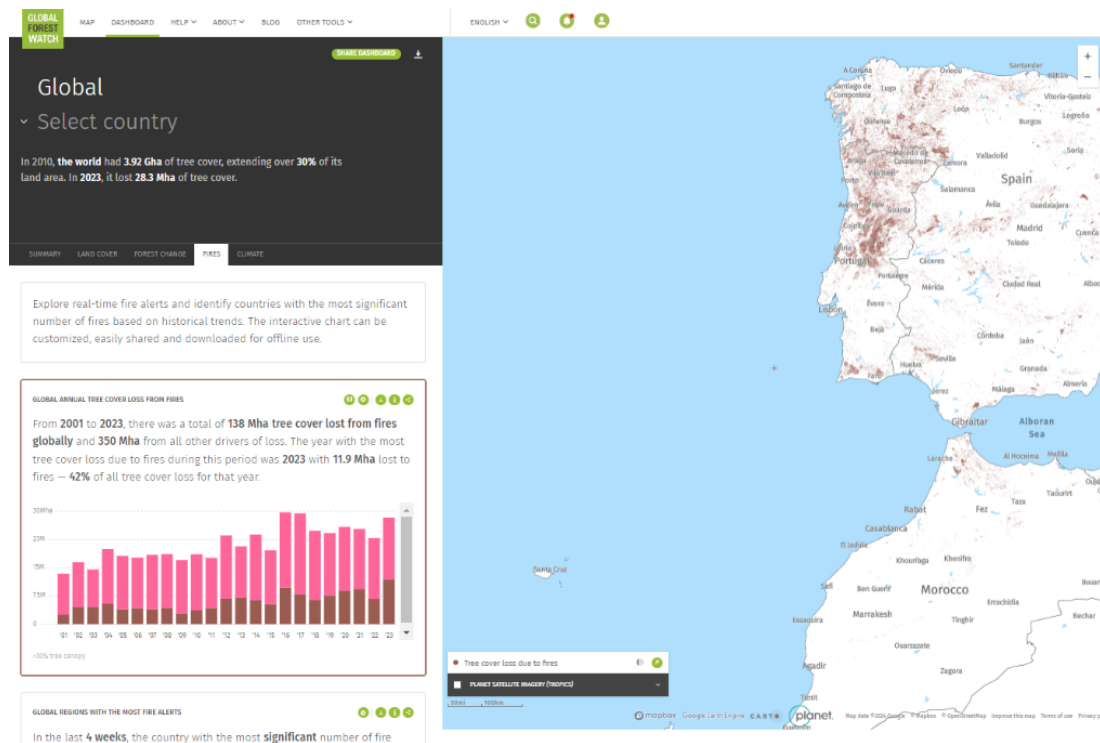


Figure 2.2 – Fire page of the Global Forest Watch Dashboard (The World Resources Institute, 2009)

The Fire Management Dashboard: Wildfire View, developed by Koele (2023) is another exploratory dashboard that provides insights on wildfires in the United States. Figure 2.3 shows the dashboard has several filters also permitting users to completely take control of the narrative and generate their own conclusions. The DNR maps dashboard includes two perspectives, wildfire and burning view. Viewers can understand the causes of fire, and develop a tailor-made analysis, making specific to the end-users needs. The dashboard has included acres burned, fire occurrences and causes as the main indicators for users to explore wildfire behaviours and occurrences in the United States. However, the dashboard is not very user-friendly and will take time to develop conclusions. Graphs are not connected, meaning when you select a filter on one graph, all other visuals on the same page do not update. This is a feature that has been integrated by Dias (2018).

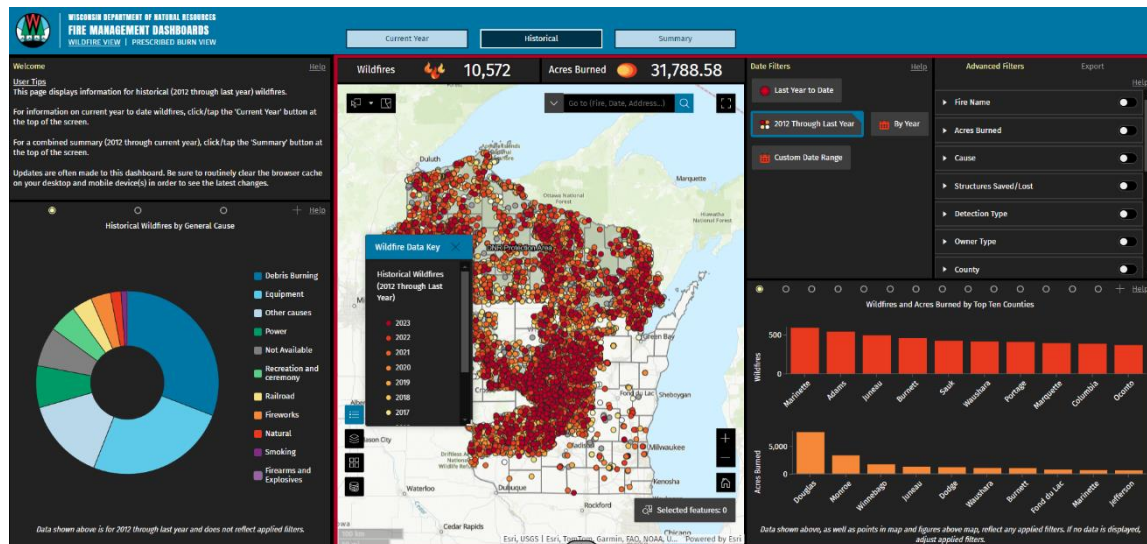


Figure 2.3 – Main page of the Fire Management Dashboard (Koele, 2023)

ICNF also have a dashboard that was developed by Manuel Dias using Power BI, see figure 2.4. According to Dias (2018), the dashboard analyses the occurrences of wildfires and areas burnt in Portugal. In contrast to Fogos.pt, Global Forest Watch and DNRmaps, the dashboard developed by Manuel Dias is a self-service tool, and an explanatory approach was adopted. The storyline is consistent and report pages have filters permitting small user interaction and exploration. This was beneficial as viewers could absorb comparisons without giving too much control in changing the narrative and building clear conclusions. There is a great variety in graphs, comparing to the other dashboards that were reviewed in this section. The author Manuel Dias has focussed on Area burned on fields, agricultural land, and forests, and total number of wildfire occurrences and false alarms as the main indicators of the dashboard. However, the data ranges from 3 years, indicating the impact of wildfires only in 2013-2015 and later expanding the analysis from internal wildfires in Portugal and comparing it internationally. The author familiarized with the data and developed a storyline based on their own interpretation of the data. Dashboard lacks some exploratory factor features, where users can have higher level of interactivity, but selection of visuals does generate a better clarity dissemination of information. A significant limitation identified with this dashboard was the inconsistency with data shown on different pages regarding the number of wildfires, where graphs showed different results.



Figure 2.4 – Dashboard analysing wildfire occurrence from 2013-2015 (Dias, 2018)

After reviewing these dashboards, several indicators can be integrated within our dashboard. Some authors have identified indicators crucial to explore wildfire behaviours and occurrences, and others interesting features to produce a better end-user experience. Indicators explored in scientific studies will also be integrated into the dashboard proposed by this thesis. See below a summary of indicators found in several different analytical approaches, all aimed at providing insight into wildfire behaviours and occurrences.

2.8. GAP ANALYSIS

The dashboard template that this study proposes will provide a faster model development and interpretability of data. Using indicators identified in each of these analytical areas – descriptive, diagnostic, predictive and prescriptive – supports in showcasing a dashboard that improves the understanding of fire activity in Portugal. So far, the related studies chapter has shown that research in fire analytics generally requires complex models making it difficult for non-technical or experts to interpret. Most studies also seemed to have challenges with data quality, where results were not entirely accurate for predictive studies (Sevinc et al., 2020; Koutsias et al., 2010; Alhammadi et al., 2022; Thompson et al., 2019).

Additionally, few research studies in Chapter 3 developed an exploratory narrative for end-users. These studies were identified in sub chapter 3.6. Prescriptive Methods (Georgoulopoulos et al., 2016; Alix et al., 2018), which conducted an exploratory analysis, leaving end-users to explore and

build their own interpretations and conclusions. End-users follow create their own narratives and explore various choices simulation in model selection and parameters during simulations. The other studies in sub chapter 3.3. Descriptive Methods, 3.4. Diagnostic Methods, 3.5. Predictive Methods created their own narrative, leaving end-users to interpret the authors results. Interpretation was concluded by following the steps data scientists and researchers did during data preprocessing, initial model implementations, performance metrics, and model outcomes. However, this study intends to introduce an exploratory dashboard template to improve user interactivity and dissemination of information. It was also vital to identify what indicators were used in different analytical models and help understand conclusions and insights that could be built within the dashboard.

Therefore, as mentioned in prior sections of this work, it will attempt to bridge this gap by building a dashboard with several visualizations and indicators, leveraging a clear understanding of wildfire behaviour and occurrence insights. The purpose is to facilitate the dissemination of information and interpretation, delivering quicker decision-making, and demonstrating that storytelling can create contextualization. End-users will have a historical breakdown of how fires have evolved in Portugal, focused on what happened and what is happening. Rather focussing on predicting wildfires, which requires resources, a comprehensive data analysis and reliable data quality, dashboards can be used to generate faster insights.

Table 2.1 – Summary of indicators identified in several studies explored

Analytical Approach	Author	Indicators
<i>Descriptive</i>	Meira Castro et al. (2020)	• Frequency of wildfires • Human population density • Forest area
<i>Descriptive</i>	Visner et al. (2021)	• Frequency of bushfire • Topographical (hotspots) • Population density • Person's correlation
<i>Descriptive</i>	Tutmez et al. (2018)	• Causes of fires (human influence and natural causes)
<i>Diagnostic</i>	Alhammadi et al. (2022)	• Temperature • Humidity & dryness • Wind speed • Dead fuel moisture • Frequency of wildfires
<i>Diagnostic</i>	Koutsias et al. (2010)	• Forest area • Frequency of wildfires • Demographic changes • Density of employees in agricultural areas
<i>Diagnostic</i>	Badia et al. (2019)	• Frequency of wildfires • Changes in forest areas • Land usage • Burned area • Rate of change and variance
<i>Predictive</i>	Bot & Borges (2022)	• Population density • Precipitation

<i>Analytical Approach</i>	<i>Author</i>	<i>Indicators</i>
		• Wind speed & direction • Temperature & humidity • Burned area • Topography • Vegetation type and density
<i>Predictive</i>	Duane et al. (2015)	• Maximum fire spread • Wildfire size • Weather conditions • Fuel factors
<i>Predictive</i>	Abid (2021)	• Temperature, humidity, light intensity & CO2 • Heat release rate • Smoke & gas rates • Wind speed for burned areas • FWI
<i>Predictive</i>	Sevinc et al. (2020)	• Humidity • Temperature • Wind speed • Distance from fire starting point to nearest settlement • Burned area • Tree species
<i>Prescriptive</i>	Thompson et al. (2019)	• Proximity of fire to populated areas • Water resources • Soil erosion • Weather and terrain conditions • Fuel type
<i>Prescriptive</i>	Georgouloupoulos et al. (2016)	• Frequency of fires • Weather forecasts • Land coverage • Fire Event Critical Index (referred as FWI) • Burning index and fire growth estimations
<i>Prescriptive</i>	Alix et al. (2018)	• Forest density factor • Slope and elevation • Wind presence • Percentage of damaged forest over time
<i>Online Public Dashboard</i>	Fogos.pt (2018)	• Temperature, Pressure, Wind, Rain, Clouds
<i>Online Public Dashboard</i>	The World Resources Institute (2009)	• Land coverage, deforestation & forest change • Frequency of fires • CO2 emissions
<i>Online Public Dashboard</i>	Koele (2023)	• Acres burned • Frequency of wildfires • Causes • Air temperature, humidity, windspeed
<i>Online Public Dashboard</i>	Dias (2018)	• Total and average Burned area (fields, agricultural land, forest) • Frequency of fires & false alarms • Duration of wildfire (Total, average) • Causes of wildfire

3. METHODOLOGY

In the context of fire management response-time and knowledge dissemination is crucial and current analytical models also require time to analyse. This study has taken inspiration from the work of Dias (2018), that was discussed in **Section 2.7 Wildfire Dashboard**, by producing a dashboard with a similar dataset, but with more indicators identified from studies explored **Chapter 3 Related Work**. The indicators that this study adopted from other studies is described in **Section 3.3 Selected Indicators from Studies Explored in Related Studies**. To achieve this outcome and answer the research question - Understanding the Occurrences of Wildfires in Portugal Through the Lens of BI (Business Intelligence) - Kimball's methodology was used as the backbone of the study in the development of the dashboard.

