

NOVA

IMS

Information
Management
School

MDSAA

Master's Degree Program in
Data Science and Advanced Analytics

Monitoring Public Procurement in Mainland Portugal
Visualization Driven Web Platform

Mariana Zina Almeida

Project Work

presented as partial requirement for obtaining a Master's Degree in Data Science and Advanced Analytics

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

Universidade Nova de Lisboa

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

MONITORING PUBLIC PROCUREMENT IN MAINDLAND PORTUGAL

Visualization Driven Web Platform

by

Mariana Zina Almeida

Project Work presented as partial requirement for obtaining the Master's degree in Data Science and Advanced Analytics, with a specialization in Business Analytics

Supervised by

Flávio Pinheiro, PhD, NOVA Information Management School
Bruno Damásio, 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, July 15th 2024

DEDICATION

To those who believed in me, even when I did not believe in myself.

ACKNOWLEDGEMENTS

I would like to thank my supervisors Prof. Flávio Pinheiro and Prof. Bruno Damásio for their guidance, support and mentorship throughout the course of my thesis. I was lucky to have such dedicated mentors who encouraged me to achieve more than I thought I could.

ABSTRACT

Transparency and accountability are critical in public procurement processes. The development of open-access visualization-driven web platforms provides an effective solution to improve comprehension and interaction with procurement data. By converting complex public procurement data into visual formats, the platform can make it easier for citizens to understand and interpret the information. By offering open access to public procurement data, the platform promotes government transparency and accountability through interactive dashboards that provide both summaries and detailed views of how public funds are utilized. This project presents the “HORKOS - Observatório de Contratação Pública Municipal”, an open-access intuitive data visualization platform to help citizens to access and understand easily Portuguese public procurement data at the municipal level. The dashboard offers four pages - General-Analysis, Municipality-Oriented Analysis, Company-Oriented Analysis, and first level of the Common Procurement Vocabulary (CPV) Analysis – that allow the user to gain a broad understanding of key metrics and trends, conduct detailed municipal-level analyses, explore data related to contracted companies, and understand procurement activities across various CPV categories. By addressing the research question, "How can a visualization-driven web platform significantly enhance the visibility of public procurement data and make analysis more accessible to citizens?", the project shows that the platform’s user-centred design, combined with strong data visualization capabilities, proved effective in providing citizens with actionable insights, promoting good governance and public trust in procurement practices. The platform can be accessed via this [link](#).

KEYWORDS

Public Procurement; Data Visualization; Transparency; Web Platform Development;
Power BI

Sustainable Development Goals (SDG):



TABLE OF CONTENTS

1. INTRODUCTION	1
2. RELATED WORK	3
2.1. UNDERSTANDING PUBLIC PROCUREMENT	3
2.2. LEVERAGING OPEN GOVERNMENT DATA WITH DATA VISUALIZATION PLATFORMS	9
3. DATA AND METHODS	14
3.1. CHARACTERIZING THE DATA	14
3.2. METHODOLOGY	19
4. RESULTS AND DISCUSSION	22
4.1. BUILDING THE PLATFORM.....	22
4.2. HORKOS: OBSERVATÓRIO DE CONTRATAÇÃO PÚBLICA MUNICIPAL	25
4.3. EVALUATION	39
4.4. EVALUATION RESULTS AND DISCUSSION	40
5. CONCLUSIONS AND FUTURE WORKS	44
5.1. LIMITATIONS AND FUTURE WORK	45
BIBLIOGRAPHICAL REFERENCES	46
APPENDIX A	51
APPENDIX B	53

LIST OF FIGURES

Figure 1 - Preprocessing steps.....	15
Figure 2 - a) Number of contracts by type of procurement between 2009 and 2022; b) Total expenditure per type of procurement between 2009 and 2022	18
Figure 3 - a) Evolution of the number of monthly contracts between 2009 and 2022; b) Evolution of the total expenditure between 2009 and 2022. In both a) and b), each bar corresponds to a month/year and vertical red lines indicate Municipal Elections held nationwide. c) geographical distribution of the total number of contracts awarded by municipality. d) geographical distribution of the total value spent in procurement by municipality.....	19
Figure 4 - Relational Model	24
Figure 5 - Initial section of the General Analysis Page.....	25
Figure 6 - Second Section of the General Analysis Page.....	26
Figure 7 - Third Section of the General Analysis Page.....	27
Figure 8 - General Analysis Page.....	28
Figure 9 - First Section of the Municipality Page	29
Figure 10 - Second Section of the Municipality Page.....	30
Figure 11 - Third Section of the Municipality Page.....	31
Figure 12 - Municipality Page.....	32
Figure 13 - First Section of the Company Page	33
Figure 14 - Second Section of the Company Page.....	34
Figure 15 - Company Page.....	35
Figure 16 - First Section of the CPV Page	36
Figure 17 - Second Section of the CPV Page.....	37
Figure 18 - CPV Page.....	38
Figure 19 - Socio-demographic characterization of the sample.....	41
Figure 20 - Illustration of the flow of the pages and visualizations.....	53

LIST OF TABLES

Table 1 - Literature Summary on Fraud and Corruption in Public Procurement.....	8
Table 2 - Literature Summary on OGD Initiatives and Platforms	13
Table 3 - Distribution of Contracts by Type of Entity	16
Table 4 - Distribution of Contract Types	17
Table 5 - Description of the types of procurement according to the EU public procurement guidelines	17
Table 6 - Description of the dimensions and fact table.....	23
Table 7 - Summary data of the questionnaire by question.....	42
Table 8 - CVP First Level Divisions	51

LIST OF ABBREVIATIONS AND ACRONYMS

BI	Business Intelligence
CPV	Common Procurement Vocabulary
CRI	Corruption Risk Index
DATUS	Dashboard Assessment Usability Model
DAX	Data Analysis Expressions
EU	European Union
EC	European Commission
EP	European Parliament
GDP	Gross Domestic Product
GPA	Government Procurement Agreement
ICT	Information and Communication Technology
IMPIC	Instituto dos Mercados Públicos, do Imobiliário e da Construção
IOT	Internet of Things
OECD	Organization for Economic Cooperation and Development
OGD	Open Government Data
OLAF	European Anti-Fraud Office
PCC	Public Contracts Code
PDA	Procurement Data Analytics
TAM	Technology Acceptance Model
WTO	World Trade Organization

1. INTRODUCTION

Nations bear the responsibility of offering public services to improve the quality of life for their citizens. Various methods exist for countries to deliver these public services, one of which involves the issuance of contracts through a process known as public procurement. Annually, public entities within the European Union (EU) allocate approximately 14% of the Gross Domestic Product (GDP) (around 2 trillion euros) on procuring services, works and supplies (European Commission, 2017). In Portugal, the weight of contract values in 2022 represented 5,23 % of GDP (IMPIC, 2023).

Transparency and access to information are fundamental aspects of any open government. Transparency denotes governments' and organizations' readiness and commitment to share information freely, allowing citizens to comprehend their activities, decisions, and procedures. On the other hand, individuals have the right to seek, receive and impart knowledge through access to information, which promotes informed judgment and active participation in civic life. Together, they form the foundation of democracy, encouraging accountability, trust, and the prevention of corruption. Public procurement is one of the government activities that is highly susceptible to fraud and corruption (OECD, 2015; Wensink & Vet, 2013). To ensure fairness and impartiality towards potential suppliers and competitors and to facilitate informed public scrutiny, it is essential to enhance transparency and access to information throughout the entire procurement process (Soylu et al., 2022).

The substantial amount and importance of information gathered by governments highlight the significance of this data as a valuable asset for enhancing public transparency (Ubaldi, 2013). Data related to public procurement contracts ought to serve as the foundational data for analytical frameworks designed to evaluate the effectiveness and impact of public sector initiatives across different areas and dimensions (Curado et al., 2021). Emerging technologies such as Big Data and Open Government Data (OGD) have become critical for ensuring and improving transparency across all aspects of the public sector (Moreno et al., 2022). OGD can be employed to aid the public in gaining a clearer insight into government activities, evaluating its performance, and ensuring accountability for any misconduct or unmet objectives (Ubaldi, 2013). By making their data available and easier to analyse, process and combine, public institutions become more transparent and allow a new level of public scrutiny (Ubaldi, 2013). In 2020, the European Commission (EC) launched the European Strategy for Data, creating a single market for data, allowing it to flow freely within the EU and across sectors.

In Portugal, the public procurement rules are established in the Public Contracts Code (PCC), published by the decree No. 111-B/2017, 31 August. As foreseen in Article 1A of PCC, several principles must be respected in the formation and execution of public contracts, including the principles of competition, publicity, and transparency. The [BASE Portal](https://www.base.gov.pt/base4)¹ centralizes information on public contracts celebrated in Portugal under the PCC regime.

Governments can find a lasting solution in data monitoring tools for extracting valuable information from procurement transactions data in the long term (World Bank Group, 2022). To address the question, *How can a visualization-driven web platform significantly enhance the visibility of public procurement data and make analysis more accessible to citizens?*, the following objectives arise: i) simplify complex public procurement data by transforming it into intuitive visual formats; ii) increase government transparency and accountability by providing open access to comprehensive procurement data analysis. User-friendly visualizations provided by interactive data monitoring tools empower citizens and other users to monitor, analyse, and present essential procurement indicators in a customizable manner (World Bank Group, 2022).

To accomplish these objectives, the present work is divided into five sections: the next chapter introduces a brief context to what is public procurement, as well as what is the type of data that is being used, some works regarding fraud and corruption in public procurement and a look over related works on public procurement data analytics. Chapter 3 presents the data with an exploratory analysis and the methodology used. Chapter 4 covers the results, with an overview of the steps of the project development. Finally, chapter 5 concludes the work with some final thoughts and a discussion about the findings, highlighting some of the main limitations of the study, possible future directions it might take and policy implications.

¹ <https://www.base.gov.pt/base4>

2. RELATED WORK

2.1. UNDERSTANDING PUBLIC PROCUREMENT

Public procurement refers to the process by which government agencies and public sector organizations acquire goods, services, or works from external suppliers. It constitutes a significant portion of the government's expenditure (Curado et al., 2021; OECD, 2015).

According to the EC, sensible, efficient, and transparent use of public procurement can play a crucial role in addressing significant challenges in Europe. Given the substantial expenditure it involves, this approach is particularly important for fostering sustainable economic growth and creating new job opportunities. Public procurement aims to achieve socio-economic development, provide economical and efficient service to citizens and simultaneously safeguard environmental resources (Agrahari & Srivastava, 2019). Public authorities can receive greater value for money and contribute to a more innovative, sustainable, inclusive, and competitive economy by improving procurement methods, all while addressing societal policy objectives (European Commission, 2017).

Since public procurement is a key economic activity of governments (OECD, 2009), it was necessary to regulate public market, particularly in the post-World War II period where the main goal was to foster co-operation and peace through cross-border transactions. Therefore, the World Trade Organization (WTO) was created with the aim of achieving "value for money" (Marcos, 2023). For Agrahari and Srivastava (2019), there are some key conditions to achieve "value for money": prevent waste and foster competition, transparency, and accountability. Several WTO members negotiated the Agreement on Government Procurement (GPA) to ensure equitable and transparent competition within government procurement markets. Being Portugal one of the EU member states, not only its public procurement legislation is highly shaped by the European Law but is also bound to the WTO GPA (Esteves de Oliveira & Pinto Correia, 2022). In 2014, the adoption of a new generation of public procurement directives, Directive 2014/24/EU, provided a framework for a more flexible procuring. Its stronger provisions on integrity and transparency target fraud and corruption (European Commission, 2017).

The EU has made a strong commitment to the development of Information and Communication Technology (ICT), which are seen as crucial for a prepared and competitive Europe for a global economy that will increasingly be digital (Silva, 2019). Therefore, the

aforementioned directive also brought along another big change for public procurement – Electronic Procurement, also known as E-Procurement. E-procurement refers to the integration of digital technologies in the replacement or redesign of paper-based procedures throughout the procurement cycle (OECD, 2015). Electronic procurement is a fundamental tool to ensure the integrity of the public procurement process. It argues, in this regard, that it is through this tool that economic operators can access the necessary information in a standardized and real-time manner, reducing contact between public officials and economic operators (Marcos, 2023). The use of digital technology in the public sector is a driver of efficiency and supports the effectiveness of policies by enabling more open, transparent, innovative, participatory, and trustworthy government (OECD, 2015).

As stated before, public procurement is one of the government activities that is highly susceptible to white collar crimes, due to its large level of funds involved and expenditure profile of public procurement programs. As opposed to other key components of public's expenditure, public procurement typically involves a low volume of high-value transactions, making it more attractive to potentially corrupt public officials (Ware et al., 2007). Many studies were made to evaluate in which way data science can help detect, characterize and predict fraud, corruption and collusion in public procurement activities. By now, some studies show that it is possible to identify these kinds of suspicious activities and even prevent them from occurring.

Literature on fraud and corruption in public procurement reflect the need to raise awareness of the detrimental impact of these illicit activities and to better comprehend the phenomena. To accomplish this, many recommendations were made, mainly the use of indicators/"red flags" to identify possible suspicious transactions. Corrupt activities in public procurement exhibit recurring patterns of illicit conduct, as they involve specific forms of economic behaviour (Kenny & Musatova, 2010). This behaviour leaves traces thus presenting an opportunity for intervention. Since these "red flags" are accumulations of traces that suggest corrupt activity, they are mainly aimed to help practitioners, investigators, and policymakers in predicting the likelihood of corruption in a specific procurement case, as well as to provide the groundwork for new evidence-based strategy to combat corruption (Ferwerda et al., 2016). Similarly, in 2011, the European Anti-Fraud Office (OLAF) published a set of fraud-related structural indicators within the EU, based on its investigative expertise gained over the years. By using anonymized cases where fraud elements were detected, OLAF were able to delve into the

specifics of each case and conduct a qualitative analysis to pinpoint some of the most important fraud indicators (European Anti-Fraud Office, 2011).