Figure 3.1 depicts an overview of Kimball's methodology, which has a proven track record in current descriptive and BI studies (Makaleng, 2022). Having used this methodology for this study ensured credibility given this study attempted to build on existing BI knowledge. Also, the end-goal of Kimball's four-step process in dimensional modelling when creating a data warehouse is also used in Power BI (Powell, 2018; Kimball et al., 2013), making the methodology relevant to the research design. Not only has it helped build the dimensional model, but it also guided the direction of research in the related studies by finding relevant indicators in various analytical works important for the dashboard.

The four-step dimensional design process of building a dimensional model is as follows: 1) Select the business process, 2) Declare the grain, 3) Identify the dimensions, and 4) Identify the facts. However, before beginning the four-step process, Kimball states that it's necessary to gather business requirements and data realities and undergo collaborative dimensional modelling workshops to build a well-structured dimensional model aligned with the end-users needs. Both these stages are highlighted by Kimball et al. (2013) as fundamental concepts required before attempting to design the Four-Step Dimensional Design Process. Both these initial stages are defined in **section 3.1 Original Data Source** and describe the chosen data collection methods used to build the dimensional model necessary for the dashboard.



Figure 3.1 – Diagram shows steps taken in the studies methodology (Kimball et al., 2013)

3.1. ORIGINAL DATA SOURCE

Two data collection methods were applied to achieve the research objectives. Firstly, business requirements and data realities were outlined by identifying what data was available, and its quality and structure needed to build the dimensional model in Power BI. During this stage, quantitative data was collected, which is explained in **section 3.1.1 Gathering Business and Data Requirements**. Secondly, collaborative workshops were taken to collect qualitative data explained in **section 3.1.2 Collaborative Dimensional Modelling Workshops**, where the needs and objectives were discussed to ensure the four-step model in the dashboard met the end-user needs and analytical requirements. This also provided the basis for creating the dimensions and fact tables.

3.1.1. Gathering Business and Data Requirements

Requirements were gathered from the Field Tech Lab, a laboratory belonging to ADAI (Associação para o Desenvolvimento da Aerodinâmica Industrial), a university in Coimbra. The laboratory is specialised in development of state-of-the-art technology in UAVs, remote sensing, AI, autonomous systems in wildfire and forestry control (Field Tech Lab, 2024) and had a need to understand the evolution of wildfires in Portugal. The expected outcomes that were defined by the Field Tech Lab for the dashboard are as follows:

- Provide an insight into the history of fires in Portugal, with an in-depth data analysis.
- Have a dashboard with capabilities to determine the fire risk in a specific area of Portugal based on multiple factors beyond meteorological indices.

During this phase, to achieve these objectives, a public dataset from ICNF's website was used to develop the dashboard. Data was collected by ICNF's Forest Fire Information Management Software called GIS. This dataset¹ includes 475,448 records, where each record refers to the occurrence of a fire incident. Data ranges from the period of 2001 until 2022 and includes 41 variables. Finally, the dataset includes several categorical variables, statistical information, various indexes and indicators, some obtained from burned area cartography, and definitions of cause categories. Refer to table 2 for more information of the description and structure of variables.

Table 3.1 – Data set description including variables, data type and description

Column	Data Type	Description
<i>Codigo_SGIF</i>	Text	Unique identifier SGIF of the rural fire
<i>Codigo_ANEPC</i>	Text	Identifier of the occurrence in the SADO database of ANEPC
<i>Ano</i>	Text	Year of the fire alert date
<i>Mes</i>	Text	Month of the fire alert date
<i>Dia</i>	Text	Day of the fire alert date

¹ <https://www.icnf.pt/florestas/gfr/gfrgestaoinformacao/estatisticas>

Column	Data Type	Description
<i>Hora</i>	Text	Hour of the fire alert date
<i>AreaPov_ha</i>	Decimal	Burned area in forest (ha)
<i>AreaMato_ha</i>	Decimal	Burned area in shrubland (ha)
<i>AreaAgric_ha</i>	Decimal	Burned area in agricultural zones (ha)
<i>AreaTotal_ha</i>	Decimal	Total burned area (forests + shrubland + agricultural zones)
<i>ClasseArea</i>	Text	Class of total burned area (ha)
<i>DataHoraAlerta</i>	Date/time	Date/time of fire alert
<i>DataHora_PrimeiraIntervencao</i>	Date/time	Data/hora da primeira intervenção
<i>DataGora_Extincao</i>	Date/time	Date/time of the first intervention
<i>Duracao_Hora</i>	Decimal	Duration of the fire (period between the date/time of alert and the date/time of extinction)
<i>IncSup24horas</i>	Text	Rural fire lasting more than 24 hours
<i>DTCCFR</i>	Whole Number	INE code of the parish where the fire started
<i>Distrito</i>	Text	District of the starting point of the fire
<i>Concelho</i>	Text	Municipality of the starting point of the fire
<i>Freguesia</i>	Text	Parish of the starting point of the fire
<i>Local</i>	Text	Location of the starting point of the fire
<i>RNAP</i>	Text	Starting location situated in an area of the National Network of Protected Areas
<i>RNMPF</i>	Decimal	Starting location situated in an area of the National Network of National Forests and Forest Perimeters
<i>X_Military</i>	Decimal	X coordinate of the starting point of the fire in Lisbon Hayford Gauss Datum (EPSG: 20790)
<i>Y_Military</i>	Decimal	Y coordinate of the starting point of the fire in Lisbon Hayford Gauss Datum (EPSG: 20790)
<i>Latitude</i>	Decimal	Latitude of the starting point of the fire (unit: decimal degrees)
<i>Longitude</i>	Decimal	Longitude of the starting point of the fire (unit: decimal degrees)
<i>X_ETRS89</i>	Decimal	X coordinate of the starting point of the fire in the PT-TM06/ETRS89 coordinate system (EPSG: 3763)
<i>Y_ETRS89</i>	Decimal	Y coordinate of the starting point of the fire in the PT-TM06/ETRS89 coordinate system (EPSG: 3763)
<i>DSR</i>	Decimal	Index of rural fire weather danger (calculated based on FWI)
<i>FWI</i>	Decimal	Meteorological index of rural fire danger (click to access explanatory article)
<i>ISI</i>	Decimal	Meteorological index of initial fire spread (click to access explanatory article)
<i>DC</i>	Decimal	Meteorological drought index (click to access explanatory article)
<i>DMC</i>	Decimal	Meteorological index of dead fuel moisture (click to access explanatory article)
<i>FFMC</i>	Decimal	Meteorological index of fine fuel moisture (click to access explanatory article)

<i>Column</i>	<i>Data Type</i>	<i>Description</i>
<i>BUI</i>	Decimal	Meteorological index of available fuel (click to access explanatory article)
<i>CodCausa</i>	Decimal	Code of the fire cause
<i>TipoCausa</i>	Text	Type of cause
<i>GrupoCausa</i>	Text	Group of causes that is based on the type of cause
<i>Descricao</i>	Text	Description of the cause
<i>FonteAlerta</i>	Text	Alert source of fire incident

To ensure the accuracy and quality of data presented by the dashboard, some missing values were treated regarding the definition of causes. Given there are several distinct values of causes and there are 249,624 of missing values, it was decided to replace *null* as *not classified*. Other columns such as *DataHora_PrimeiraIntervencao*, *RNAP* and *RNMPF* were removed given the quantity of missing values was extremely high and columns were not needed. Additionally, *X_Military*, *Y_Military*, *X_ETRS89*, *Y_ETRS89* were removed from the model given they did not provide any purpose to the analysis.

3.1.2. Collaborative Dimensional Modelling Workshops

To understand the end-users needs, workshops were conducted with Carlos Viegas the head of Field Tech Lab and assistant professor at University of Coimbra. These were interactive sessions where Business Needs (BN) were discussed and important Business Questions (BQ) raised, ensuring that the dashboard met the end-users needs. Given the expertise of the Field Tech Lab in wildfires and requirement to understand its evolution in Portugal, collaboration between the subject matter experts and data modelers was critical and prevent the development of the dashboard in isolation. It's important to consider that these questions can be generalized by other locations and are not just limited to Portugal. BQ raised within each BN are formatted to be in "What happened" and "What is happening" given the research question was focused on a descriptive approach. Considering the workshop sessions, the following needs and questions were collected:

BN 1: General Overview summarizing key findings and insights.

It was expressed by the Field Tech Lab that it was necessary to include an overview page. Essentially to indicate key insights that would be further analysed throughout the dashboard were users could start to familiarize themselves with districts that have a high occurrence (hotspots) and low occurrences (coldspots) of wildfires. Also discussed in **Background Chapter 2**, it is imperative that users start to feel enthusiastic about the information they read, and the objective is to get their attention in wanting to further explore.

1. What is the historic evolution of wildfires in Portugal?
2. What are the hotspots and coldspots in Portugal?
3. What is the number of wildfires per Cause?

4. What were the top 3 years presenting the most danger in wildfires in Portugal?
5. What were the highest number of false alarms per year?
6. What was the frequency of wildfire by season?

BN 2: Map showing the geographic evolution of wildfire intensity and severity.

Include a Map to help visualize the occurrences of wildfires throughout Portugal, including several different filters. Each filter should users better understand the frequency of larger wildfire throughout Portugal, during what seasons and districts. Additionally, adding a Cause filter also supports users to understand if causes are specific to locations in the North, Center and South of Portugal.

7. What was the county that had the highest frequency of wildfires by county?
8. What was the severity and intensity of wildfires by month?
9. What was the severity and intensity of wildfires by cause?
10. What was the severity and intensity of wildfires by class area burned?
11. What was the severity and intensity of wildfires in Portugal?
12. What was the location that had the highest frequency of wildfires that burned over 1000 hectares?

BN 3: Determine the geographical areas in Portugal most affected by fires.

Analysing regions in Portugal that have a higher frequency of fire occurrences will help Fire Management departments and organizations to allocate resources more efficiently. This will also help gain context to support the development of fire prevention strategies for regions that have significant levels of fire occurrences. Understanding the difference between occurrences, intensity and duration of fire geographically will also indicate possible hotspots within Portugal over the years.

13. What are the top districts with highest frequency of fires?
14. What was the area burned by district?
15. Are there any emerging trends indicating shifts in fire hotspots within Portugal?
16. Which district has the longest duration of fire incidents?
17. What is the classification of a possible fire occurrence in a district?
18. Which topologies (fields, forests, agricultural land) have the most area burned and how have they evolved over time and by district?

BN 4: Compare perspectives to gain an understanding intensity and severity of wildfires.

Areas that have high occurrences of wildfires does not necessarily mean they also have longer durations, or larger areas burned. Thus, it would be interesting to compare these three indicators and gain an understanding of where they are varying, during what periods and districts. This would further raise questions and exploiting areas that the Field Tech Lab could investigate.

19. What districts had the highest false fire alarms?

20. What districts had the highest max duration?
21. What districts had the highest max area burned?
22. What was the comparison of number, duration and area burned from wildfires by year?
23. What was the impact of wildfires by season geographically?
24. What district had the highest number of wildfires that last more than 24 hours?

BN 5: Analyse the historical patterns of fire incidents across Portugal.

Analysing whether the number of fire occurrences is increasing or decreasing will help organisations interpret the effectiveness of current fire prevention strategies, and allocation of resources. Understanding this need is vital, given if fire occurrences are increasing, it will create more research questions and indicate areas where further investigation is needed. Hence the importance of also analysing the top causes of wildfires and if they were accidental, intentional or a natural phenomenon.

31. What are the top 5 worst years in Portugal that had most fires?
32. Which year had the highest average of area burned and fire duration?
33. What are the top causes of fires by year?
34. Which is the cause that most contributes to fire incidents?
35. Is the area burned and frequency of wildfires increasing or decreasing?
36. Has the average duration of fires, and total area burned increased or decreased?

BN 6: Understand how seasons can impact the behaviour and occurrences of wildfires.

Fire behaviours and occurrences vary depending on seasonal factors. During summer the weather is warmer and naturally increases the risk and possibility of wildfires. However, understand the months of transition indicate when organization and even the public need to be more cautious. For example, this may also support fire prevention such as not being permitted by law to burn fires or any use of fires during certain periods of summer.

37. What was the number of fires per season with more than 24-hour duration?
38. Which month had the highest frequency of wildfires?
39. What time of day had the highest number of fire incident alerts?
40. What season had the highest risk of wildfire during the last 20 years?
41. What month had the highest area burned?
42. What month had the highest duration of wildfire?

BN 7: Analyse the risk of wildfires in Portugal.

To understand the risk of wildfires in Portugal, it's important to explore how weather conditions and fuel moisture levels. These following indexes found within the dataset described in **Section 3.1.1 Gathering Business and Data Requirements** – FWI (Fire Weather Index), ISI (Initial Spread Index), BUI

(Buildup Index), DC (Drought Code), DMC (Duff Moisture Code), and FFMC (Fine Fuel Moisture Code) – provides a different perspective and insight into fire behaviour-based weather conditions and fuel moisture levels. Other indicators such as area burned, duration and number of wildfire occurrences also helps mitigate control over wildfire spread and intensity. This can help anticipate fire behaviour during certain regions and time periods by conducting a historical analysis and understanding the variation of indexes and indicators over time.

43. What districts in Portugal have the highest index rates for fire dangers?
44. What was the fire danger by months and season?
45. What is the fire risk by causes?
46. What month displayed the worst DC, DMC and FFMC value?
47. What month displayed the worst ISI and BUI value?
48. Filtering by 24h was the worst value of FWI in a month?

3.2. FOUR-STEP DIMENSIONAL DESIGN PROCESS

After having gathered business requirements and undergoing collaborative workshop sessions with the Field Tech Lab, the next steps were to implement the Four-Step Dimensional Design Process outlined by Kimball.

3.2.1. Identify the Business Process

Selecting the Business Process is understanding the low-level activity performed by an organization (Kimball et al., 2013). The process is focussed on documenting wildfire events, which is what end-users want to analyse. It's important to understand different perspectives, on the impact that wildfires have had, and to better understand their behaviour and occurrences in Portugal. From various examples retrieved from current dashboards and scientific studies, geographical, temporal, and seasonal analysis are explored, enabling a more profound understanding of wildfire patterns. Subsequently, this approach not only benefits strategies and decision-making, but also supports the future development of more effective wildfire management practices.

3.2.2. Identify the Grain

Declaring the grain means defining the lowest level of detail of a specific row associated with the fact table and understanding exactly what a row represents (Kimball et al., 2013). The granularity of the dashboard was defined during the collaborative workshops to build a model to that captures each individual wildfire event in Portugal from 2001 to 2022 by cause, start-date, end-date, and location. This is based on the most detailed level of the Dim Causes dimension, described in **section 3.2.3**, and start date, end-date and district that were attributed kept in the fact table, rather creating dimensions. The granularity for date and location data within the fact table was set to seconds and location by district, ensuring each incident is uniquely identifiable and can be analysed based on specific characteristics like start date, location, size, duration, and causes. Focusing on the granularity

of single events allows for analysis based on the end-users needs to explore wildfire trends, patterns, and outliers over time.

3.2.3. Identify the Dimensions

Dimensions are single valued and contain descriptive attributes aimed at building context to the numerical data in fact tables. They provide the “who, what, where, when, why, and how” in events (Kimballs et al., 2013). This dimensional model will include three dimensions: Dim_Causes, Dim_Location and Dim_DateHourAlert. Creating this dimension helped improve model performance, given there was less data stored the fact table. Additionally, hierarchies were created within each of the dimensions to enable drill-down and role-up analysis by users regarding causes of wildfires. Essentially, this method defined by Kimball has benefited users to explore data interactively by slicing and dicing across different dimensions, helping to identify trends and patterns, and setting filters to meet the users needs.

3.2.4. Identify the Fact

The fact tables are the measurements resulting from wildfire events and a single row is unique to all other rows in the fact table (Kimballs et al., 2013). The uniqueness of each row was also defined in **section 3.2.2 Identify the grain** and any fact or measurement not deemed important from the collaborative workshop, were not included in this model. Identifying the facts means answering “What is the process measuring?”. The following facts were identified:

- Number of wildfires
- Duration of wildfires
- Total area burned
- Area burned in fields
- Area burned in forest
- Area burned in agricultural land

The output of this study was not to produce a DW, instead the aim is to produce a showcase of a dashboard that includes indicators identified in other analytical studies. Therefore, key steps defined by Kimball were highlighted to justify the structure of dimensional model structure, necessary to build the dashboard.

3.3. SELECTED INDICATORS FROM STUDIES EXPLORED IN RELATED STUDIES

As mentioned before, this study aims at introducing indicators identified in analytical studies discussed in **Chapter 3 Related Studies**. The following indicators that were selected based on the information available from the dataset, and based arguments authors used in their research. These indicators will also meet the business needs identified during the workshops with the Field Tech Lab to create a dashboard that improves the understanding of behaviours and occurrences of wildfires in Portugal. This dashboard will integrate examples of indicators found in descriptive, diagnostic, predictive, prescriptive studies. These indicates were selected, given they are transversal to all

studies and project within this research area; but it does not intend construct predictive outcomes. The focus is to develop a descriptive approach support end-users such as those in the Field Tech Lab on understanding “what happened?” and “what is happening?”.

Not all indicators identified in the research were included because it would clutter the dashboard with too much data. Also, for some indicators it was not possible to collect data, considering the resources and time disposed to this project.

3.3.1. Wildfire Danger Index (FWI)

The FWI is a framework used for assessing the danger of wildfires based on weather conditions, estimating the potential of fire severity in designated areas (Varela, 2018). The FWI system has been researched by many studies (2022; Abid, 2021; Georgouloupoulos et al., 2016; Fernandes & Palheiro, 2015) for several geographical locations and is proven to be effective for calculating the severity and danger of wildfires. It is composed by 6 indicators and according to Fernandes & Palheiro (2015), they are defined as follows:

- **FFMC:** Reflects the moisture content of surface litter and other fuels, indicating how easily they could ignite.
- **DMC:** Represents the moisture content of decomposed organic material below the litter, affecting fire spread.
- **DC:** Indicates the moisture content of deep, compact organic layers, providing a sense of long-term drying effects on fire potential.
- **ISI:** predicts the rate that fire would spread immediately after ignition based on wind speed and the FFMC.
- **BUI:** integrates DMC and DC to estimate the amount of combustible material available.
- **FWI:** integrates ISI and BUI into a single indicator that revealing the potential fire risk.

3.3.2. Statistical indicators

These following indicators are simple statistical calculation that help determine Hotspots:

- **Total Area burned:** indicates the area burned in hectares (ha) including all geographic terrains (Koutsias et al., 2010; Bot & Borges, 2022; Badia et al., 2019; Sevinc et al., 2020; Koele, 2023; Dias, 2018; Duane et al., 2015).
- **Total area burned in fields:** indicates the total area burned just in fields (Dias, 2018).
- **Total area burned in Forests:** indicates the total area burned just in forests (Dias, 2018).
- **Total area burned in Agricultural land:** indicates the total area burned just in agricultural land (Dias, 2018).
- **Total Duration of wildfires:** indicates duration of wildfire immediately after ignition until its put out (Dias, 2018).
- **Average duration of wildfires:** indicates the average duration of wildfires (Dias, 2018).

- **Number of wildfires:** indicates frequency of wildfires and supports identifying hotspots in Portugal (Meira Castro et al., 2020; Visner et al., 2021; Alhammadi et al., 2022; Koutsias et al., 2010; Badia et al. 2019; Georgouloupoulos et al., 2016; The Fire Management Dashboard, 2009; Manuel dias, 2018).
- **Number of False alarms:** indicates how many false fire alerts occurred (Dias, 2018).

These following indicators were not found in any study. However, they were discussed during the collaborative workshops with the Field Tech Lab:

- **Minimum area burned:** indicates the minimum area burned from a wildfire.
- **Maximum area burned:** indicates the minimum area burned from a wildfire.
- **Minimum duration of wildfire:** indicates the minimum duration of a wildfire.
- **Maximum duration of wildfire:** indicates the maximum duration of a wildfire.
- **Average area burned:** indicates the average area burned from all geographic terrains.

3.3.3. Causes of Wildfires indicators

The causes of wildfires are indicators used in the dashboard showing the specific cause and description of the event. Fires are categorized by various causes from both human interference or from a natural phenomenon.

- **Frequency of Cause:** indicates the frequency of wildfires for a specific cause (Koele, 2023; Tutmez et al., 2018; Dias, 2018).

4. RESULTS & DISCUSSION

This chapter discusses the dashboard model section (section 4.1), explaining the output of Kimbal's four-step approach, calculated measures and parameters created. The dashboard results (section 4.2) are then presented, linking each page of the dashboard to a BN, indicating how BQ's were answered, and with what indicators and combination of visualizations. Section 4.3 Indicators and 4.3.1 Comparative study analysis is essentially a benchmark analysis of the indicators included in the dashboard developed by this study compared with online dashboards explored in the Related Studies sub-chapter Wildfire Dashboards (section 2.8). Essentially highlighting the significance of this research in improving the understanding of wildfire behaviours and occurrences in Portugal.

By conducting this research and implementing the chosen methodology, this study was able to build a dashboard with key indicators identified in the studied explored during the Related Studies Chapter. The results section also addressed how the dashboard has integrated an appropriate narrative and visualization to close the research gap that was discussed throughout the study.

4.1. DASHBOARD MODEL

Considering the dimensional model created in Kimball four step approach, a dashboard with many indicators and visualizations, with interactive features was created. This section describes the model developed, including the calculated measures and parameters created to produce the necessary visualization. Figure 4.1 shows the star-schema model that was created for the purpose of this study.

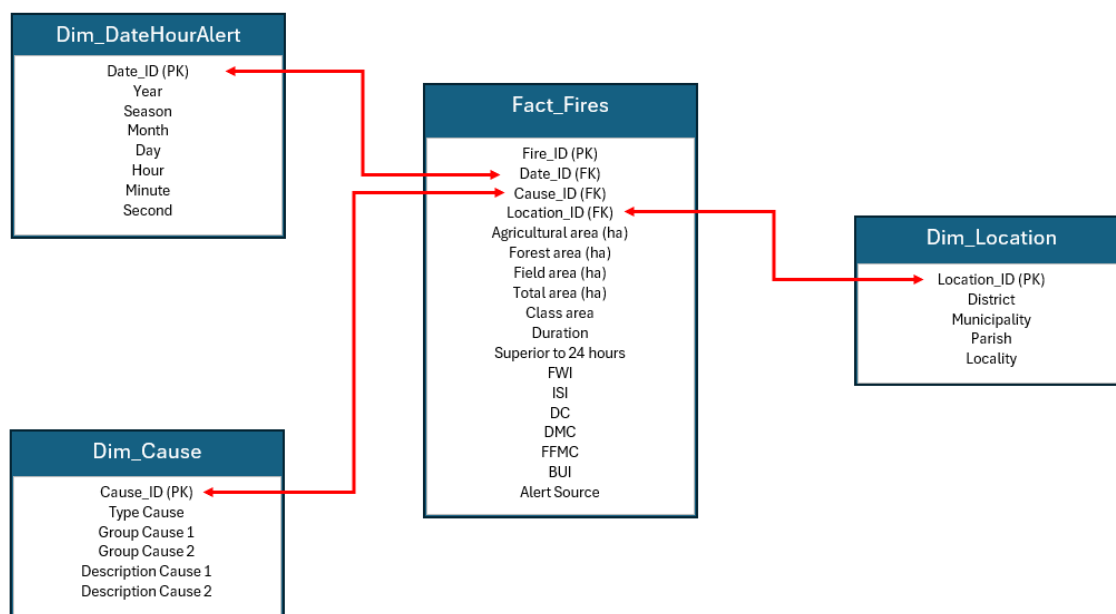


Figure 4.1 – Dimensional model created referring to Star-Schema

4.1.1. Calculated Measures

To ensure the BN and BQ were answers, the following calculated measures were created, which can be found in table 3. These measures were indicators used in the dashboard, necessary to provide a historical analysis, identify trends and align to the BN defined by the Field Tech Lab. These measures were all created in the Measure table, please refer to table 3.