Ware et al. (2007) explored the vulnerabilities that occur in the public procurement process and presented some of the most observed forms of corruption, such as kickbacks, bid rigging and the use of shell companies. The identification of important indicators is frequently carried out through interviews with experts and users of public procurement and with statistical methods to assess their significance (Ferwerda et al., 2016). Wensink & Vet (2013) developed, for the EC, a special methodology to compute the risk index of corruption using logistic regression. Nevertheless, the results proved to be unsatisfactory due to the limit number of samples and indicators. Ferwerda et al. (2016) showed that, by creating sufficiently large and balanced dataset of corrupt and clean procurement cases, it is feasible to test the efficacy of some of the most notable “red flags” proposed by experts on corruption and public procurement. Moreover, Fazekas et al. (2016) addressed the lack of reliable indicators of corruption by developing an indicator of high-level institutionalised corruption at a contract level using publicly available electronic public procurement records in Hungary. This indicator, called Corruption Risk Index (CRI), relies on 13 “red flags” that are able to predict likely corrupt outcomes in regression analysis. The results show that companies with a higher corruption risk had relatively greater profitability, higher ratio of contract value to initial estimated price, an increased probability of being managed or owned by politicians, and a higher likelihood of being registered in tax havens, compared to firms scoring lower on the corruption index.

One of the issues raised in this matter is how can data science help in identify and predict suspicious public procurement transactions? Many authors conducted studies to evaluate in which way data analytics and machine learning techniques were useful to detect fraud and corruption in public procurement transactions. Cheng et al. (2017) used community detection algorithms to identify bid rigging behaviours in water projects. The results show that closely connected nodes with relatively large weights indicate that the enterprises in the communities have cooperative behaviour and are likely bidding fraudulently. Likewise, Fazekas & Tóth (2016) used public procurement data from Hungary and established a network with a subset of actors and the relationships between them. This network analysis allowed them to identify companies that conduct extensive contracts within their community rather than with other enterprises outside their communities.

Fazekas & Wachs (2020) analysed the effects of corruption in the structure of public procurement markets and how changes in government might disrupt corruption, using government data of Hungary and Czech Republic. To detect evidence on suspicious behaviour, the authors used network framework to test the qualitative hypotheses. The results revealed that government turnover does affect corrupt contracting networks and that corruption in public procurement is mostly about excluding non-favoured suppliers. Additionally, the authors observed that buyers with a high CRI and low-competitive clustering experience considerably bigger changes in their contracting relationships throughout government turnovers than other buyers. In the same manner, by representing public procurement markets as bipartite networks, Wachs et al. (2021) aimed to understand the distribution of corruption risk in EU member states. The study measures the degree to which a single bid differs across different communities inside a network. Significant evidence that corruption within the public sector is not arbitrarily distributed was found. Highly centralized markets tend to have a higher corruption risk, which is consistent with other's authors findings.

Text-mining has been used to detect indications of corruption in finance, insurance, medicine and other areas. Rabuzin & Modrušan (2019) investigated how this machine learning method could be applied to public procurement and detect indications of corruption. Findings prove that it is possible to detect signs of corruption in tender documentation suggesting that a system built around these models would be a useful tool for public procurement monitoring specialists. Nevertheless, the authors noted that the low power of the model was due to a lack of data that could have been used as dependent variables.

Regarding Portuguese public procurement data, Curado et al. (2021) proposed a framework to study quantitatively local public expenditure borrowing methods from urban scaling laws. The study demonstrates that public procurement expenditure scales sublinearly with population size, corroborating previous findings in the urban scaling laws literature. Furthermore, an upward trend was found in the annual variation of the scaling coefficients. Such trend is linked to socio-political events that took place in Portugal and to a revision of the municipal finances law. The study also showed a new regional characterization of municipalities through their similarities in procurement activity. Moreover, authors argue that studying public procurement data at a municipal level brings a lot of value as this kind of data offers a detailed categorization of the spending patterns of municipalities, and since this data is standardized across EU members it is possible to compare regional activities.

All these studies showcase one thing in common: the potential benefits of adopting a data science perspective to OGD. However, it still faces some challenges. Data needs to be enhanced in quality and interoperability in order to improve machine learning efforts.

The findings from research on fraud and corruption in public procurement can inform the development of effective public procurement visualization platforms, which could enhance trust and accountability within these systems. Although the platform created in this project may not directly address fraud and corruption neither visualize illicit activities, understanding the dynamics and prevalence of such issues is crucial for designing successful visualization tools. Ensuring that the presented data is clear, accurate, and comprehensive for the average citizen will support better scrutiny and oversight. This, in turn, promotes transparency and accountability in public procurement processes. Consequently, this approach contributes to a more trustworthy and efficient public procurement system, even if the platform itself does not directly address fraud and corruption.

Table 1 - Literature Summary on Fraud and Corruption in Public Procurement

Paper	Year	Dataset	Conclusions
(Ware et al., 2007)	2007	Not Applicable	Corruption in public procurement is a significant issue, with various stages in the procurement process being vulnerable to corrupt practices. Implementing anti-corruption measures in procurement requires a structured approach like a road map, which helps identify vulnerabilities and potential points of corruption, aiding in combating corrupt practices effectively.
(Wensink & Vet, 2013)	2013	Procurement data of 8 EU Member States	The study emphasizes the importance of measuring the costs of corruption as a foundation for effective anti-corruption policies. It is estimated that corruption explains over two-thirds (69%) of direct public losses in public procurement.
(Fazekas et al., 2016)	2016	Hungarian procurement data	The Corruption Risk Index (CRI) developed in this paper allows for the measurement of corruption at the microlevel using objective data, demonstrating its feasibility and validity.
(Fazekas & Tóth, 2016)	2016	Hungarian procurement data	The study demonstrates the feasibility and usefulness of measuring corruption using transaction-level public procurement data and linking it to state capture, offering a new analytical framework for research and policy on anticorruption, budget deficit, market competition, and quality of democracy.
(Ferwerda et al., 2016)	2016	Procurement data of 8 EU Member States	By testing the 28 indicators against control groups, it was found that 18 red flags are significantly correlated with corrupt public procurements, emphasizing the effectiveness of certain indicators in detecting corruption.
(Cheng et al., 2017)	2017	Data related to water project bidding	Results show that closely connected nodes with relatively large weights indicate that the enterprises in the communities have cooperative behaviour and are likely bidding fraudulently.
(Rabuzin & Modrušan, 2019)	2019	Croatian procurement data	Machine learning detects corruption in public procurement through text-mining techniques.
(Fazekas & Wachs, 2020)	2020	Hungarian and Croatian procurement data	High corruption risk buyers in public procurement markets tend to have sparser network neighbourhoods, leading to fewer contracting relationships with suppliers, indicating corruption is about the exclusion of non-favoured firms. Political change significantly impacts corrupt relationships.
(Wachs et al., 2021)	2021	Procurement data of 26 EU Member States	Highly centralized procurement markets tend to have higher corruption risk, indicating a strong relationship between market centralization and corruption risk.
(Curado et al., 2021)	2021	Portuguese procurement data	Public procurement expenditure scales sub linearly with population size, indicating an economy of scale for public spending as cities increase in population size. When looking at the municipal Scale Adjusted Indicators, they identified a richer local characterization of municipalities based on the similarity of procurement activity.

2.2. LEVERAGING OPEN GOVERNMENT DATA WITH DATA VISUALIZATION PLATFORMS

Before E-Procurement, data regarding public procurement was difficult for citizens to access and, as a result, there was a common belief that several illegalities were taking place, as well as a significant lack of transparency. While not being the only motivation, corruption was one of the main reasons behind starting ODG initiatives (Attard et al., 2015), like the ones mentioned below.

Allied to the concept of Open Governance is the movement of Open Government Data (OGD), which is related to the earlier Right to Information movement, also known as Freedom of Information (Afful-Dadzie & Afful-Dadzie, 2017). First conceived in 2007, the OGD movement relies on social and economic reasoning to encourage the release of government data, focusing on data held in government databases. The Right to Information movement, officially approved in 1946 by a United Nations General Assembly Resolution, supports the idea that people have the right to access information. Despite the differences between the two movements, both aim to increase government transparency (Ubaldi, 2013).

Supporters of OGD argue that offering free and easy access to public data increases democratic involvement and allows individuals to build new products and services based on well-informed and reliable public data (Janssen et al., 2012; Maude, 2012). Moreover, the EC believes that data has an enormous potential to address the needs of individuals and, therefore, create value for the economy and society (European Commission, 2020). In 2019, the European Parliament (EP) and the Council of the EU approved a new directive, called Open Data Directive², which requires the distribution of government data in free and open forms. The Directive's overarching goal is to improve the EU's data economy by expanding the amount of public sector data available for re-use, guaranteeing fair competition and easy access to public sector information, and improving cross-border data-driven innovation. A year later, it was launched the European Strategy for Data, intending to create a single market for data that will guarantee Europe's global competitiveness and data sovereignty. As identified by Zuiderwijk et al. (2016), the quality of the data published determines the success of OGD projects. To guarantee the quality of the data made available, this strategy is based on four pillars, which include the Data Governance Act, which aims to increase the accessibility of data and promote sharing of data among various sectors and EU nations; and the Data Act, ensuring fairness by

² Directive (EU) 2019/1024

defining regulations regarding the use of data produced by Internet of Things (IoT) devices and ensuring consistency between data access rights.

According to Gonzalez-Zapata & Heeks (2015), there are four different perspectives on OGD: i) in *the Bureaucratic Perspective*, OGD is seen as a government policy to support the service delivery of public sector data; ii) *the Technological Perspective* conceives OGD as a technological innovation centred on the development of processes, formats and standards for handling public sector data; iii) *the Political Perspective*, associated with the concept of open government, where access to public sector data is a fundamental right to improve transparency and accountability of government officials; and lastly iv) *the Economic Perspective*, where OGD is a mean to further economic growth. Taking into account the purpose of this project, I am going to focus on the second perspective mentioned above.

Following the literature, OGD platforms or OGD portals are government data systems in charge of gathering, publishing, and processing OGD (Moreno et al., 2022). Zuiderwijk-van Eijk (2015) states that OGD portals' main activities are divided into five categories:

- *Searching and finding data*: users should be able to access government data from a variety of sources, to access metadata in multiple languages, and the platform should facilitate filtering and sorting search results.
- *Analysis*: the OGD infrastructure should have specifications that allow users to analyse the available datasets, such as, providing data to support users in understanding and interpreting the datasets in various formats.
- *Visualization*: the platform should provide and integrate visualizations tools.
- *Interaction*: users must be able to interact with other users, to receive notifications about changes in datasets and to give feedback or discuss datasets.
- *Quality analysis*: the infrastructure should enable users to discuss the quality of a dataset and compare the dataset quality with other datasets or other versions.

Furthermore, some authors claim that the majority of conventional open data systems lack more advanced data analysis tools and offer basic upload and download capabilities (Charalabidis et al., 2014).

In Portugal, data on public contracts celebrated under the PCC regime is centralized in BASE Portal. This OGD portal supports the claim stated above, as does not offer advanced data analysis tools. The purpose of this project is to develop an interactive and intuitive platform for

analysing Portuguese public procurement data at the municipal level. Covering data from 2009 to 2022 for mainland Portugal, this platform will allow users to filter information and draw their own conclusions.

As we have already established, public procurement is a key method for allocating public funds. Thus, it is essential to improve transparency and monitoring throughout the procurement process (Siciliani et al., 2023). This can be done through Procurement Data Analytics (PDA). World Bank (2022) defines PDA as the use of data to generate useful insights and evidence to monitor outcomes, guide reform efforts, and assess the impact of reforms and strategies in public procurement.

While governments have been implementing OGD policies to make their data publicly available, the availability of this data is insufficient as the majority of the data provided is in the form of raw datasets. This is primarily due to a lack of skills and expertise among the majority of the population to collect, process, and transform this data into useful information (Ansari et al., 2022; Eberhardt & Silveira, 2018; Graves & Hendler, 2013). To overcome this technical barrier, OGD platforms can be complemented with visualizations. Visualization has been in use for centuries since it constitutes a mechanism to analyse and comprehend information in a simple way. It has been demonstrated to be beneficial in understanding large quantities of data across various fields (Graves & Hendler, 2013). Therefore, the use of data visualization enables the user to skip the technical procedures of collecting, process and transform data by providing a straightforward data access point. However, some portals fail to evaluate whether their visualizations are in line with best practices, resulting in designs that are difficult for citizens to understand and engage with. Ensuring that interested parties comprehend the communicated information is fundamental (Oliveira et al., 2023). This need becomes even more critical when dealing with public data, as understanding is essential for promoting transparency, strengthening democracy, and increasing public scrutiny (Sundberg & Gidlund, 2022). Oliveira et al. (2023) argue that incorporating plain language alongside visualizations in graphic design enhances citizens' understanding of information, corroborating the findings of Barcellos et al. (2022).

As summarized in Table 2, the studies highlighted in this section revolve around how important visualization is for conveying information about public data as transparency and public scrutiny are essential to strengthen democracy.

This literature review identifies one significant gap in the existing research: the lack of advanced data analysis tools to analyse OGD data available. Some studies have already addressed this gap in other countries such as in Albany (Graves & Hendler, 2013), Colombia (Moreno et al., 2022), India (Agrahari & Srivastava, 2019), and South Korea (Lee et al., 2017). However, it appears that there are no existing studies that specifically address this gap in the context of Portuguese public procurement. The current project aims to bridge this gap by creating an open-access intuitive data visualization platform to help citizens to access and understand easily Portuguese public procurement data.

Table 2 - Literature Summary on OGD Initiatives and Platforms

Paper	Year	Dataset	Conclusions
(Graves & Hendler, 2013)	2013	Not Applicable	Visualizations are crucial for understanding and communicating large amounts of data, especially for individuals without technical expertise. Most visualization tools do not adequately address the unique requirements of Open Government Data.
(Gil García et al., 2012)	2012	Not Applicable	Transparency portals serve as a key platform for publishing government-related information, enabling citizens to access data and report any irregularities they observe.
(Agrahari & Srivastava, 2019)	2019	Indian procurement data	The significance of benchmarking in the public procurement process and proposes an IT-driven tool for the same. The application of data visualization through e-Procurement data not only enables benchmarking but also contributes to performance improvements, competitiveness, effectiveness, and transparency in public procurement.
(Moreno et al., 2022)	2022	Colombian procurement data	Through a focused case study on monitoring public procurement data, the research showcases the significance of the platform in validating data quality and integrity within the public sector, emphasizing its role in supporting government transparency.

3. DATA AND METHODS

3.1. CHARACTERIZING THE DATA

In Portugal, the decree No. 18/2008 of 29 January, foresaw the creation of a portal to publish public contracts, with the aim of promoting transparency in public procurement. Since January 1st 2009 public contracts in Portugal must be awarded using electronic platforms and published through BASE portal. BASE portal, managed by *Instituto dos Mercados Públicos, do Imobiliário e da Construção* (IMPIC), is an open-access governmental public repository that centralizes information on public contracts celebrated in Portugal under the PCC regime.. Public administration bodies and other regardless of their public or private nature, whose purpose is to fulfil needs in the public interest are subject to the public procurement regime. The data used in this project comprises information about 1.399.315 public contracts celebrated in Mainland Portugal under the PCC regime spanning from 2009 to 2022.

Portugal is geographically organized in two different ways: statistically and administratively. The statistical organization, known as the Nomenclature of Territorial Units for Statistics (NUTS), was established by Eurostat in the early 1970s to standardize regional data collection, compilation, and dissemination across countries. It is subdivided into three levels: NUTS I, NUTS II, and NUTS III. NUTS I divides Portugal into its mainland territory and its two autonomous regions, the Azores and Madeira. NUTS II split the country into 9 regions, and NUTS III further divides it into 26 administrative units, including 24 intermunicipal authorities and the two autonomous regions (INE, n.d.). These NUTS regions were created for statistical purposes only; they cannot be directly associated with public contracts, nor do they have directly elected governments. On the other hand, the administrative divisions in Portugal do have these elements. Administratively, Portugal is organized into three levels: the primary level consists of 18 districts, the secondary level includes 308 municipalities (278 on the mainland and 30 in the islands), and the tertiary level comprises 3,091 parishes (INE, n.d.). The administrative division is the oldest territorial division in the country (INE, n.d.).

Municipalities are Local Administrative Units with the most consistently defined regional boundaries, whose governance body is elected by universal suffrage (INE, n.d.). Therefore, they constitute good candidates for studying public procurement. As Azores and Madeira are autonomous regions endowed with political-administrative statutes and, they won't be taken

into consideration for this project. Hence, this analysis is going to be focused on the public procurement contracts issued by the 278 mainland municipalities.

In order to select the contracts issued by each municipality, it was necessary to identify the Fiscal Number of each municipality. This was manually curated with the support from the Yearly Financial Booklet of Portuguese Municipalities (Fernandes et al., 2022). The curated data is available online through a public Github repository³ for future reference. The preprocessing steps, illustrated in the figure below, include:

- Drop unnecessary variables to the analysis;
- Identifying the Fiscal Number of each municipality;
- Add the CPV's description;
- Classify the contracts into supplies, works or services according to the UE public procurement guidelines;
- Aggregating municipal companies to the parent Municipality;
- Discard all non-Municipality related contracts;
- The value of contracts of companies owned by more than one municipality was split according to each municipality's stake in the company.

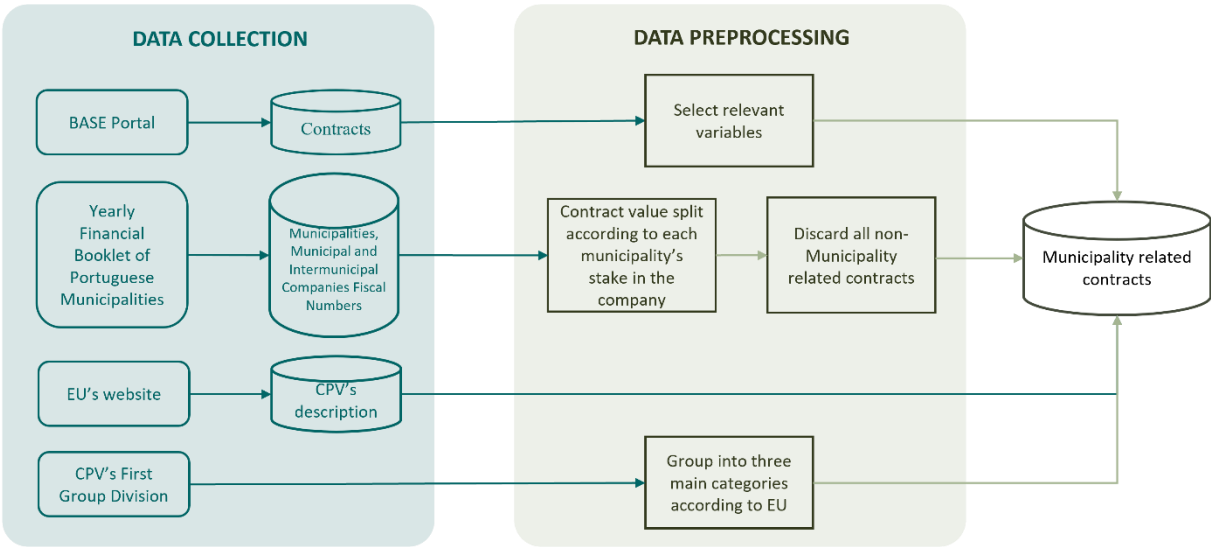


Figure 1 - Preprocessing steps

³ <https://github.com/marianazinaalmeida/Portuguese-Public-Procurement-Platform>

The final dataset comprises 411.495 contracts between 2009 and 2022, worth over 34 billion euros. Table 3 provides a comprehensive analysis of the distribution of contracts across different types of municipal entities: municipalities, municipal companies and intermunicipal companies. Municipalities hold most contracts, indicating their primary role in local governance and service delivery. Following are municipal companies, indicating their significant, albeit smaller, role compared to municipalities. Inter-municipal companies hold the smallest share at 3%, suggesting a more limited or specialized function in the regional context.

Table 3 - Distribution of Contracts by Type of Entity

	Number of Contracts	Number of contracts (%)	Total Expenditure (€)
Municipality	345.642	84%	27.467.921.142
Municipal Company	53.003	13%	3.849.303.161
Inter-municipal Company	12.850	3%	491.682.845

This study analyses the procurement activities and the allocation of public funds in each municipality. Furthermore, we analyse these transactions based on the three distinct types of procurement they represent. According to the UE public procurement guidelines there are three different types of procurement, as described in Table 4: Supplies, Services or Works (European Commission, 2018). This division is only possible because of the Common Procurement Vocabulary (CPV). The CPV was developed by the EU and introduces a classification framework for public procurement, with the goal of standardising the terminology used by contracting authorities and entities to characterize the subject matter of procurement contracts. It corresponds to a tree-based structure consisting of a main vocabulary and a supplementary vocabulary (European Union, n.d.). Table 7 in Appendix A lists the 2-digit (first level) of the classification.

Table 4 - Description of the types of procurement according to the EU public procurement guidelines

Type of Procurement	Description
Supplies	Public contracts having as their object the purchase, lease, rental or hire purchase with or without option to buy, of products such as stationery, vehicles, or computers.
Services	Public contracts other than public works or supply contracts having as their object the provision of services such as consultancy, training, or cleaning services.
Works	Public contracts having as their objective either the execution, or both the design and execution, of works, for example building or civil engineering works such as a road or sewage plant.

Table 5 presents an overview of the distribution of the contracts among the three most common types of procedure, as well as the total spent in each one.

Table 5 - Distribution of Contract Types

	Number of contracts	Distribution of the number of contracts by type of contracts (%)	Total Expenditure by type of contracts (€)	Distribution of the expenditure by type of contracts (%)
Direct Award General Regime (1)	290.559	71%	9.536.095.715	33%
- Supplies	80.722	28%	1.674.899.511	6%
- Services	149.266	51%	4.792.280.602	17%
- Works	60.571	21%	3.068.915.600	11%
Prior Consultation (2)	64.156	16%	2.741.685.580	9%
- Supplies	20.664	32%	506.993.872	2%
- Services	29.052	45%	1.038.316.411	4%
- Works	14.440	23%	1.196.375.296	4%
Public Tender (3)	47.368	12%	14.218.146.430	49%
- Supplies	12.580	27%	1.320.591.614	5%
- Services	15.694	33%	3.234.840.976	11%
- Works	19.094	40%	9.662.713.839	33%
Other Types of Contracts (4)	9.412	2%	2.520.421.258	9%
Total (1+2+3+4)	411.495	100%	29.016.348.983	100%

What is interesting in this data is that the Direct Award General Regime comprises the majority of the contracts being analysed in this project: 71% of the total number of contracts. However, this regime only represents 33% of the total expenditure. On the other hand, Public Tender comprises almost half of the total expenditure (49%) despite comprising only 12% of the total number of contracts. This means that Public Tender contracts tend to involve significantly larger sums of money on average compared to Direct Award General Regime contracts.

The trend of the number of contracts awarded in each procurement type per year is shown in panel a) of Figure 1. Overall, both Supplies and Services exhibit an upward trend over the years under analysis. The former has gradually increased since 2012, while the latter had a sharp rise since the same year. However, the number of contracts awarded in Works presents fluctuations over time, with no clear trend observed. Moving to panel b) of Figure 1, it illustrates the total expenditure by procurement type over the 14-year period. It is clear that all three types reached a peak in 2011. However, the peak expenditure for Supplies was not as high as the that registered in Services and works. Subsequently, both Supplies and Services demonstrated a steady increase, while Works continue to fluctuate, having a substantial increase in 2019 reaching another peak in 2021.

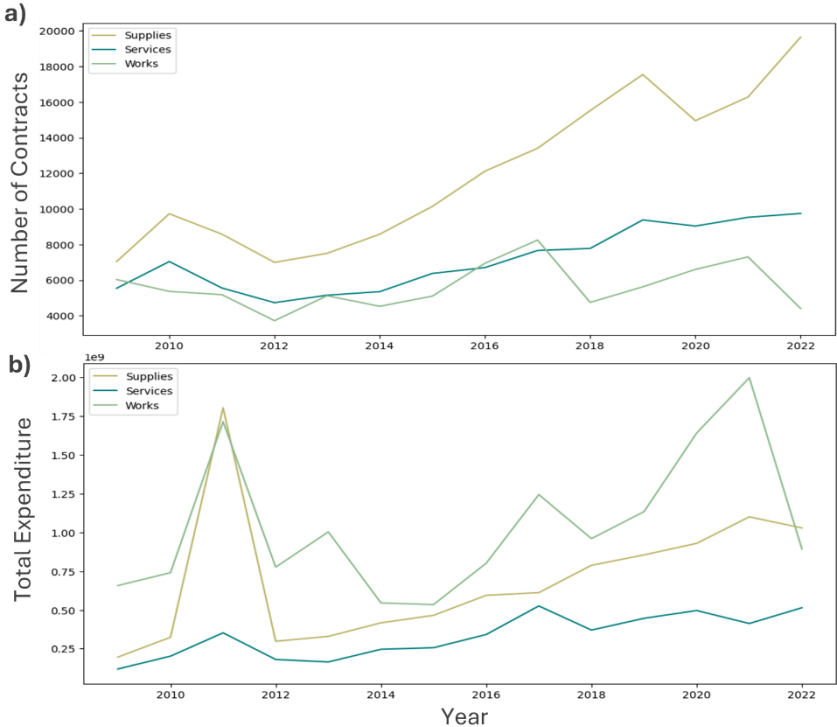


Figure 2 - a) Number of contracts by type of procurement between 2009 and 2022; b) Total expenditure per type of procurement between 2009 and 2022

Panel a) of Figure 2 presents the monthly number of contracts awarded by municipalities over the years under analysis. From this data, we can observe that it seems to be a trend among municipalities to increase the number of procurement contracts issued in the months preceding elections. Such trend can also be seen in the total expenditure as illustrated in panel b).