Table 4.1 – Shows the calculated measures created in Power BI

Calculated Measure	Description
<i># Fire</i>	Number of fire incidents
<i># False Alarms</i>	Number of false alarms (no fires)
<i>BUI</i>	Average Buildup Index indicating the available fuel for combustion
<i>BUI Description</i>	Based on the Media BUI index value a definition is provided on the interval
<i>DC</i>	Average Drought Code indicating the dryness
<i>DC description</i>	Based on the Media DC index value a definition is provided on the interval
<i>DMC</i>	Average Duff Moisture Code indicating moisture content
<i>DMC Description</i>	Based on the Media DMC index value a definition is provided on the interval
<i>FFMC</i>	Average Fine Fuel Moisture Code indicting flammability of fuels
<i>FFMC Description</i>	Based on the Media FFMC index value a definition is provided on the interval
<i>FWI</i>	Average Fire Weather Index indicating potential fire spread
<i>FWI Description</i>	Based on the Media FWI index value a definition is provided on the interval
<i>ISI</i>	Average Fire Weather Index indicating potential fire spread
<i>ISI Description</i>	Based on the Media ISI index value a definition is provided on the interval
<i>Max area burned</i>	Max area burned by fires (hectares "ha")
<i>Max duration of fire</i>	Max duration of fire (in hours)
<i>Average area burned</i>	Average area burned by fires (hectares "ha")
<i>Average fire duration</i>	Average duration of wildfire (in hours)
<i>Minimum area burned</i>	Minimum area burned by fires (hectares "ha")
<i>Minimum duration of fire</i>	Minimum duration that fire incidents lasted in hours
<i>Danger of fire (FWI)</i>	Classification of the Fire Weather Index (Low, Medium, High, etc..)
<i>Total Agricultural land burned</i>	Total area burned in agricultural zones by fires (hectares "ha")
<i>Total area burned (ha)</i>	Total area burned by fires (hectares "ha")
<i>Total duration (h)</i>	Total duration of fire incidents in hours
<i>Total fields burned</i>	Total area burned in shrubland by fires (hectares "ha")
<i>Total forest burned</i>	Total area burned in forests by fires (hectares "ha")

4.1.2. Parameters

Three field parameters were created to generate better user visualizations and interpretation of insights. Given the research gap this study identified in current analytical models found visualizations are limited, which restricted the dissemination of data, field parameters improved the interpretation of insights. Including field parameters provided simplicity and clarity, where users can easily select indicators they want, whilst avoiding clutter. This significantly increased user interactivity and report customization, tailoring it to their needs, permitting users to switch perspectives and making quick comparisons and interpretation of results.

Figure 4.2 shows 3 parameters that were created to filter by calculated measures. The dynamic slicer filters by these 3 indicators: number of fire incidents, total duration of wildfires and total area burned. The average perspectives Slicer filter by these 5 indicators: average area burned, average duration of wildfire, max duration of wildfire, max area burned, number of wildfires. Finally, the danger indicators slicer filters by these 6 indicators: BUI, DC, DMC, FFMC, FWI and ISI.

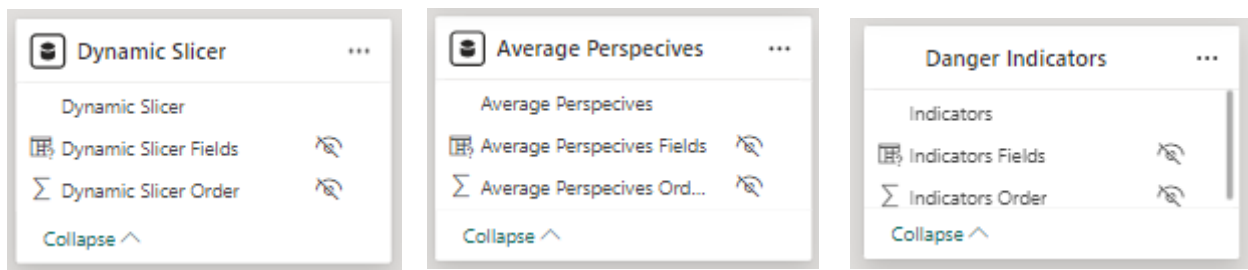


Figure 4.2 – Field parameters used for filtering by calculated measures helping users to quickly compare different perspectives of data and insights

4.2. DASHBOARD RESULTS

The Wildfire Dashboard developed by this study, was one of the research outcomes described, to create narrative and contextualise users on wildfire occurrences and trends in Portugal. It was identified that current online dashboards have limited visualizations, restricting the dissemination of data, and do not leveraging a clear understanding of insights. The objective was to exploit what current dashboards online were missing, using indicators identified in scientific studies under the same field of study.

These indicators were then aligned with the needs identified from the Field Tech Lab to develop a dashboard template, with a combination of indicators, that could then be reproduced by other works in the future. Subsequently, each page of the dashboard answers one of the needs expressed by the Field Tech Lab during the collaborative workshops. BQs are then answered within visualization and by using slicers to take snapshots of data.

This following section will describe the results of the functionalities implemented within each page, describing how it answered the related questions. Indicators were also identified regarding which BQ was answered and through which visualization.

4.2.1. Overview

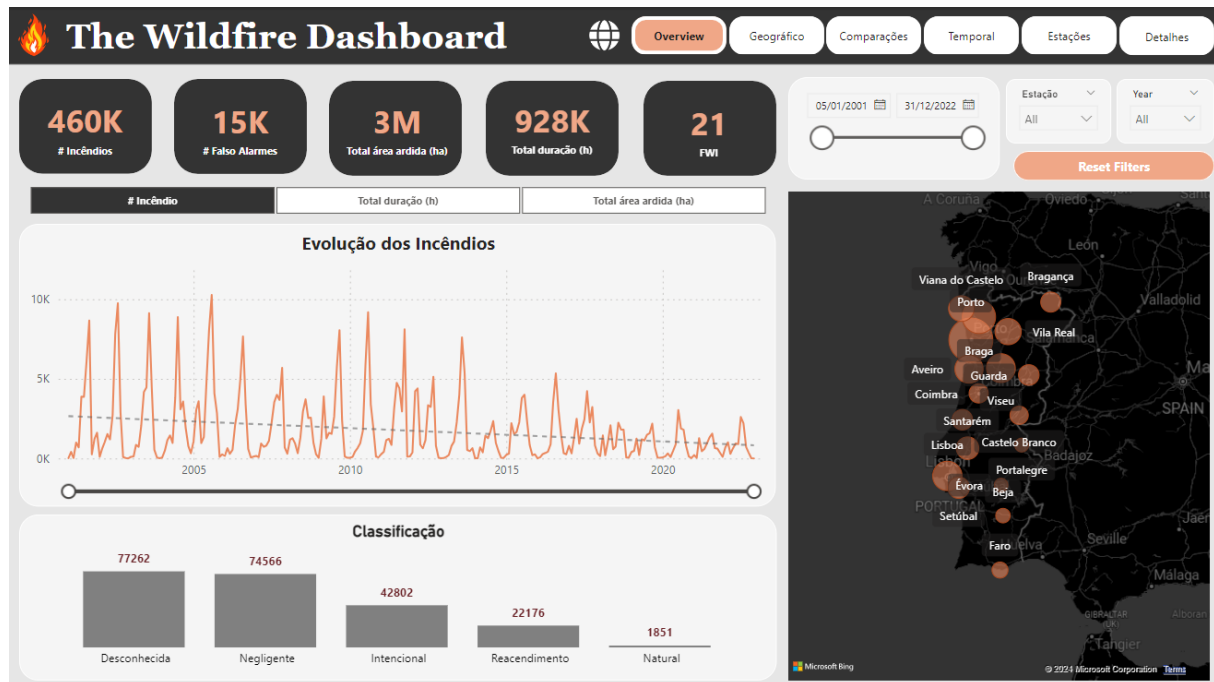


Figure 4.3 – First page of dashboard

Figure 4.3 shows the first page Overview is where users familiarize themselves with the data, quickly grasping on insights such as: number of wildfires, number of false alarms, total area burned, total duration of wildfire and the FWI index. The objective of this page was to get user attention in wanting to further explore the dashboard. Key figures were included to raise curiosity and make users start to feel enthusiastic about the visuals they are going to explore. This is aligned with the first BN: General Overview summarizing key findings and insights.

The dimensional model was also created in Power BI to give users a high level of interactivity, where selecting specific elements of visuals directly affect surrounding visuals and graphs. Considering the BN's defined by the Field Tech Lab, this page considers elements of each BN, particularly to start building the narrative and provide the user with a starting point. This was done by including scorecards with aggregated values, creating a line graph showing the historical evolution of wildfires, integrating a map to visualize wildfires by location and showing the classification of wildfires by cause.

Three perspectives were also included in a slicer – number of wildfires, total duration and total area burned – answering the BQ of whether the frequency and intensity of fires is increasing or decreasing. A similar approach was also used according to the dashboard developed by Dias (2018).

Frequency is related to the number of incidents, and the intensity considers insights grasped from both total duration and total area burned. Selecting on one of these perspectives will affect the entire data displayed on the Overview page. Date filters were also included so users can select years and have snapshots of specific time periods.

Table 4.2 – Displays most critical results for each BQ of the first BN

Business Questions	Result
1. What is the historic evolution of wildfires in Portugal?	Both number and duration of wilders is decreasing, and total area burned has been the same. Overall, over the last 20 years there were 460,000 wildfires in Portugal, 15,000 false alarms, a total area burned of 3,000,000 hectares, total duration of 928,000 hours.
2. What are the hotspots and coldspots in Portugal?	Porto, Castelo Branco, Guarda and Vila Real.
3. What is the number of wildfires per Cause?	Unclassified 77,262; Negligent 74,566; Intentional 42,802; 22,176; Rekindling 22,176; Natural 1851. Most wildfire occurrences, occurred due lack of precaution from people, rather from natural causes.
4. What were the top 3 years presenting the most danger in wildfires in Portugal?	2013 with FWI value of 27, 2006 and 2016 with 26.
5. What were the highest number of false alarms per year?	2011 had 2262 false fire alarms, the highest number of fire alarms.
6. What was the frequency of wildfires by season?	Summer had the highest frequency of wildfires at 239,000, autumn with 123,000, spring with 71,000, and lastly winter with 27,000.