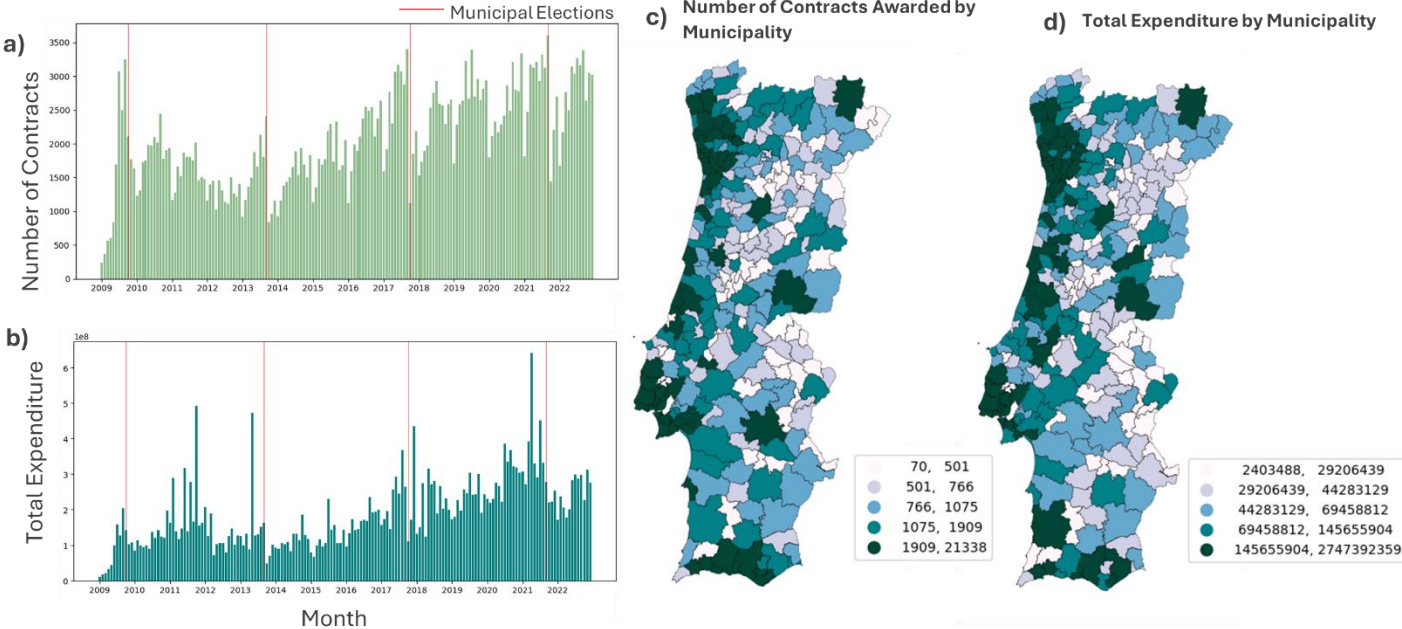


Figure 3 - a) Evolution of the number of monthly contracts between 2009 and 2022; b) Evolution of the total expenditure between 2009 and 2022. In both a) and b), each bar corresponds to a month/year and vertical red lines indicate Municipal Elections held nationwide. c) geographical distribution of the total number of contracts awarded by municipality. d) geographical distribution of the total value spent in procurement by municipality

3.2.METHODOLOGY

Portugal, as a member of the EU, recognizes the matter of public procurement as a strategic issue. A survey conducted by OECD (2015) has shown that Portugal measures the results of its strategy to promote green public procurement. Since 2008, all elements relating to the formation and execution of public contracts must be communicated on BASE portal (www.base.gov.pt). This enables the contracts to be monitored and followed up by the government and all Portuguese citizens, bringing transparency to these processes. While this portal enables users to access all public procurement records, it lacks any exploitation capabilities. There is a gap in tools for processing and visualizing public procurement information to support anyone who is interested in studying potential issues with the procurement process (Moreno et al., 2022). To bridge the gap aforementioned and to contribute to public scrutiny, in this project was developed a visualization platform, called “Observatório de Contratação Pública Municipal”

(HORKOS), that provides a series of pages that display descriptive analytics results to assist in public procurement evaluation.

This chapter outlines the methodological approach used to build the HORKOS platform. The methodology involves a systematic process consisting of the following steps:

Dashboard Development

The dashboard was developed in Microsoft Power BI due to its capabilities of integrating data from different sources, data visualization capabilities and user-friendly interface. As further explained in the next chapter, these features make Power BI an excellent tool for creating an interactive and informative dashboard that allows the data to be accessible to a wide audience. The design of the dashboard prioritized clarity, usability and interactivity, allowing every user, including the ones with low analytical skills, to make the most of it. The layout includes a main overview page summarizing key metrics, a detailed page for analysing procurement in each municipality, a page for analysing procurement activities in each contracted company, and a last page for analysing procurement activities in each first-level CPV. These pages allow the user to gain a broad understanding of key metrics and trends, conduct detailed municipal-level analyses, explore data related to contracted companies, and understand procurement activities across various CPV categories.

In order to import the cleaned and pre-processed data detailed in the previous chapter, a data warehouse was created, as explained in the upcoming chapter. This data warehouse uses a star schema to organize the information effectively. The fact table represents the contracts, containing transaction details, while the dimension tables provide descriptive attributes such as procurement types, contract types, and metadata about municipalities. Key measures, including the total procurement value, were calculated using Data Analysis Expressions (DAX).

Analysis and Visualization

The analysis focused on several key metrics such as total procurement value, and consequently the aggregate value of procurement activities per municipality, the total number of contracts, the most contracted companies by number of contracts, the municipalities that contract the most, most common first level CPV categories. To visualize and effectively communicate these insights, various visualizations were employed including bar charts, line charts and columns charts to depict trends over time, and treemaps to display the first level CPV categories in the contracts awarded by the municipality.

To improve user interactivity and analytical flexibility, several features were added. A date slicer allows users to filter data for any period between 2009 and 2022 or to focus on a specific year. Additionally, a navigation bar in the top right corner provides easy access to all the pages of the platform. Furthermore, each visual is paired with a text box next to the plot that briefly explains the plot's contents, ensuring that even users with limited analytical skills can effectively use the platform.

Evaluation and Validation by Final Users

The platform was evaluated through a qualitative questionnaire developed based on two existing questionnaires, as further explained on chapter 4.3. The customized questionnaire was designed to assess usefulness, ease of use, learnability, satisfaction, user interface and content. It also includes questions aimed at collecting feedback from users, which will be considered for future improvements to refine the dashboard's usability and functionality. Participants were asked to use the dashboard and then complete the questionnaire.

By following these steps, the HORKOS platform aims to deliver a robust tool for analysing and visualizing Portuguese public procurement data, at the municipal level, ultimately promoting transparency and accountability.

4. RESULTS AND DISCUSSION

4.1. BUILDING THE PLATFORM

Restating what was mentioned in Chapter 2, OGD often has a significant limitation: its distribution format makes it challenging for non-technical users to access and comprehend large volumes of data. Typically, the data is offered as raw datasets, creating a technical obstacle that limits its accessibility and usability. Therefore, this project aims to offer an easy-to-use data access point that allows users to skip all preliminary steps, such as collecting relevant data, loading and reading it with the right tools, searching for necessary information, and visualizing the results (Graves & Hendler, 2013). To address these challenges, I opted for using a Business Intelligence software solution.

In an era of digital transformation, Business Intelligence (BI) plays a pivotal role in meeting the needs of modern organizations. To demonstrate how BI is used in data analysis, Microsoft has introduced Power BI, which is a comprehensive business analysis service featuring data visualization and BI functionalities for users. Power BI is a cloud-based reporting service developed by Microsoft that enables the development of reports from vast amounts of data. It offers a collection of tools that connect data from different sources into a unified platform. In the context of public procurement, where data often comes from different sources, Power BI can aggregate and consolidate this data. In addition, its data analysis and visualization functionalities allow the data to be understandable and accessible to a wide audience.

To visualize the data in Power BI, it was first established a data warehouse. A data warehouse is a form of data management system that is designed to support BI activities, particularly analytics. It is built to store large quantities of historical data and enable fast and complex queries across all the data, making it easier to retrieve information and generate reports. In order to structure the data in the data warehouse, it was implemented a dimensional model. Dimensional modelling is a technique aimed at optimizing databases for efficient data retrieval. First defined by Kimball (1996), this approach categorizes data into “facts” and “dimensions”. Facts represent the core entities for analysis—in this case, contracts—while dimensions provide descriptive attributes, such as contract types and procedures. The dimensional model can be structured in various ways based on project requirements.

The first step in using the Kimball & Ross (1996) methodology is to identify the business process, which, in this case, are the contracts celebrated by the municipalities, municipal and

intermunicipal companies. This means that the grain or the smallest unit of analysis will be the contracts. After identifying the business process, the next step is to identify the dimensions that will provide additional context and attribute to the model. In this project, the dimensions identified are Date, Company, Contract Type, Procedure Type, Municipalities and CPV's. A more detailed description of the dimensions and fact table is provided in Table 6.

Table 4 - Description of the dimensions and fact table

Table	Variable Name	Data Type	Description
DIM Date	SK Date	Whole number	Primary key
	Date	Date	Full date of the contract transaction
	Day	Whole number	Day when the transaction of the contract was published
	Month	Whole number	Number of the month when the transaction of the contract was published
	Year	Whole number	Year when the transaction of the contract was published
DIM Company	BK Company	Whole number	Fiscal number of the contracted companies and primary key
	company_name	Text	Name of the contracted companies
DIM Contract	SK_contract_type	Whole number	Primary key
	contract_type	Text	Type of contract
DIM Procedure	SK_procedure	Whole number	Type of procedure
	procedure_type	Text	Description of the procedure
DIM Municipalities	BK_mun	Whole number	Administrative division code and primary key
	Municípios	Text	Name of the municipalities
	nif_contracting_agency	Whole number	Fiscal number of the municipalities
DIM Cpvs	BK_cpvs	Text	CPV number and primary key
	Cpvs_description	Text	Description of the cpv
	category	Text	CPV category
	group	Text	First level of the CPV
FACT Contracts	FK_company	Whole number	Foreign key from DIM Company
	FK_contract_type	Whole number	Foreign key from DIM Contract
	FK_cpvs	Text	Foreign key from DIM Cpvs
	FK_Date	Whole number	Foreign key from DIM Date
	FK_municipalities	Whole number	Foreign key from DIM Municipalities
	FK_procedure	Whole number	Foreign key from DIM Procedure
	contract_id	Whole number	Id of the contract
	execution_deadline	Whole number	Deadline of the contract in days
	initial_price	Decimal number	Initial price of the contract
	final_price	Decimal number	Final price of the contract
adjusted_price	Decimal number	Adjusted price of the contract according to each municipality's stake in the company	

Since this project is using Power BI, the choice of model schema is a crucial step. Each visual in Power BI generates a query that is sent to the model. Therefore, it is important to have a schema that enhances query performance and that fits well with star schema principals. Dimension tables support filtering and grouping, while fact tables support summarization. This schema offers a denormalized database that's easily understandable, even for non-technical users. The star schema allows users to easily locate and access the relevant information, simplifies DAX by making them more readable and faster to write, improves performance as it is designed to handle large volumes of data efficiently and enables faster data refreshes, reducing processing time and supporting scalability for future adjustments. Figure 4 illustrates the dimensional model created for the project.

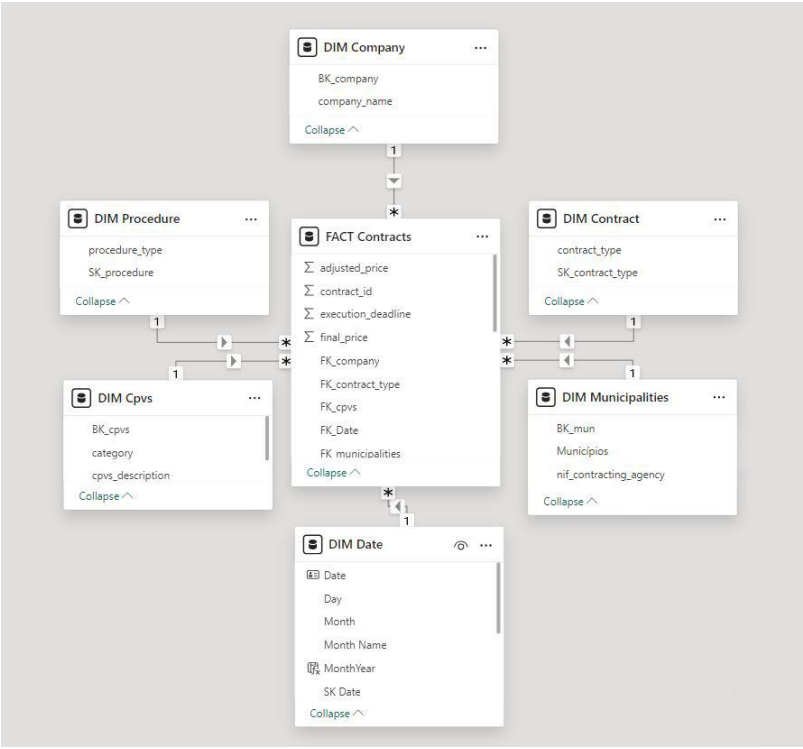


Figure 4 - Relational Model

After establishing the data warehouse, several visualizations that are useful for extracting information were chosen. These are available through pages, each tailored to a specific area of interest:

- The first page offers a general analysis, encompassing all the data.
- The second page focuses on municipality-oriented analysis, allowing users to select a specific municipality for analysis.

- The third page explores contracted companies.
- The fourth page provides a comprehensive overview of a specific first level CPV.

A more detailed description of the report's content and technical aspects are provided in the following section. An illustration of the flow of the pages and a summary of the visualizations is present in Figure 20 in Appendix B.

4.2.HORKOS: OBSERVATÓRIO DE CONTRATAÇÃO PÚBLICA MUNICIPAL

The first page (Figure 8) provides a comprehensive overview of the data, designed to give users a broad understanding of the key metrics and trends. This page is structured with various visualizations and interactive elements to enhance the user's analytical experience.

In the initial section (Figure 5a), the cover displays three Key Performance Indicators (KPIs): the number of contracts, the number of companies involved, and the number of municipalities analysed. These KPIs give users a snapshot of the overall scope of the data.

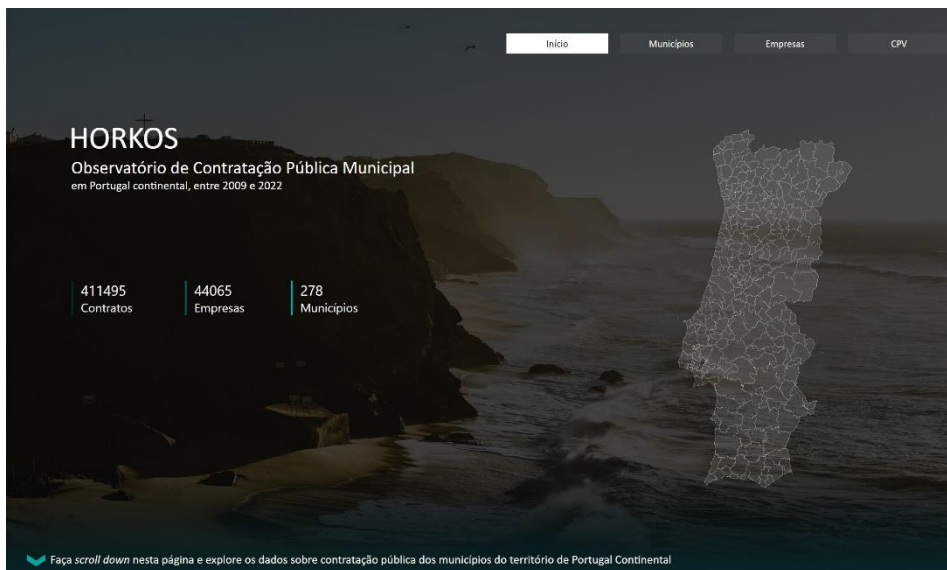


Figure 5 - Initial section of the General Analysis Page

Moving to the second section (Figure 6), the analysis delves deeper with two informational cards. One card presents the total number of contracts signed by all municipalities and municipal companies, while the other card shows the total expenditure on these contracts. Additionally, this section includes two important charts. From left to right, the first chart is a decomposition tree, breaking down the total number of contracts by procedure type and contract type. This visualization highlights the procedure with the most contracts and identifies the most common contract type. The second chart is a bar chart that identifies which municipalities have the highest number of contracts, providing a clear comparison across different regions.

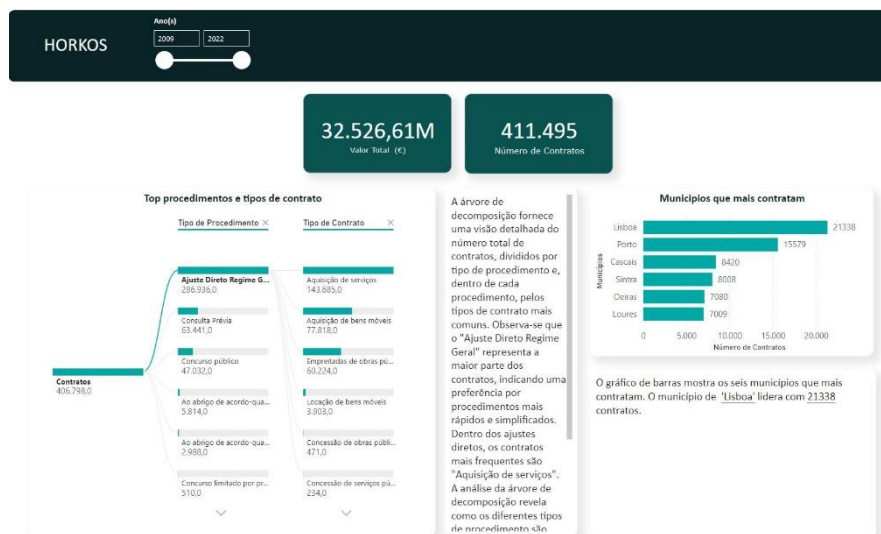


Figure 6 - Second Section of the General Analysis Page

In the third section (Figure 7), the focus shifts to more detailed breakdowns with four vertically arranged charts. The top chart is a bar chart listing companies with the highest number of contracts, offering insights into which companies are most frequently engaged. The next chart is a column chart depicting the ranking of municipalities concerning their expenditure in public procurement considering its population (based on (Curado et al., 2021)). This chart includes a slicer in the top right corner, that allows users to view the information by type of procurement. Below this, two column charts display the trends over time: one shows the evolution of the number of contracts and, the other shows the total amount spent over time. The blue dotted lines indicate when the municipal elections took place. Both charts include a toggle button in the top right corner, allowing users to switch to a line chart to analyse these trends by procurement category. Adjacent to the contract number chart, there are four cards displaying values. These cards are used for bunching analysis, helping to understand changes in the number of contracts before elections.

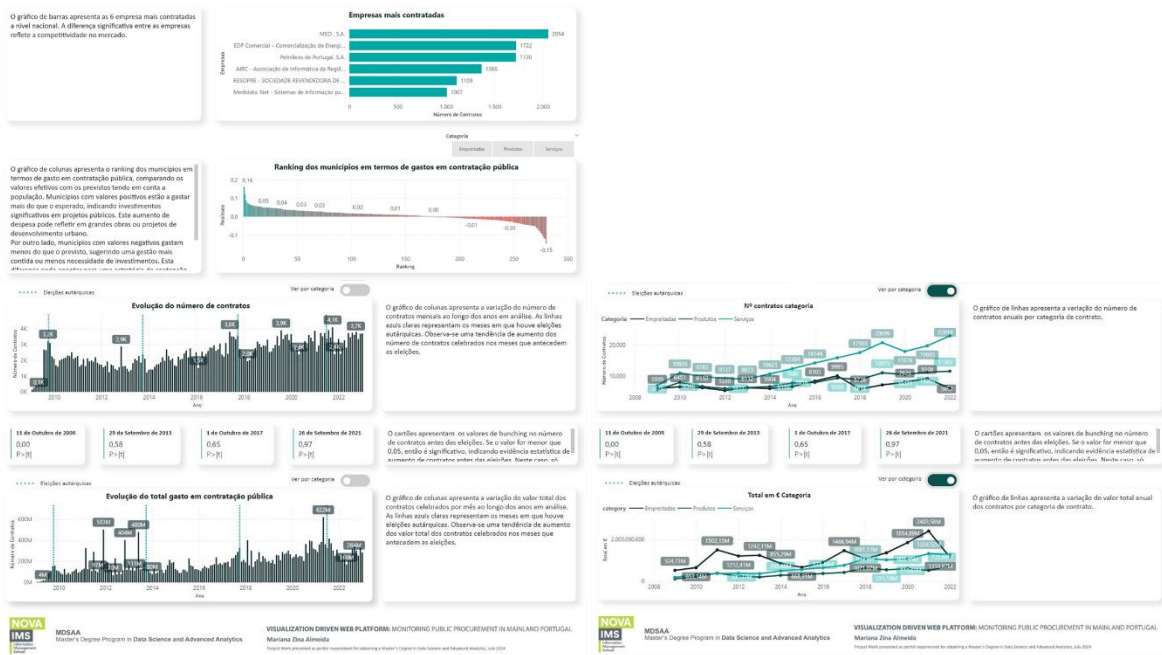


Figure 7 - Third Section of the General Analysis Page

To enhance user interactivity and analytical flexibility, some aspects were incorporated. A date slicer (Figure 6) allows users to filter the data for any period between 2009 and 2022, or to focus on a specific year. Additionally, a navigation bar located in the top right corner enables users to easily navigate through all the pages in the platform. This navigation bar is present in all the pages. Furthermore, every visual is accompanied by a text box located beside each graph. This text box provides a brief description of what is being conveyed in the graph. This feature ensures that even users without high analytical skills can understand and make the most of the platform. This feature is common in all the pages.

These interactive elements and detailed visualizations are designed to provide a thorough and flexible analytical experience, enabling users to explore and understand the data from multiple perspectives.

Inicio Municípios Empresas CPV

HORKOS

Observatório de Contratação Pública Municipal
em Portugal continental, entre 2009 e 2022

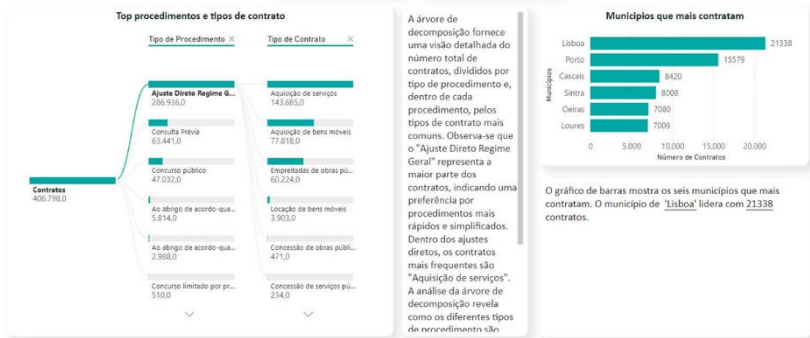
411495 Contratos | 44065 Empresas | 278 Municípios

✓ Faça scroll down nesta página e explore os dados sobre contratação pública dos municípios do território de Portugal Continental

HORKOS 2009 2022

32.526,61M
Valor Total (€)

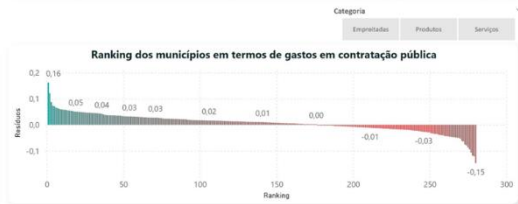
411.495
Número de Contratos



O gráfico de barras apresenta as 6 empresas mais contratadas a nível nacional. A diferença significativa entre as empresas reflete a competitividade no mercado.



O gráfico de colunas apresenta o ranking dos municípios em termos de gasto em contratação pública, comparando os valores efetivos com os previstos tendo em conta a população. Municípios com valores positivos estão a gastar mais do que o esperado, indicando investimentos significativos em projetos públicos. Este aumento de despesa pode refletir em grandes obras ou projetos de desenvolvimento urbano. Por outro lado, municípios com valores negativos gastam menos do que o previsto, sugerindo uma gestão mais contida ou menos necessidade de investimentos. Esta diferença pode resultar de um modelo de contratação...



O gráfico de colunas apresenta a variação do número de contratos mensais ao longo dos anos em análise. As linhas azuis claras representam os meses em que houve eleições autárquicas. Observa-se uma tendência de aumento do número de contratos celebrados nos meses que antecedem as eleições.



O cartões apresentam os valores de bunching no número de contratos antes das eleições. Se o valor for menor que 0,05, então é significativo, indicando evidência estatística de aumento de contratos antes das eleições. Neste caso, há...



O gráfico de colunas apresenta a variação do valor total dos contratos celebrados por mês ao longo dos anos em análise. As linhas azuis claras representam os meses em que houve eleições autárquicas. Observa-se uma tendência de aumento do valor total dos contratos celebrados nos meses que antecedem as eleições.

Figure 8 - General Analysis Page

The Municipality-Oriented (Figure 12) page provides a detailed, municipality-oriented analysis of public procurement activities. It provides a comprehensive overview of public procurement activities within each municipality, offering, for example, key insights into spending patterns.

The first section (Figure 9) serves as an introductory overview. It features images of each municipality, a map pinpointing its location, and 11 socio-economic indicators: resident population, gross income, average monthly earnings of employees under contract, municipal property tax, municipal tax on onerous property transactions, individual income tax, single circulation tax, value of payment operations, value of transactions for the sale of family housing, the scale-adjusted indicator showing whether the municipality is spending more or less on procurement than predicted, and the respective ranking of the municipality regarding that indicator. These indicators allow for a socio-economic characterization of the selected municipality. This section also includes a slicer with a search bar that allows users to select the municipality they want to analyse, enabling a customized and relevant examination.



Figure 9 - First Section of the Municipality Page

In the second section (Figure 10), the page mirrors the structure of the General Analysis page by presenting three informational cards. One card presents the name of the municipality that is being analysed, the other shows the total number of contracts signed by the selected municipality and its municipal companies, and the next card details the total expenditure on these contracts. Additionally, this section contains two vertically arranged bar charts. The first

bar chart highlights the companies with the highest number of contracts with the selected municipality, offering insights into key business relationships. The second bar chart lists the most common procurement procedures used by the municipality, providing an understanding of the prevalent procurement practices.

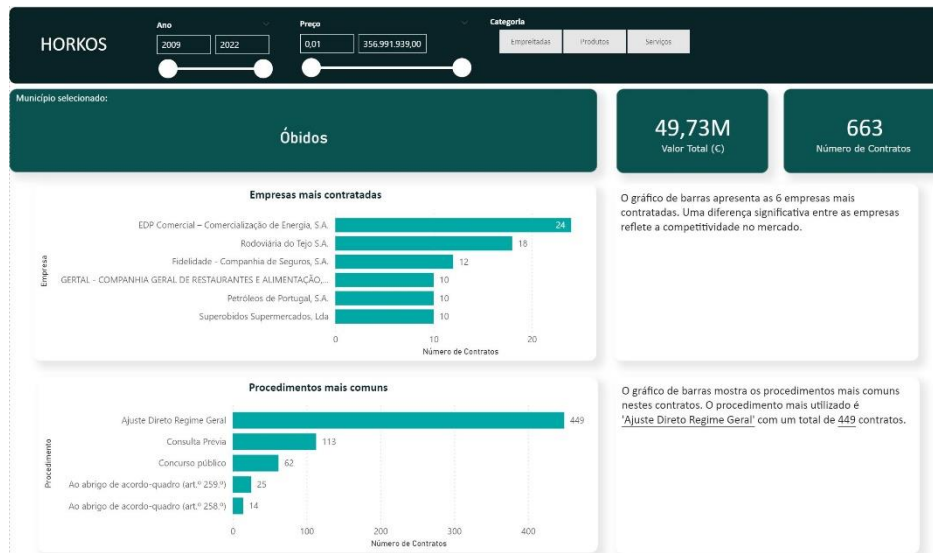


Figure 10 - Second Section of the Municipality Page

The third section (Figure 11) shifts to a more detailed analysis with two column charts that track trends over time. One chart displays the evolution in the number of contracts, and the other shows the evolution of the total expenditure. Blue dotted lines indicate the dates of municipal elections, adding context to the trends. Both charts feature a toggle button in the top right corner, allowing users to switch to a line chart view for a more detailed analysis of these trends by procurement category. At the bottom of the page, there is a treemap. The treemap provides a detailed view of the most frequent 1st-level CPV codes in the contracts awarded by the municipality. This analysis reveals diversity in the contracting areas by the municipality, with certain categories dominating in terms of contract volume. This information is crucial for understanding municipal investment priorities and resource management, highlighting where there is greater contractual activity and potential areas for improvement or expansion.