4.2.2. Map

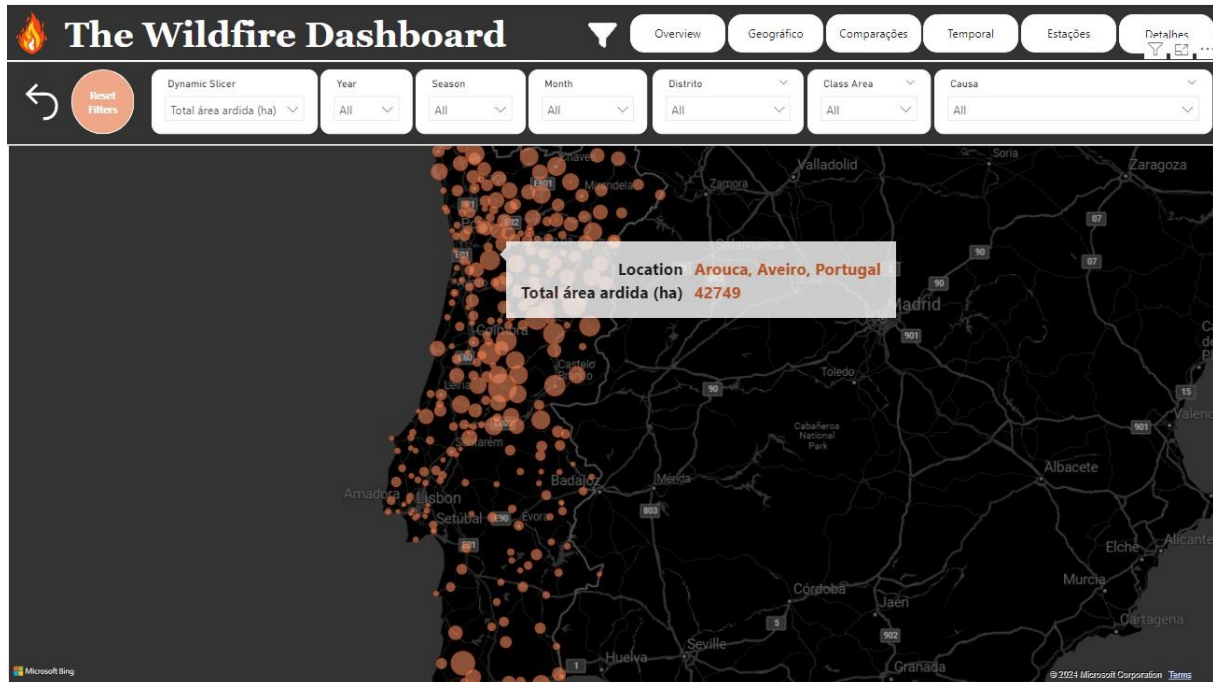


Figure 4.4 – second page of dashboard

The Map page was focussed on the second BN, where a similar design and approach was analysed in Fogos.pt (Forgot.pt, 2018). The objective was to provide users with an interactive map showing the geographic distribution of wildfires in Portugal. Users explore the data purely using slicers to take snapshots of the data presented on the Map. This page only includes a map visual and several slicers where users can customize their analysis.

The dynamic slicer shown in figure 4.4 was used to show 3 different perspectives on the map, where users can instantly change the insights to show the number of wildfires, duration or area burned. Power BI allows users to select slices of information, making this page highly dynamic and interactive.

Fogos.pt also integrated the same type of map visualization, with different indicators than the dashboard developed in this study: temperature, pressure, wind, precipitation, and clouds. However, Fogo.pt does not show the number of wildfires, nor the area burned and duration of wildfires, which makes it difficult for users to gain understanding on the occurrences of wildfires in Portugal. Instead, The Wildfire Dashboard presented by this study used Fogos.pt as a use-case and included indicators focussed on the BN and BQ outlined in the **Methodology Chapter**.

Table 4.3 – Displays most critical results for each BQ of the second BN

Business Question	Results
7. What was the county that had the highest frequency of wildfires by county?	Over the 20 years of data (2001-2022) Paredes in Porto had an overall of 11,672 wildfire occurrences.
8. What was the severity and intensity of wildfires by month?	In general, the north of Portugal had the highest duration and area burned.
9. What was the severity and intensity of wildfires by cause?	Both highest causes unclassified and negligent displayed wildfire presences in the north
10. What was the severity and intensity of wildfires by class area burned?	Classes above 1000 hectares were found in central Portugal, and the remaining in the north.
11. What was the severity and intensity of wildfires in Portugal?	North displayed highest numbers of Number, duration and area burned of wildfires.
12. What was the location that had the highest frequency of wildfires that burned over 1000 hectares?	Over the 20 years of data (2001-2022), Castelo Branco had 13 wildfire occurrences that burned more than 1000 hectares of land. Land included forest, field and agricultural land.

4.2.3. Geographic

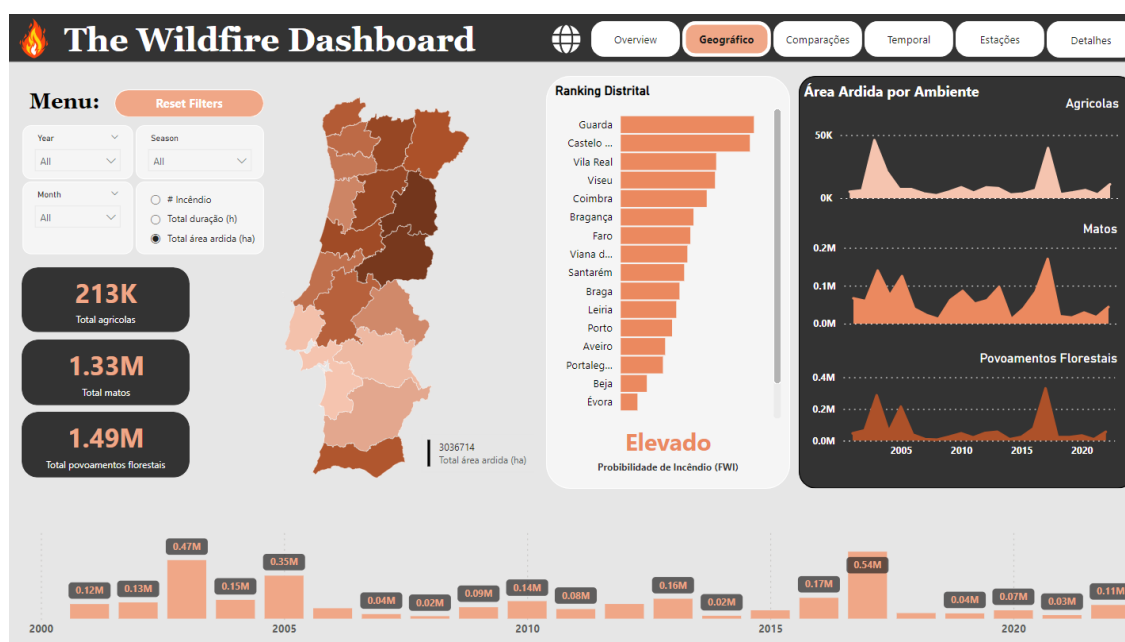


Figure 4.5 – third page of dashboard

In comparison to the Overview and Map page, the Geographic page is significantly more cluttered. Both dashboard from The World Resource Institute (2009) and Koele (2023) were also significantly cluttered, however, users have already absorbed insights from the first two initial pages, meaning they are now simply diving deeper into specific details from a Geographic perspective. The Fire Management Dashboard developed by Koele (2023) had several graphs that were not linked and provided various insights, and The World Resource Institute (2009) was hard to follow, given there was too much text and graphs are very detailed.

On the contrary, the Wildfire Dashboard developed by this study, broke down the dashboard into several pages, where this presented a page only focussed on geographic insights. All visuals are linked, meaning selecting specific elements on visuals directly affect information displayed on surrounding visuals. This was one of the gaps this study identified in prior dashboards, where the focus was to build a clear storyline for users whilst providing exploratory functionalities by having slicers and interactives graphs.

Figure 4.5 shows the Graph selected provide a great variety of visuals that should make it more visually appealing, increasing their need to further explore, but also making insights clearer and more intuitive. Users can use graph to also filter information on the dashboard. For example, selecting the column 2005 on the vertical bar chart filters the page to only show data on all visuals from 2005.

Table 4.4 – Displays most critical results for each BQ for the third BN

<i>Business Question</i>	<i>Results</i>
<i>13. What are the top districts with highest frequency fires?</i>	Porto indicated the highest number of wildfires in Portugal with 98,366, followed by Braga with 52,479 and Lisbon with 39,272.
<i>14. What was the area burned by district?</i>	In contrast to BQ 13 Porto had one of the but it had one of the lowest total areas burned. Total area burned was more critical in Guarda and Castelo Branco.
<i>15. Are there any emerging trends indicating shifts in fire hotspots within Portugal?</i>	There is no clear indication of hotspot shifts, results seem quite unpredictable. However, Guarda and Castelo Branco almost every year were considered one of the highest hotspots.
<i>16. Which district has the longest duration of fire incidents?</i>	Porto indicated the most durations of wildfires, almost every year.
<i>17. What is the classification of a possible fire occurrence in a district?</i>	Considering Porto has the highest number of wildfire occurrences, it displayed an elevated level of wildfire danger. This is seen given other district such as Guarda indicate much higher rates of area burned, where Porto had less.

Business Question

18. Which topologies (fields, forests, agricultural land) have the most area burned and how have they evolved over time and by district?

Results

Forest wildfires are the most critical, followed by fields and finally agricultural lands. Area burned shows to be unpredictable and not follow a specific pattern. Forests had a total of 1.49 million area burned fields had a total of 1.33 million hectares burned, and agricultural had 213 thousand.

4.2.4. Comparisons

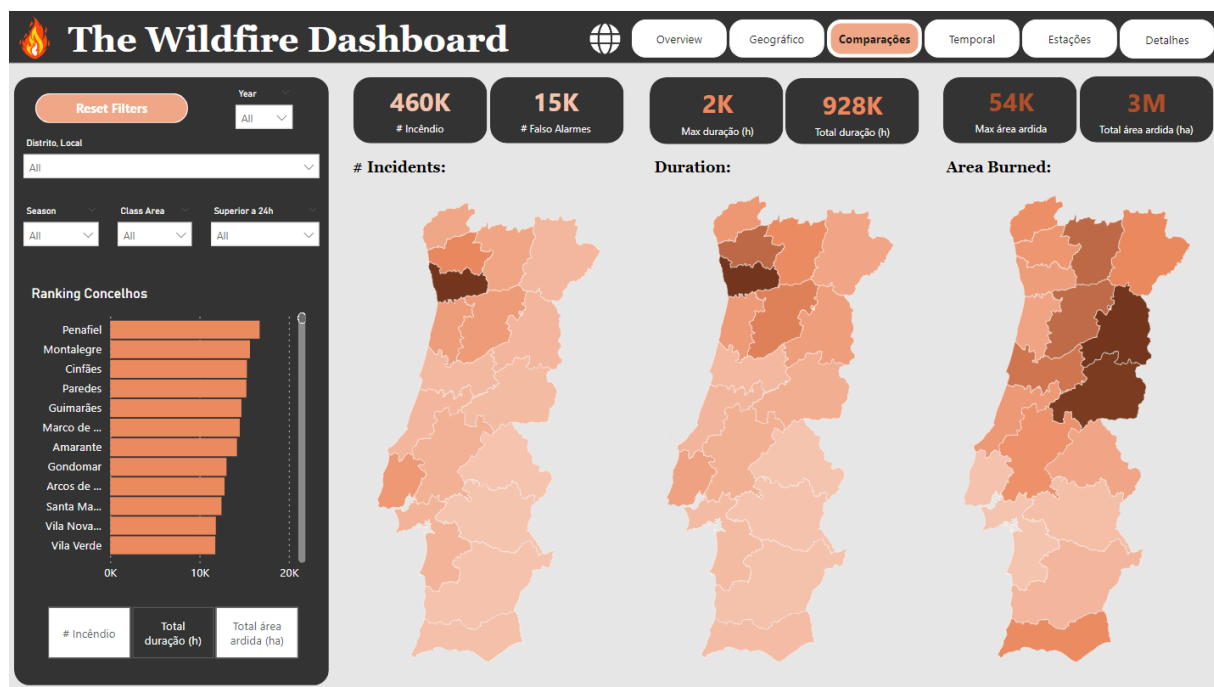


Figure 4.6 – Fourth page of dashboard

Figure 4.6 shows the fourth page shows comparisons on the frequency of wildfires, duration of wildfires, and area burned. One of the business needs described by the Field Tech Lab was to explore how insights would vary between these 3 indicators to improve the understanding of the behaviour of wildfires and their occurrences. This page shows 3 maps, each displaying one of the indicators, showing that not areas that have more wildfire occurrences do not necessarily have higher durations of wildfires and areas burned.

Scorecards at the top have been included to display aggregated values and a municipal ranking has been integrated into a horizontal bar chart. Whenever users click on specific districts on the map, it will be ranking the municipalities, based on the filtering options: number of wildfires, duration of wildfires and area burned.

In terms of design, this page particularly similar to the dashboard developed by Dias (2018), where 3 maps were used alongside a ranking visual. The same structure of displaying 3 maps was kept, given it was a perfect technique of presenting insights without generating clutter. However,

this page adapted to the needs presented by the field tech lab and shows data since 2001 until 2022, whereas Dias (2018) is only focussed on the years 2013-2015.