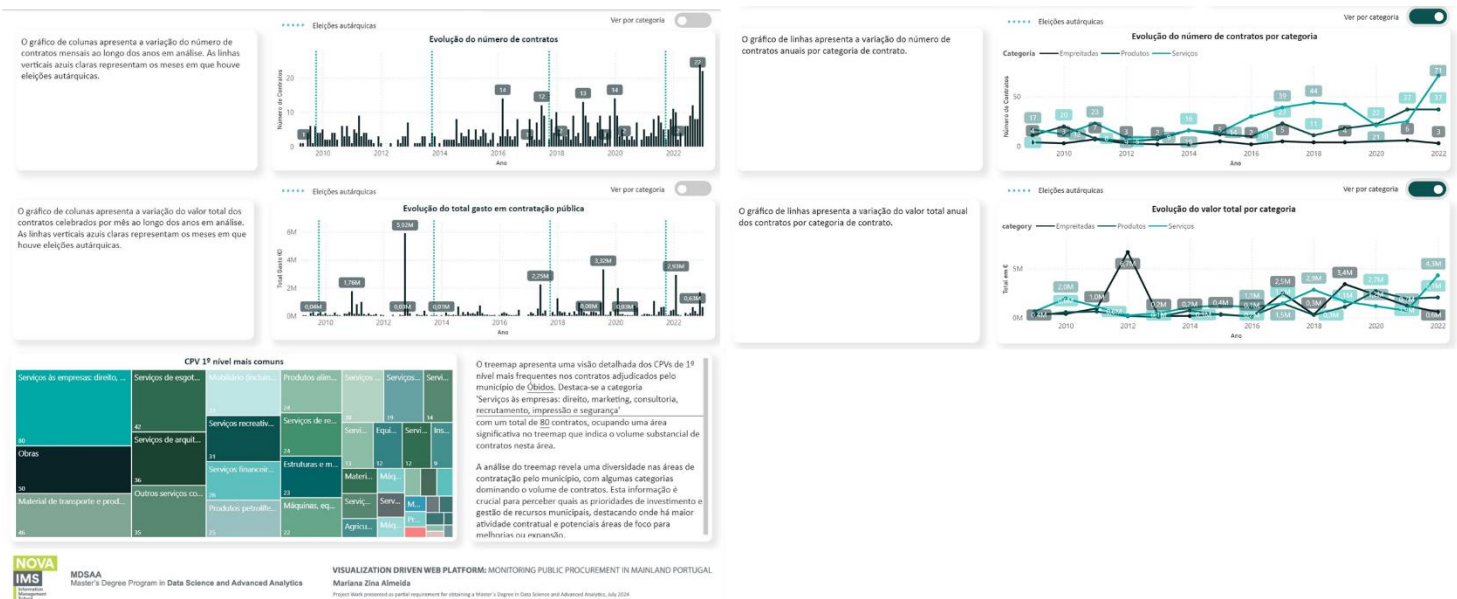


Figure 11 - Third Section of the Municipality Page

To improve user interactivity and analytical flexibility, several features have been added. A date slicer (Figure 10) lets users filter the data for any time frame between 2009 and 2022 or focus on a specific year. Additionally, a price slicer allows users to filter the data by any price range or focus on a specific price and a category slicer allowing users to filter the data by type of procurement.

These interactive elements and visualizations are designed to provide a thorough and flexible analytical experience, enabling users to explore and understand the public procurement activities of each municipality from multiple perspectives.

The company page (Figure 15) provides a detailed analysis of public procurement activities for each company. This page features the same set of visualizations as previous pages, along with three additional informational cards. These cards display the name of the company being analysed, the total number of contracts it has signed, and the total amount earned from these contracts, (Figure 13). Users can refine their analysis using available filters for date and company, allowing for targeted data exploration. The company slicer has a search bar to make searching easier and faster.

Following this, two vertically arranged bar charts are included. The first bar chart highlights the municipalities that contract with this company the most, providing insight into key municipal relationships. The second bar chart lists the most common procurement procedures used in contracts with this company, enlightening prevalent contracting methods. Below these charts, and similarly to the previous page, there is a treemap providing a detailed view of the most frequent 1st-level CPV codes that constitute the areas of activity of the selected company.

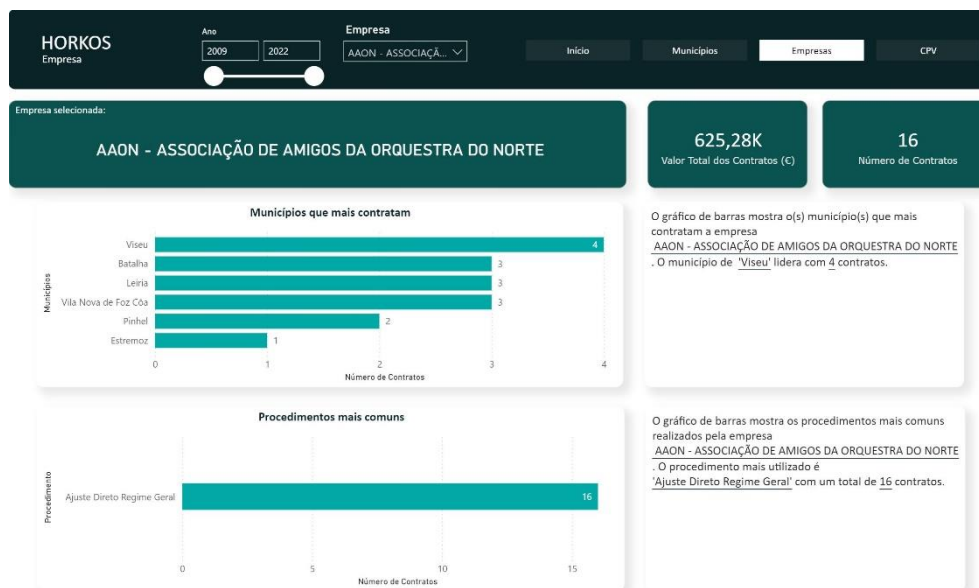


Figure 13 - First Section of the Company Page

In the next section of the page, two line charts track trends over time. One chart shows the evolution of the number of contracts awarded to this company, while the other illustrates the total amount earned from these contracts. This temporal analysis helps users understand how the company's procurement activities and earnings have changed over time.

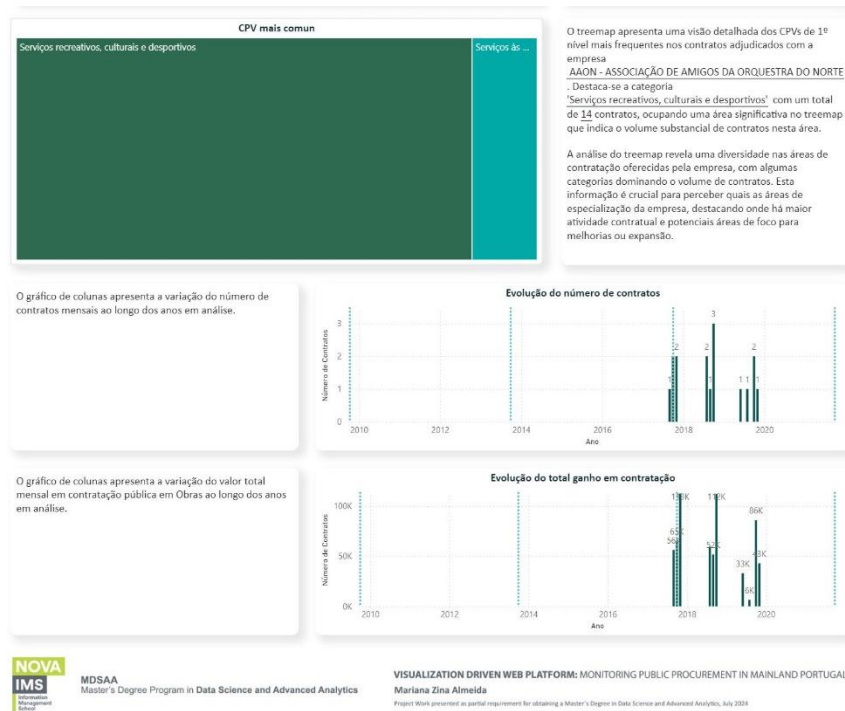


Figure 14 - Second Section of the Company Page

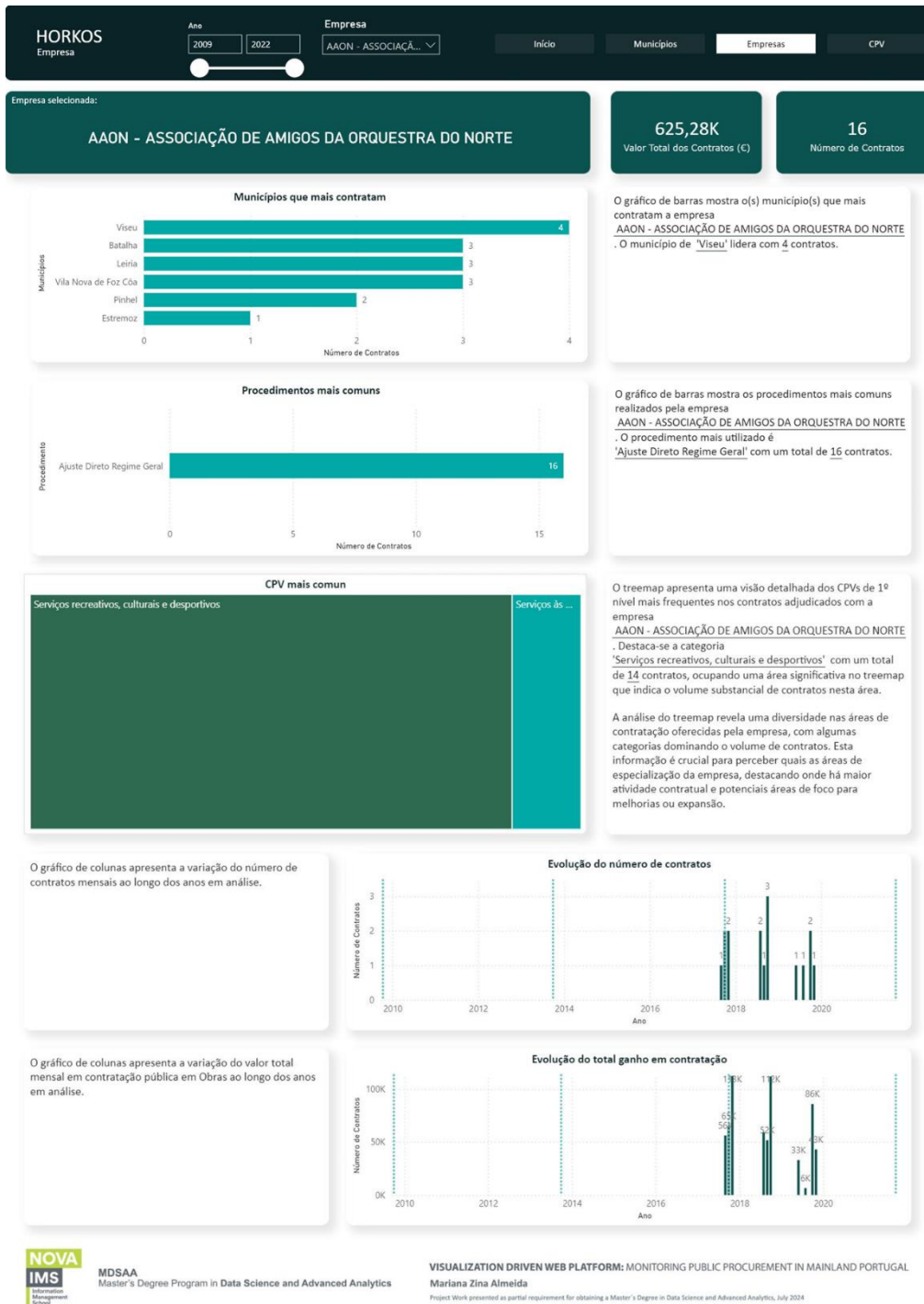


Figure 15 - Company Page

The CPV analysis page (Figure 18) provides a wide-ranging insight into the public procurement activities for every first-level CPV. Similar to previous pages, at the top, the user has a slicer where they can select a specific CPV to analyze, and the page will display a series of visualizations tailored to the selected CPV category. This page features a similar set of visualizations as previous pages. The page includes three primary bar charts for a deeper dive in procurement patterns (Figure 16): the first displays the municipalities that have awarded the most contracts within the selected CPV, offering insights into key municipal players and their procurement activities; the second bar chart showcases the most frequently used procurement procedures for contracts issued in the chosen CPV, giving an overview of what procurement procedures are most typical within this category; the last bar chart shows the companies that have won the most contracts in the selected CPV, offering a clear view of the leading suppliers in this procurement area.

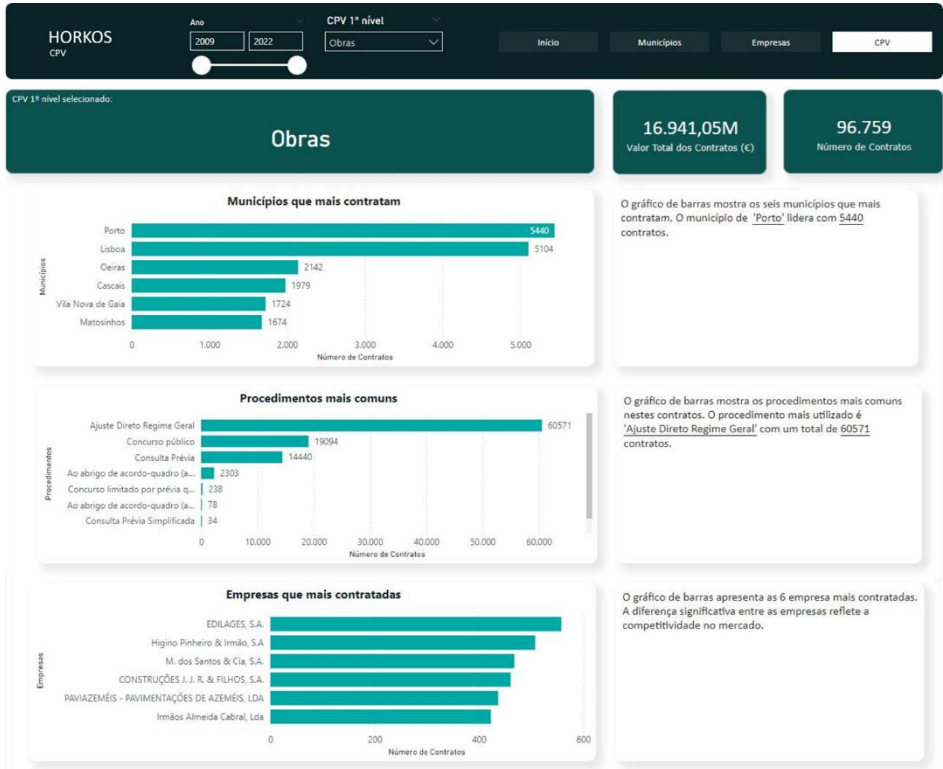


Figure 16 - First Section of the CPV Page

In the next section of the page (Figure 17), two time series column charts track trends over time: the first one tracks the evolution of the number of contracts awarded in the selected CPV, providing a temporal view of procurement activity; the second depicts the total expenditure associated with these contracts, allowing users to view how spending has evolved over time

within the selected CPV. These visualizations, along with an interactive filter for date, allow users to conduct targeted analyses of procurement activities, enabling them to gain valuable insights into public procurement trends and practices.

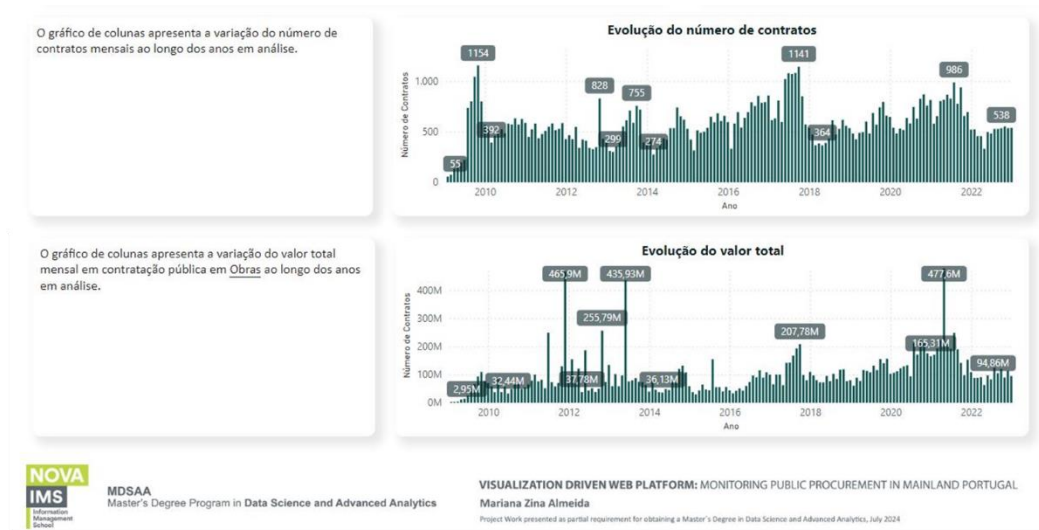


Figure 17 - Second Section of the CPV Page



4.3. EVALUATION

Assessing the acceptance and use of a new technology, in this case a new data analysis open-source platform, is important for ensuring its effectiveness, usability and integration (Davis, 1989). A study conducted by Almasi et al. (2023) identified and suggested a more specific set of usability evaluation criteria for dashboards, based on a literature review of existing questionnaires used for this purpose. The criteria presented is as follows: (i) *usefulness*, defined as satisfying a customer's needs or giving a competitive advantage with the product's features or benefits; (ii) *operability*, referring to a user's ability to operate and manage a dashboard to complete their tasks; (iii) *learnability*, a quality of software interface that allows users to quickly understand it and effectively use all its features and functions; (iv) *ease of use*, a fundamental concept describing how easily users can use a dashboard; (v) *suitability for tasks*, that helps determining if users can identify whether a product or system meets their needs; (vi) *improvement of situational awareness*, is about knowing what is happening and what could happen next; (vii) *satisfaction*, referring to the satisfaction with the features, functions and ease of use of a dashboard; (viii) *user interface*, includes visual and interactive tools such as color coding for data visualization, bar graphs, data labels, geographic maps, customizable searched, data sorting and filtering, etc.; (ix) *content*, concerning the amount and quality of data shown on a dashboard; and (x) *system capabilities*, that involves assessing how well software meets work-related needs (Almasi et al., 2023). This is the result of a combination of several questionnaires. Since this study was conducted regarding dashboards in healthcare, and in order to adapt it to this project, a questionnaire was created based on the suggestions of this study and based on the following frameworks: Technology Acceptance Model (TAM) (Davis, 1989) and Dashboard Assessment Usability Model (DATUS) (Antunes, 2020). The questionnaire consists of 20 questions rated with a Likert scale. Each item is rated on a scale from 1 (strongly disagree) to 5 (strongly agree) plus NA (Not Applicable) and is worth 5 points. Furthermore, there are two-open ended questions provided for participants to share their thoughts:

- 1. Is there any additional information you would like to see included on the dashboard?
- 2. Do you have any other feedback or suggestions you would like to share?

Following TAM, some questions were used for measuring perceived usefulness and following DATUS, some questions were used for measuring operability, suitability for tasks,

satisfaction, user interface, content and system capabilities. The proposed questions are summarized in table 7.

Table 5 - Questionnaire

#	Questions	Strongly disagree			Strongly agree		
		1	2	3	4	5	NA
PU1	HORKOS allows to analyze Portuguese public procurement data at a municipal level	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
PU2	HORKOS enhances transparency and accountability in public procurement at a municipal level	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
PEOU1	I find HORKOS clear and understandable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
PEOU2	Learning to operate HORKOS would be easy for me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
PEOU3	I find HORKOS to be flexible to interact with	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D1	Overall, I am satisfied with how easy it is to use this dashboard	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D2	It was simple to use this dashboard	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D3	I feel comfortable using this dashboard	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D4	It was easy to learn to use this dashboard	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D5	The information provided with this dashboard is clear	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D6	It was easy to find the information I needed	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D7	The information displayed in the dashboard is easy to understand	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D8	The organization of the information on the dashboard is clear	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D9	Whenever I make a mistake using the dashboard, I recover easily and quickly	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D10	The interface of this dashboard is pleasant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D11	I like using the interface of this dashboard	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D12	This dashboard has all the functions and capabilities I expect it to have	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D13	Overall, I am satisfied with this dashboard	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D14	Data on the dashboard is easy to read	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
D15	Visual encoding of data is consistent throughout the dashboard	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4.4. EVALUATION RESULTS AND DISCUSSION

The evaluation was completed with a sample of 10 Portuguese citizens. This sample comprises citizens with an average age of 39 years. Figure 19 shows the distribution of the respondents per age, gender, civil status and academic qualification.

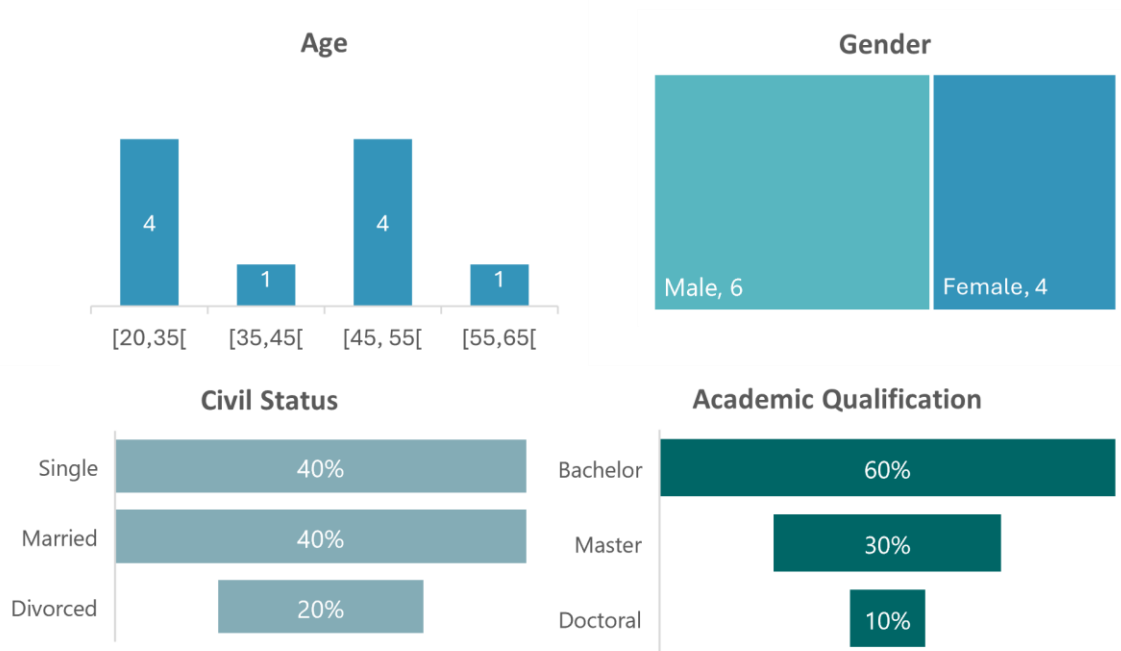


Figure 19 - Socio-demographic characterization of the sample

Based on the analysis of the sample, there is a diverse demographic spread. The age distribution shows that 80% of respondents are between 20-35 and 45-55 years old, with 40% in each age group. Gender-wise, the sample consists of 60% males and 40% females. Regarding civil status, 40% of participants are single and another 40% are married. In terms of academic qualifications, 60% have a bachelor's degree, 30% have a master's degree, and 10% have a doctorate.

Recalling the proposed questionnaires, TAM and DATUS, the platform's final grade will be calculated by summing the average scores of each question, resulting in a score on a scale from 0 to 100. The summarized results are presented in the following table:

Table 6 - Summary data of the questionnaire by question

Question ID	Mean	Median	Standard Deviation
PU1	4,70	5	0,437
PU2	4,90	5	0,286
PEOU1	4,70	5	0,437
PEOU2	4,50	4;5	0,477
PEOU3	4,80	5	0,381
D1	4,70	5	0,437
D2	4,50	4;5	0,477
D3	4,90	5	0,286
D4	4,60	5	0,467
D5	4,70	5	0,437
D6	4,44	4	0,471
D7	4,60	5	0,467
D8	4,20	4	0,714
D9	4,44	4	0,471
D10	4,80	5	0,572
D11	4,67	5	0,447
D12	3,80	4	0,934
D13	4,60	5	0,467
D14	4,80	5	0,381
D15	4,80	5	0,381
Total	92,16	-	-

The platform has a score of 92,16 out of 100.

The findings from the evaluation revealed several key insights. First, a great majority of the responses to the TAM questions indicate that most users found the platform useful for monitoring public procurement activities. This perception is important because it greatly influences whether or not people will use the system in future. The users appreciated the ability of the platform to visualize complex procurement data into easily interpretable formats, facilitating informed decision-making.

Concerning usability and user experience, the DATUS evaluations showed that the dashboard has been evaluated positively. Users found the interface of the dashboard pleasant, intuitive and straightforward, enabling them to explore data without extensive training. Each visual has a brief explanation of what is being conveyed in the graph. This feature ensured that even users without high analytical skills could understand and make the most of the platform.

The low value on question D12 reveals that users are not entirely satisfied with the platform's functionalities and would like to see more features implemented, as well as some new capabilities. This is why the last two open questions are so important. Regarding the DATUS open questions, some users left some feedback and suggestions for improving the platform. The users suggested:

- In addition to the text boxes beside the visuals, there could be an informative page explaining how the platform works so that users with low analytical skills can make the most of it, what it is for and with contacts.
- Adding indicators that would enable the data to be related to the political parties represented in the municipal executives.
- The first part of the homepage should display highlights of the results, with the remainder of the page presenting the full results for a more in-depth analysis.
- Having links to other similar websites.
- Having the capability to export the information to and excel or pdf file.
- Having the capability to compare municipalities or companies.

These suggestions will be considered for future work.

By making procurement data accessible and understandable, the visualization-driven web platform significantly enhances transparency. Citizens can scrutinize public procurement transactions and detect patterns, prevalent contracting methods and the diversity in the contracting areas by the municipality, for example. Although the platform does not provide real-time updates and the data used is already out of date (as the last year in the analysis is 2022), the vast information coverage helps to create a more open and transparent environment by reducing the secrecy that often surrounds public spending.

One key feature of the platform is its open-access nature, allowing any citizen to access the data without restriction. Such open access ensures that the platform is more than just a tool for policymakers and monitoring organizations, it also enables citizens to stay informed about public procurement activities. Therefore, the platform promotes a more inclusive and participatory approach to governance by allowing all community members to scrutinize, monitor and contribute to accountability transactions.