Table 4.5 - Displays most critical results for each BQ for the fourth BN

<i>Business Question</i>	<i>Results</i>
<i>25. What districts had the highest false fire alarms?</i>	Aveiro had a total of 3399 false fire alarms, Porto came in second with 2476 and Braga with 2429.
<i>26. What districts had the highest max duration?</i>	Both Evora and beja had 2000 hours meaning that the fire with the longest duration came from both locations. To be specific, both locations had a fire that lasted around 83 days of duration during the last 20 years.
<i>27. What districts had the highest max area burned?</i>	The following locations displayed the fires that had burned the most area. First Coimbra with 54k Castelo Branco 34k, Leiria with 30k, and faro with 27k.
<i>28. What was the comparison of number, duration and area burned from wildfires by year?</i>	Over the years the duration of wildfires seems to follow the same pattern as number of occurrences. When the occurrence of wildfire is high in one district it also displays high duration. Area burned however indicates different results and does not follow the same pattern.
<i>29. What was the impact of wildfires by season geographically?</i>	During winter, the north indicates most of the frequency, duration and area burned regarding wildfires and Spring also indicates the same results. However, as the season progresses central and south of Portugal start to display higher levels of area burned.
<i>30. What district had the highest number of wildfires that last more than 24 hours?</i>	Filtering wildfires by 24 hours, Porto is no longer the top district, meaning most fires in Porto are considered small. Viseu had the highest number of wildfires that lasted more than 24 hours and then Guarda.

4.2.5. Temporal

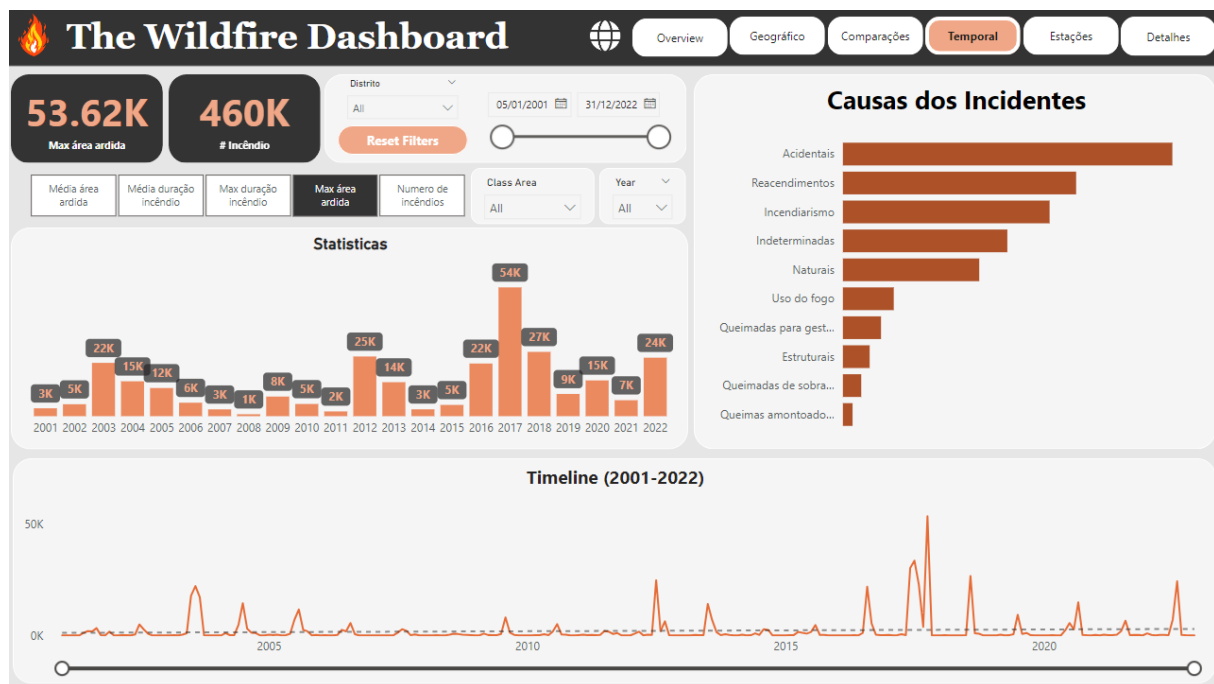


Figure 4.7 – Fifth page of dashboard

Figure 4.7 shows the fifth BN, analysing the historical patterns of fire incidents across Portugal. Determining if the number of fire occurrences is increasing or decreasing, and understand which years and causes had the worst wildfire impacts.

A temporal analysis was found in all dashboards analysed (Fogos.pt, 2018; The World Resources Institute, 2009; Koele, 2023; Dias, 2018). When considering descriptive analytics, a temporal analysis is almost assumed to be present, given the objective is to understand “what happened?” and “what is happening?”.

Like prior pages, the temporal page also includes scorecards. These visuals are excellent at producing instant insights, summarizing aggregated values making it easy for user to conclude and compare data. This page has integrated similar indicators to Dias (2018), however there were slight adaptation to answer the BQ. Dias (2018) showed the total area burned and duration of wildfire, this page focussed instead on the frequency of incidents, and the average and maximum duration and area burned from wildfires. The maximum indicators essentially helped users determine which year and cause had the worst incident in terms of duration and area burned.

The dashboard presented by this study includes data from 2001 to 2022, which is more than the online dashboards analysed, and visually, the design is more user-friendly. Whereas The World Resources Institute (2009) dashboard has integrated artificial intelligence to generate paragraph summarizing data displayed on graphs. On the contrary, the dashboard presented by this study intended on using visuals to minimize clutter and make users determine their own conclusions.

One of the main initiatives this can help users interpret is the level of resources allocated to extinguish fires by year and cause. However, a slicer for selecting districts was also included to identify geographical areas more prone to higher intensity of fires. This was a key need identified in several studies analysed in the Related Studies Chapter that was not expressed by the Field Tech Lab. Subsequently a slicer was included, given it does not occupy space, but if needed can provide snapshots of data.

Table 4.6 – Displays most critical results for each BQ for the fifth BN

<i>Business Question</i>	<i>Results</i>
<i>31. What are the top 5 worst years in Portugal that had most fires?</i>	2005 had a total of 41k wildfires, followed by 2009 with 29k and 2001, 2002 and 2011 had 28k.
<i>32. Which year had the highest average of area burned and fire duration?</i>	2017 had the highest average area burned by wildfire at 26 hectares per wildfire and 2010 and 2011 had the highest average duration at 3 hours per wildfire.
<i>33. What are the top causes of fires by year?</i>	By year most wildfires were undetermined regarding their cause followed by people that deliberately started wildfires. Both these categories were almost always in the top 3.
<i>34. Which is the cause that most contributes to fire incidents?</i>	The cause that most contributed to wildfire occurrences is underdetermined. However, in second it is classified as intentional.
<i>35. Is the area burned and frequency of wildfires increasing or decreasing?</i>	In terms of number of wildfires, they are decreasing but their intensity is the same. Area burned in individual wildlife seems to be growing, where 2017 has the wildfire with the highest area burned, followed by 2018 with the second wildfire at 27k hectares.
<i>36. Has the average duration of fires, and total area burned increased or decreased?</i>	The average duration of wildfire has decreased starting from 2011, which did not surpass 2 hours. Before 2011 there were several years with above 2 hours per wildfire. Regarding average area burned, there is no specific linear pattern and does not follow a clear positive or negative correlation.

4.2.6. Seasons

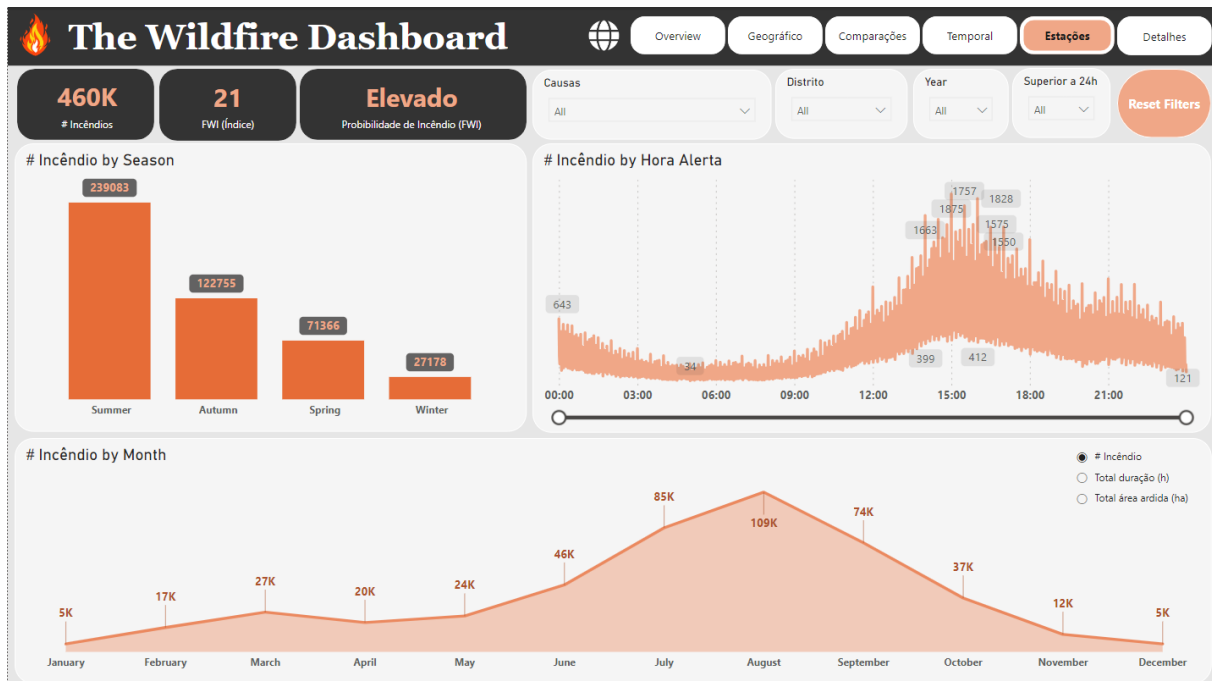


Figure 4.8 – sixth page of dashboard

Figure 4.8 shows the sixth page of the dashboard. The purpose of this page was to generate a season analysis of the impact of wildfire in Portugal. Its common knowledge that the area burned, duration and occurrence of wildfires during the summer is much higher than during winter. However, filters such as Causes and Districts were included to also further the analysis of user when focussing on specific seasons.

No online dashboard analysed in the related studies section (Fogos.pt, 2018; The World Resources Institute, 2009; Koele, 2023; Dias, 2018) included a season analysis. However, based on the BN presented by the Field Tech Lab, it was included to help determine transitional month. These are months where the increase of occurrences, duration and area burned is significant.

This page shows three graphs, showing different perspectives, one focussed solely on the four seasons, the second on the Months and the final visual focussed on the hours during the day. Based on the BNs three indicators were included in a slicer: number of wildfires, total duration of wildfires, total area burned. Selecting on one of these indicators will show that indicator on all visuals.

Table 4.7 – Displays most critical results for each BQ for the sixth BN

Business Question	Results
37. What was the number of fires per season with more than 24h duration?	Summer was in first place with 2377 wildfires that lasted more than 24 hours, with autumn in second with 839, spring with 118 and winter with 27.
38. Which month has the highest frequency of wildfires?	August had the highest frequency of wildfires with 109k, followed by July with 85k and September with 74k.
39. What time of day had the highest number of fire incident alerts?	During the day, the peak of wildfire alerts occurred during 15:00 to 16:00. Within this hour, around 1000 wildfires occurred per min during the last 20 years.
40. What season had the highest risk of wildfire during the last 20 years?	Summer has the highest risk of wildfires in Portugal, where the FWI index was 26 and 239k wildfires occurred during summer in the last 20 years.
41. What month had the highest area burned?	Considering the time between 2001 and 2022, August has been the month with wildfires causing the highest area burned. Total area burned for August is 1.3 million hectares, with July in second by far with 600 thousand.
42. What month had the highest duration of wildfire?	August was the month between the year 2001 and 2022 that had the highest duration of wildfires. A total of 265 thousand hours.

4.2.7. Details

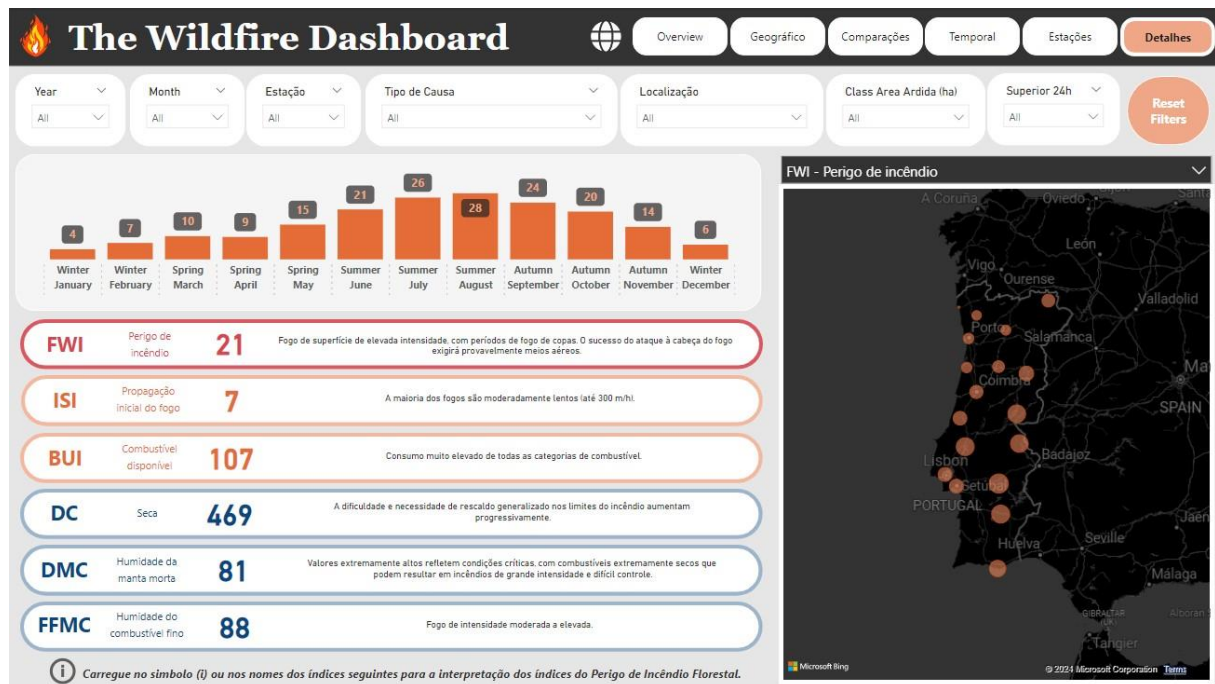


Figure 4.9 – Seventh page of dashboard

The main objective of this page was to analyse the danger of wildfires in Portugal, which was done by exploring weather conditions and fuel moisture levels, considering the following indexes: – FWI, ISI, BUI, DC, DMC and FPMC. These indicators were found on several studies analysed in the Related Studies Chapter (Bot & Borges, 2022; Abid, 2021; Georgouloupoulos et al., 2016; Fernandes & Palheiro, 2015). The online dashboard also analysed (Fogos.pt, 2018; The World Resources Institute, 2009; Koele, 2023; Dias, 2018) did not include these indicators.

Referring to figure 4.9, one of the results of the Details page was the exploratory element provided to the end-user and answering the BN 7. After having analysed and absorbing insights generated from visuals of prior pages, users should have become familiar with the dataset, and understanding the behaviours and occurrences of wildfires in Portugal. So, integrating an exploratory approach on the final page, poses an interesting challenge for users, investigating the impact of wildfires on districts in Portugal; based on the result of indexes.

These indexes were used as indicators to help end-users determine the impact of weather conditions and fuel moisture levels on areas that had high frequency, duration and area burned of wildfires. Users are given the definition of the value of each index and a specific number. Also, by clicking on the acronyms of each index takes users to a detailed study online, describing the class, definition and what each index means in more detail.

It is important to note that these indicators are used to predict possible occurrences and behaviours of wildfires, depending on the value of each index. For this study, they were used instead for a descriptive analysis. Having data for more than 20 years of wildfire occurrences provided a good basis on analysing the historic trends of dangers in Portugal, per district.

Table 4.8 – Displays most critical results for each BQ for the seventh BN

Business Question	Results
43. What districts in Portugal have the highest index rates for fire dangers?	Evora displayed an FWY value of 33 which was the highest followed by Santarem and Beja with 31.
44. What was the fire danger by months and season?	August indicates the month with the highest danger of possible wildfire occurrences, where the value is at 28.
45. What is the fire risk by causes?	Natural and rekindling fires (fires that restarted) displayed the highest FWY values of 24.
46. What month displayed the worst DC, DMC and FPMC value?	The worst month was September with a DC value of 705, DMC of 97 and FPMC of 88. Following august that displayed a DC value of 631, DMC was 121 and FPMC was 89. September was the month considered most dry and lack of humidity in land to prevent wildfires.
47. What month displayed the worst ISI and BUI value?	August displayed the worst ISI value of 7 and BUI of 157 followed by September with ISI value of 7 and BUI value of 137. August had worse conditions of wildfire potentially starting given the available fuel combustion.

Business Question	Results
48. <i>Filtering by 24h was the worst value of FWI in a month?</i>	Filtering by 24 hours duration wildfires, August still maintained as the most dangerous month with possible wildfire occurrence and intensity, with a value of 36 FWI. July came in second with 35 and September with 30.

4.3. DISCUSSION

The Wildfire Dashboard developed by this study, included a total of 20 indicators that were identified during the related studies chapter 3. According to the scientific research explored, its vital to include these following indicators, as the purpose of this research was to exploit a gap that current analytical models lack narrative and contextualisation of wildfire occurrences and trends. Visualizations were limited, restricting the dissemination of data, not leveraging a clear understanding of insights. This section outlines each indicator integrated into the dashboard. Specifying their relevance to each scientific research, discussing its importance to generate insights and improve the understanding of wildfire behaviours and occurrences.

4.3.1. Comparative study analysis of Indicators

This dashboard indicates that there have been 460,000 wildfires in Portugal, 15,000 false alarms, a total area burned of 3,000,000 hectares, total duration of 928,000 hours and a danger index of 21. Porto was the district that suffered most wildfires, but not the most area burned or district with most duration of wildfires. Generally, considering the timeline graph in the Overview page, wildfires seem to be decreasing based on the trend line. However, the area burned seems to be quite inconsistent and significantly affected by specific wildfire occurrences. For example, the timeline graph found in figure 4.3, if filtered by area burned, the top years are 2003, 2005, and 2017, which is also supported by online evidence in newspaper articles (Carvalho, 2017).

According to the results of the dashboard, the final page Details on figure 4.9, indicates that on average throughout the 22 years (2001-2022), Portugal has an FWI value of 21, classified as high danger. This is a mid-level classification, where in summer, value jumps to 27, which is classified as very higher danger. Using the map visualization, with the summer value selected in the seasonal slicer, various regions around Portugal display an extreme value of wildfire danger, with Evora in first at 39. However, further filtering of the data to 2022 years, the general FWI index displays values of 35, with having the highest values in Beja and Bragança at 44. Other studies used the FWI system as metrics to predict wildfires within their models (Georgouloupoulos et al., 2016; Abid 2021), whereas this dashboard had a more visual approach. It's important to note that these indicators are typically used in predictive models, as its purpose is to predict wildfire danger, however they can also be used to track historical dangers of wildfires in Portugal. Interestingly, on the Geographic, figure 4.5, and using the dynamic slicer, an actual perspective can be concluded from the number of wildfires, duration of wildfires and area burned, and compared to what the FWI value was indicating in 2022. Based on these three indicators, neither Beja or Braganca was found on the top 5 of number of wildfires, duration of wildfires, and areas burned.

Figure 4.6 shows the comparisons between 3 indicators, frequency of fires, duration and area burned. Filtering by year 2025 shows that in 2015 there were 19,627 wildfires registered in Portugal, which is slightly more than double that was registered in the dashboard developed by Dias (2018), which was 15,851, refer to figure 2.4. This most likely was the case if different preprocessing methods and Dias (2018) focus on developing an explanatory dashboard, meaning that he has dictated the storyline and tailored the data to present the insights he concluded. One of the potential limitations is that more attention should be allocated to preprocessing. However, according to online sources, 2015 had a total of 18,945 wildfire occurrences (Sampaio, 2022) and other official source from GNR indicate 16,296 from January to August in 2015 (Larguesa, 2015). Figure 4.5 shows the district ranking. Both dashboards, this study and Dias (2018), display the same top 3 districts in 2013 for top area burned: 1st Viseu, 2nd Bragança and 3rd Vila Real. The land type that suffered the most area burned in Portugal are forests, followed by fields, and then agricultural land, also indicated by Dias (2018).

Figure 4.3 indicates the number of false fire alarms, which is displayed on the score card. By selecting the slicer value to 2013, in the overview page it showed a total of 4 false alarms, in 2014 there were 0, and in 2015 there were 16. In contrast to Dias (2018), 2013 showed a total of 3,728 false alarms, in 2014 there were 1,753, and in 2015 there were 2,925 false alarms. These dramatic differences are most likely results related to the dataset and how it was preprocessed. The dashboard developed by this study considered false alarms as any row on the dataset that had an area burned value 0.

According to figure 4.7 from the temporal page, 2017 had the highest area burned of 54,000, which makes sense because during that year Portugal recorded one of the worst fires during its history. In 2011, the longest wildfire was recorded, with a total of 2500 hours. Both indicators max area burned, and max duration identify the worst individual wildfires. The average duration of wildfires seems to be decreasing, where the average drops from 3.0 in 2011 to 1.3 in 2012 and maintains similar value throughout the rest of the years. Average area burned shows some inconsistencies, as does the total area burned, and seems to be affected by specific event of intense wildfires, such as the one recorded in 2017.

Figure 4.8 shows insights from a seasonal perspective, which neither of the dashboards that were discussed in 2.7 Wildfire Dashboards included. Even though its natural that summer has a higher occurrence of wildfires, the objective of this page was to understand transitions between months from Spring to Summer. For example, the area chart shows that the increase in wildfires from June to July is the highest. Additionally, most wildfires occur from 2pm-5pm during summer.

The final indicators implemented in this dashboard indicates that most wildfires recorded are unclassified in Portugal. This could be interpreted as the unpredictability of wildfires, and number of variables that influence their occurrence. However, it was found that in Portugal most wildfire occurrences, occur due lack of precaution from people and because they were intentional. Negligent wildfires stand at a total of 74,566 and intentional wildfires at 42,802, and natural wildfires in 1851 from 2001 to 2022. This indicates that in Portugal wildfires are mostly caused by human interference, rather from natural causes. By clicking on the Classification graph on the overview page, see figure 4.3, the timeline graph intentional wildfires indicate an increase in wildfire trend for negligent and intentional wildfires. This indicates that wildfire occurrences caused by humans is increasing.

CONCLUSIONS AND FUTURE WORKS

With this study, a dashboard was developed considering indicators that were identified in various scientific works, specifically in the field of descriptive, diagnostic, predictive and prescriptive studies. It was found that current analytical models lacked narrative and contextualisation of wildfire occurrences and trends. Visualizations were limited, restricting the dissemination of data, and studies were difficult to follow for users that lacked technical skills.

The related studies highlighted that most studies had challenges with data quality, where results were not entirely accurate for predictive studies (Sevinc et al., 2020; Koutsias et al., 2010; Alhammedi et al., 2022; Thompson et al., 2019) and that research in fire analytics generally required complex models to develop, consuming time, also making it difficult for non-technical or experts to interpret.

Furthermore, The Field Tech Lab provided guidance on important aspects and needs that a dashboard should include when analysing the historical evolution of wildfire in Portugal, improving the understanding of wildfire behaviour and occurrences in Portugal. Subsequently, a public dataset was found providing data necessary to develop the dashboard template with a combination of indicators identified in each of the analytical areas explored in the Related Studies Chapter 2.

As a results, a dashboard template was created, using Kimball's methodology Four-Step Dimensional Design Process, where 20 indicators were included to answer seven business needs defined by The Field Tech Lab.

Despite the insights and results generated, this study has limitations, including its reliance on the public dataset from ICNF. One of the objectives was to produce a dashboard template including indicators that could be further developed by future works. However, not all studies will have access to the same dataset, limiting the reproducibility of this methodology. Studies can follow the same business needs described in section 3.1.2. Collaborative Dimensional Modelling Workshops but will require a different dataset.

Future studies may also want to integrate other indicators not just related to wildfires but also to population sizes, weather conditions and ecological factors. This can help indicate other variables that influence the behaviour of wildfires and its occurrence. The Dashboard developed by this study was limited by time and resources, meaning it was only possible to use the public dataset from ICNF. However, including other indicators will integrate more perspectives into the dashboard, providing more context and areas of exploration. This study was able to collect various indicators that described wildfires occurrences, but it was challenging to collect wildfire behaviours, given various studies that were analysed used different indicators to explain behaviour. To improve the level of insights, future studies may require collecting a greater variety of indicators to explain wildfire behaviour better.

This study focused on integrating exploratory features, such as including filters and interactive visuals, and introducing indicators from identified as important from other scientific studies. Future studies may want to focus more on the data visualization aspect of storytelling. The primary objective of this study was introducing indicators and exploratory features that facilitated the dissemination of information. Finally, future studies may want to explore a model that can stream data in real time.

BIBLIOGRAPHICAL REFERENCES

- Vázquez-Varela, C., Martínez-Navarro, J. M., & Abad-González, L. (2022). Traditional Fire Knowledge: A Thematic Synthesis Approach. *Fire*, 5(2), 1-36. <https://doi.org/10.3390/fire5020047>
- Zou, W., Jing, W., Chen, G., Lu, Y., & Song, H. (2019). A Survey of Big Data Analytics for Smart Forestry. *IEEE Access*, 7, 46621–46636. <https://doi.org/10.1109/ACCESS.2019.2907999>
- Thompson, M. P., Wei, Y., Calkin, D. E., O'Connor, C. D., Dunn, C. J., Anderson, N. M., & Hogland, J. S. (2019). Risk Management and Analytics in Wildfire Response. *Current Forestry Reports*, 5(4), 226–239. <https://doi.org/10.1007/s40725-019-00101-7>
- Meira Castro, A. C., Nunes, A., Sousa, A., & Lourenço, L. (2020). Mapping the Causes of Forest Fires in Portugal by Clustering Analysis. *Geosciences*, 10(2), 53. <https://doi.org/10.3390/geosciences10020053>
- Visner, M., Shirowzhan, S., & Pettit, C. (2021). Spatial Analysis, Interactive Visualisation and GIS-Based Dashboard for Monitoring Spatio-Temporal Changes of Hotspots of Bushfires over 100 Years in New South Wales, Australia. *Buildings*, 11(2), 37. <https://doi.org/10.3390/buildings11020037>
- Tutmez, B., Ozdogan, M. G., & Boran, A. (2018). Mapping forest fires by nonparametric clustering analysis. *Journal of Forestry Research*, 29(1), 177–185. <https://doi.org/10.1007/s11676-017-0417-4>
- Alhammadi, H., Alsereidi, A., Aldhanhani, M., & Talib Amer, Dr. S. (2022). Contribution of Weather Conditions on the Spread of Wildfires with Management System for Preparedness Enhancement. *5th European International Conference on Industrial Engineering and Operations Management*, 2122–2134. <https://doi.org/10.46254/EU05.20220408>
- Koutsias, N., Martínez-Fernández, J., & Allgöwer, B. (2010). Do Factors Causing Wildfires Vary in Space? Evidence from Geographically Weighted Regression. *GIScience & Remote Sensing*, 47(2), 221–240. <https://doi.org/10.2747/1548-1603.47.2.221>
- Badia, A., Pallares-Barbera, M., Valldeperas, N., & Gisbert, M. (2019). Wildfires in the wildland-urban interface in Catalonia: Vulnerability analysis based on land use and land cover change. *Science of The Total Environment*, 673, 184–196. <https://doi.org/10.1016/j.scitotenv.2019.04.012>
- Bot, K., & Borges, J. G. (2022). A Systematic Review of Applications of Machine Learning Techniques for Wildfire Management Decision Support. *Inventions*, 7(1), 15. <https://doi.org/10.3390/inventions7010015>
- Duane, A., Piqué, M., Castellnou, M., & Brotons, L. (2015). Predictive modelling of fire occurrences from different fire spread patterns in Mediterranean landscapes. *International Journal of Wildland Fire*, 24(3), 407. <https://doi.org/10.1071/WF14040>
- Abid, F. (2021). A Survey of Machine Learning Algorithms Based Forest Fires Prediction and Detection Systems. *Fire Technology*, 57(2), 559–590. <https://doi.org/10.1007/s10694-020-01056-z>

Sevinc, V., Kucuk, O., & Goltas, M. (2020). A Bayesian network model for prediction and analysis of possible forest fire causes. *Forest Ecology and Management*, 457, 117723. <https://doi.org/10.1016/j.foreco.2019.117723>

Thompson, M. P., Wei, Y., Calkin, D. E., O'Connor, C. D., Dunn, C. J., Anderson, N. M., & Hogland, J. S. (2019). Risk Management and Analytics in Wildfire Response. *Current Forestry Reports*, 5(4), 226–239. <https://doi.org/10.1007/s40725-019-00101-7>

Georgouloupoulos, N., Mariolis, I., Rogotis, S., Ioannidis, D., & Tzovaras, D. (2016). Fire Event Analytics: A Visual Analytics Tool to Assist Wildland Fire Management. In E. Banissi, M. Sarfraz & M. Fakir (Eds.), *13th International Conference Computer Graphics, Imaging and Visualization (CGiV)*, 89 – 294. Conference Publishing Services. <https://doi.org/10.1109/CGiV39034.2016>

Alix, G., Li, J., & Yu, B. (2018). *Forest Fire Modelling and Analytics: A Comprehensive Study*. Electrical Engineering and Computer Science, York University, 1-12. <https://www.eecs.yorku.ca/~papaggel/courses/eecs4414/docs/project/sample-project/final-report.pdf>

Pina, J. (2018) Fogos.pt. <https://fogos.pt/>

The World Resources Institute. (2009). Global Forest Watch. <https://www.globalforestwatch.org/dashboards/global/>

Field Tech Lab. (2024). ADAI Field Tech Lab. <https://ceif.adai.pt/>

Koele, C. (2023). Fire Management Dashboards: Wildfire View. https://dnrm.wi.gov/WAB/WildfireOccurrence_Dashboard/

Dias, M. (2018). Análise de Ocorrências de Incêndios e Áreas Ardidadas. <https://app.powerbi.com/view?r=eyJrIjojODAwNmY5MDItNGIzZS00NTIxLWl5MjAtNzAyNzk0OWE4NDQzliwidCI6ImU0YmQ2OWZmLWU2ZjctNGMyZS1iMjQ3LTQxYjU0YmEyNDkwZSIsImMiOiJh9>

Beighley, M, Hyde, A.C. (2018) Portugal Wildfire Management in a New Era Assessing Fire Risks, Resources and Reforms. 1-52. https://www.isa.ulisboa.pt/files/cef/pub/articles/2018-04/2018_Portugal_Wildfire_Management_in_a_New_Era_Engish.pdf

European Commission. (2021). Supporting prevention of wildfires in Portugal. Reform Support Team. https://reform-support.ec.europa.eu/what-we-do/green-transition/supporting-prevention-wildfires-portugal_en

Sun, J., Qi, W., Huang, Y., Xu, C., & Yang, W. (2023). Facing the Wildfire Spread Risk Challenge: Where Are We Now and Where Are We Going? *Fire*, 6(6), Article 6. <https://doi.org/10.3390/fire6060228>

Rajapaksha, R.M.M.S.M. & Ariyawansa, Renuka & Jayakody, Dush & Basnayake, Benedict. (2023). Forest Fire Prediction Models in Portugal: A Comprehensive Review. *3rd International Conference of Sri Lanka Technological Campus*. https://www.researchgate.net/publication/378241084_Forest_Fire_Prediction_Models_in_Portugal_A_Comprehensive_Review

Safe Communities Portugal (2017, October 12). Pedrógão Grande report identifies serious failures in the response to the fire | Safe Communities Portugal. *Safe Communities Portugal*. <https://www.safecommunitiesportugal.com/report-early-warning-could-have-prevented-most-of-the-64-deaths-recorded-in-the-pedr%C3%B3g%C3%A3o-grande-fire/>

Holroyd, M (2022, September 13). Eleven people acquitted of negligence over deadly 2017 wildfires in Portugal. *Euronews*. <https://www.euronews.com/2022/09/13/eleven-people-acquitted-of-negligence-over-deadly-2017-wildfires-in-portugal#:~:text=Eleven%20people%20acquitted%20of%20negligence%20over%20deadly%202017%20wildfires%20in%20Portugal,-More%20than%2040&text=At%20least%2066%20people%20died,hit%20central%20Portugal%20in%202017.>

Khun, T., Trisnawarman, D., & Tony. (2023). Fire Dashboard Design in Bandung. *International Journal of Application on Sciences, Technology and Engineering*, 1(1), Article 1. <https://doi.org/10.24912/ijaste.v1.i1.17-22>

Pham, B. T., Jaafari, A., Avand, M., Al-Ansari, N., Dinh Du, T., Yen, H. P. H., Phong, T. V., Nguyen, D. H., Le, H. V., Mafi-Gholami, D., Prakash, I., Thi Thuy, H., & Tuyen, T. T. (2020). Performance Evaluation of Machine Learning Methods for Forest Fire Modeling and Prediction. *Symmetry*, 12(6), Article 6. <https://doi.org/10.3390/sym12061022>

Makaleng, L.D. (2022). A framework for implementing a scalable business intelligence system [Masters dissertation, North-West University]. At Boloka Institutional Repository. <https://repository.nwu.ac.za/bitstream/handle/10394/39336/Makaleng%20LD%20Final.pdf?sequence=1&isAllowed=y>

Fernandes, P., & Palheiro, P. (2015). *Interpretação dos índices do Sistema Canadano de Indexação do Perigo de Incêndio Florestal*. <https://doi.org/10.13140/RG.2.1.3683.8568>

Varela. (2018). Fire Weather Index (FWI) classification for fire danger assessment applied in Greece. *Tethys, Journal of Weather and Climate of the Western Mediterranean*, 15, pp. 31-40 <https://doi.org/10.3369/tethys.2018.15.03>

Sampaio, G. (2022). Portugal foi o país com maior área ardida em toda a Europa desde 2015, Realça-se no Facebook. Polígrafo. Sapo.pt. <https://poligrafo.sapo.pt/fact-check/portugal-foi-o-pais-com-maior-area-ardida-em-toda-a-europa-desde-2015-realca-se-no-facebook/>

Larguesa, A (2015). Área ardida aumentou 400% nos incêndios de 2015. *Jornal de Negócios*. https://www.jornaldenegocios.pt/economia/detalhe/area_ardida_aumentou_400_nos_incendios_d_e_2015

Carvalho, P. (2017). Estatísticas negras dos fogos em Portugal: 2003, 2005 e 2017. *Eco. SAPO*. <https://eco.sapo.pt/2017/06/18/estatisticas-negras-dos-fogos-em-portugal-2003-2005-e-2017/>

Budhathoki, N. (2023) Data Bias vs. Model Bias in Machine Learning. Medium. <https://pub.aimind.so/data-bias-vs-model-bias-in-machine-learning-75d62400a89b>

Powell, B. (2018). Mastering Microsoft Power BI (First Edition). Packt Publishing Ltd.

https://books.google.pt/books?hl=en&lr=&id=krZTDwAAQBAJ&oi=fnd&pg=PP1&dq=kimball+methodology+Four-Step+Dimensional+Design+Process+%22Power+BI%22&ots=vpbeKmvvfN&sig=tey0TGcOIKLf9DM2f8pcYGpXQn0&redir_esc=y#v=onepage&q=kimball%20methodology%20Four-Step%20Dimensional%20Design%20Process%20%22Power%20BI%22&f=false

Kimball, R., Ross, M. (2013). The Data Warehouse Toolkit. (Third Edition). John Wiley & Sons, Inc.