5. CONCLUSIONS AND FUTURE WORKS

Transparency and accountability have become essential in public procurement. The use and exploitation of OGDs are becoming increasingly important, and technologies are critical to support the proper application of these concepts. The development of visualization-driven web platforms provides an effective solution to improve comprehension and interaction with procurement data. Even though many governments may feel discouraged to use and disclose public procurement data, because of its association with the anti-corruption agenda, a public procurement data dashboard can contribute to strengthening the national procurement system when they are part of a larger country engagement and communication with government counterparts (World Bank Group, 2022). In this project, it was presented the “HORKOS - Observatório de Contratação Pública Municipal”, an open-access intuitive data visualization platform to help citizens to access and understand easily Portuguese public procurement data at the municipal level.

The development of the HORKOS platform has resulted in a comprehensive tool that offers insightful and user-friendly visualizations across four pages. The General Analysis page provides an overview of the data, designed to give users a broad understanding of the key metrics and trends. The Municipality oriented analysis page enables a detailed analysis of the data at the municipal level, allowing users to delve deeper into public procurement activities within each municipality, offering, for example, key insights into spending patterns. The Company oriented page focuses on procurement data related to contracted companies. Finally, the CPV analysis page breaks down procurement data by the first level of the CPV codes, enabling an understanding of the procurement activities across the various categories.

Regarding the research question: *How can a visualization-driven web platform significantly enhance the visibility of public procurement data and make analysis more accessible to citizens?*, the current work has achieved the contribution outlined in the introduction: the visualization-driven web platform met expectations in terms of improving visibility of public procurement data and making analysis more accessible to citizens. Its user-centred design, combined with strong data visualization capabilities, proved effective in providing citizens with actionable insights and encouraging a public participation in governance, as well as a more transparent and accountable procurement environment. As governments around the world embrace digital transformation, platforms like these are powerful instruments in promoting good governance practices and ensuring public trust in public expenditure management.

5.1.LIMITATIONS AND FUTURE WORK

Several limitations were identified during the course of this project. Firstly, the current version of the dashboard may not fully meet the needs of all stakeholders, including policymakers, researchers, and the general public. Each group has different requirements for the level of detail and specific features they need. Secondly, the data on this dashboard is already outdated as of the writing of this thesis, with the last year of analysis being 2022. Additionally, the dashboard relies on the accuracy of the data, so any inconsistencies or gaps can impact the reliability of the insights it provides.

Future work could address these limitations by improving data quality and incorporate the suggestions given by the users in the evaluation stage. This will be crucial to fine-tune the dashboard's interface and functionality so it can better meet their needs. To address the issue of outdated data, one potential direction is the incorporation of real-time data processing. The dashboard could be connected to a database that would update every time a public procurement transaction was published in BASE Portal. Furthermore, the present version of the dashboard only has two machine learning techniques and does not include advanced predictive analytics. The integration of these technologies could provide better insights and forecasts. Moreover, expanding the scope of the dashboard to include more detailed analysis would favor the intention to use the platform. Finally, regarding Power BI itself, the project encountered performance limitations. Power BI struggles to process more than 30,000 rows without running into issues in some queries, which necessitated the simplification of certain visuals related to contracted companies. To address this, one potential solution could be to optimize the data models by aggregating data at higher levels, thus reducing the number of rows processed in each query. Additionally, the use of alternative tools or platforms better suited for handling large datasets could be considered, to help mitigate these performance constraints.

BIBLIOGRAPHICAL REFERENCES

A European Strategy for Data (2020).

Agrahari, A., & Srivastava, S. K. (2019). A data visualization tool to benchmark government tendering process: Insights from two public enterprises. *Benchmarking*, 26(3), 836–853. <https://doi.org/10.1108/BIJ-06-2017-0148>

Almasi, S., Bahaadinbeigy, K., Ahmadi, H., Sohrabei, S., & Rabiei, R. (2023). Usability Evaluation of Dashboards: A Systematic Literature Review of Tools. In *BioMed Research International* (Vol. 2023). Hindawi Limited. <https://doi.org/10.1155/2023/9990933>

Ansari, B., Barati, M., & Martin, E. G. (2022). Enhancing the usability and usefulness of open government data: A comprehensive review of the state of open government data visualization research. *Government Information Quarterly*, 39(1). <https://doi.org/10.1016/j.giq.2021.101657>

Antunes, R. S. da S. (2020). *DATUS: Dashboard Assessment Usability Model-a case study with student dashboards*.

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, 32(4), 399–418. <https://doi.org/10.1016/j.giq.2015.07.006>

Barcellos, R., Bernardini, F., & Viterbo, J. (2022). Towards defining data interpretability in open data portals: Challenges and research opportunities. *Information Systems*, 106. <https://doi.org/10.1016/j.is.2021.101961>

Charalabidis, Y., Loukis, E., & Alexopoulos, C. (2014). Evaluating second generation open government data infrastructures using value models. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2114–2126. <https://doi.org/10.1109/HICSS.2014.267>

Cheng, T., Liu, T., Meng, L., & Wang, C. (2017). *The Analysis of Water Project Bid Rigging Behavior Based on Complex Network*.

Curado, A., Damasio, B., Encarnação, S., Candia, C., & Pinheiro, F. L. (2021). Scaling behavior of public procurement activity. *PLoS ONE*, 16(12 December). <https://doi.org/10.1371/journal.pone.0260806>

- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information. In *Source: MIS Quarterly* (Vol. 13, Issue 3).
- Eberhardt, A., & Silveira, M. S. (2018, May 30). Show me the data! A systematic mapping on open government data visualization. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3209281.3209337>
- Esteves de Oliveira, R., & Pinto Correia, C. (2022). *Country Comparative Guides 2022 - Portugal Public Procurement*.
- European Anti-Fraud Office. (2011). *Compendium of Anonymised Cases - Structural Actions*.
- European Commission. (2017). *Making Public Procurement work in and for Europe*. https://ec.europa.eu/commission/priorities/jobs-growth-and-investment/investment-plan_en
- European Commission. (2018). *Public Procurement Guidance For Practitioners*. <https://doi.org/10.2776/461701>
- European Union. (n.d.). *Main and Supplementary Vocabulary of CPV*. Retrieved April 9, 2024, from <https://ted.europa.eu/en/simap/cpv>
- Fazekas, M., & Tóth, I. J. (2016). From Corruption to State Capture: A New Analytical Framework with Empirical Applications from Hungary. *Political Research Quarterly*, 69(2), 320–334. <https://doi.org/10.1177/1065912916639137>
- Fazekas, M., Tóth, I. J., & King, L. P. (2016). An Objective Corruption Risk Index Using Public Procurement Data. *European Journal on Criminal Policy and Research*, 22(3), 369–397. <https://doi.org/10.1007/s10610-016-9308-z>
- Fazekas, M., & Wachs, J. (2020). Corruption and the network structure of public contracting markets across government change. *Politics and Governance*, 8(2), 153–166. <https://doi.org/10.17645/pag.v8i2.2707>
- Fernandes, M. J., Camões, P., & Jorge, S. (2022). *ANUÁRIO FINANCEIRO DOS MUNICÍPIOS PORTUGUESES 2022*.

- Ferwerda, J., Deleanu, I., & Unger, B. (2016). Corruption in Public Procurement: Finding the Right Indicators. *European Journal on Criminal Policy and Research*, 23(2), 245–267. <https://doi.org/10.1007/s10610-016-9312-3>
- Gil García, J. R., Helbig, N., Ojo, A., Ferriero, D. S., Pardo, T., Qian, H., Association for Computing Machinery, & ACM Digital Library. (2012). *Anti-Corruption Online Monitoring Systems in Brazil*.
- Gonzalez-Zapata, F., & Heeks, R. (2015). The multiple meanings of open government data: Understanding different stakeholders and their perspectives. *Government Information Quarterly*, 32(4), 441–452. <https://doi.org/10.1016/j.giq.2015.09.001>
- Graves, Alvaro., & Hendler, J. (2013). *Visualization Tool for Open Government Data*. 136–145.
- IMPIC. (2023). *Contratação Pública em Portugal 2022*. www.base.gov.pt
- INE. (n.d.). Retrieved March 30, 2024, from <https://www.ine.pt/xportal/ine/portal/portlets/html/conteudos/listaContentPage.jsp?BOUI=6251013&xlang=PT>
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, 29(4), 258–268. <https://doi.org/10.1080/10580530.2012.716740>
- Kenny, C., & Musatova, M. (2010). “Red Flags of Corruption” in World Bank Projects An Analysis of Infrastructure Contracts. <http://econ.worldbank.org>.
- Kimball, R., & Ross, M. (1996). *The Data Warehouse Toolkit*.
- Lee, J., Park, G.-L., Han, Y., & Yoo, S. (2017). Big data analysis for an electric vehicle charging infrastructure using open data and software. 252–253.
- Lyra, M. S., Curado, A., Damásio, B., Bação, F., & Pinheiro, F. L. (2021). Characterization of the firm–firm public procurement co-bidding network from the State of Ceará (Brazil) municipalities. *Applied Network Science*, 6(1), 1–10. <https://doi.org/10.1007/s41109-021-00418-y>

- Lyra, M. S., Damásio, B., Pinheiro, F. L., & Bacao, F. (2022). Fraud, corruption, and collusion in public procurement activities, a systematic literature review on data-driven methods. *Applied Network Science*, 7(1), 83. <https://doi.org/10.1007/s41109-022-00523-6>
- Marcos, I. A. C. (2023). *A Contratação Pública em Portugal e Espanha* [Instituto Politécnico da Guarda]. <http://hdl.handle.net/10314/8586>
- Maude, F. (2012). *Open data white paper : unleashing the potential*. The Stationery Office.
- Moreno, A., Molano-Pulido, J., Gomez-Morantes, J. E., & Gonzalez, R. A. (2022). ADACOP: A Big Data Platform for Open Government Data. *ACM International Conference Proceeding Series*, 369–375. <https://doi.org/10.1145/3560107.3560310>
- OECD. (2009). *OECD principles for integrity in public procurement*. OECD.
- OECD. (2015). *Government at a Glance 2015*. OECD. https://doi.org/10.1787/gov_glance-2015-en
- Oliveira, R., Cappelli, C., & Oliveira, J. (2023). Assessment of public information understanding using plain language and data visualization. *ACM International Conference Proceeding Series*, 228–234. <https://doi.org/10.1145/3614321.3614353>
- Rabuzin, K., & Modrušan, N. (2019). Prediction of public procurement corruption indices using machine learning methods. *IC3K 2019 - Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, 3, 333–340. <https://doi.org/10.5220/0008353603330340>
- Siciliani, L., Taccardi, V., Basile, P., Di Ciano, M., & Lops, P. (2023). AI-based decision support system for public procurement. *Information Systems*, 119. <https://doi.org/10.1016/j.is.2023.102284>
- Silva, B. M. T. de A. (2019). *Contração Pública Eletrónica e Procedimentos Por Convite Estudo de Caso na Assembleia da República*. <http://hdl.handle.net/10400.5/19106>
- Soylu, A., Corcho, Ó., Elvesæter, B., Badenes-Olmedo, C., Yedro-Martínez, F., Kovacic, M., Posinkovic, M., Medvešček, M., Makgill, I., Taggart, C., Simperl, E., Lech, T. C., & Roman, D. (2022). Data Quality Barriers for Transparency in Public Procurement. *Information (Switzerland)*, 13(2). <https://doi.org/10.3390/info13020099>

- Ubaldi, B. (2013). *Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives* “OPEN GOVERNMENT DATA: TOWARDS EMPIRICAL ANALYSIS OF OPEN GOVERNMENT DATA INITIATIVES.” <https://doi.org/10.1787/5k46bj4f03s7-en>
- Wachs, J., Fazekas, M., & Kertész, J. (2021). Corruption risk in contracting markets: a network science perspective. *International Journal of Data Science and Analytics*, 12(1), 45–60. <https://doi.org/10.1007/s41060-019-00204-1>
- Ware, G. T., Moss, Shaun., Campos, J. E., & Noone, G. P. (2007). Corruption in Public Procurement: A Perennial Challenge. In J. E. Campos & S. Pradhan (Eds.), *The Many Faces of Corruption: Tracking Vulnerabilities at the Sector Level* (pp. 295–334).
- Wensink, W., & Vet, J. M. de. (2013). *Identifying and Reducing Corruption in Public Procurement in the EU*. www.pwc.com
- World Bank Group. (2022). *Using Data Analytics in Public Procurement: Operational Options and a Guiding Framework*.
- Zuiderwijk, A., Janssen, M., & Sussha, I. (2016). Improving the speed and ease of open data use through metadata, interaction mechanisms, and quality indicators. *Journal of Organizational Computing and Electronic Commerce*, 26(1–2), 116–146. <https://doi.org/10.1080/10919392.2015.1125180>
- Zuiderwijk-van Eijk, A. M. Gerarda. (2015). *Open data infrastructures : the design of an infrastructure to enhance the coordination of open data use*.

APPENDIX A

Table 7 - CVP First Level Divisions (source: <https://www.publictendering.com/cpv-codes/list-of-the-cpv-codes/>)

id	Description
03	Agricultural, farming, fishing, forestry, and related products
09	Petroleum products, fuel, electricity, and other sources of energy
14	Mining, basic metals and related products
15	Food, beverages, tobacco, and related products
16	Agricultural machinery
18	Clothing, footwear, luggage articles and accessories
19	Leather and textile fabrics, plastic, and rubber materials
22	Printed matter and related products
24	Chemical products
30	Office and computing machinery, equipment and supplies except furniture and software packages
31	Electrical machinery, apparatus, equipment, and consumables; lighting
32	Radio, television, communication, telecommunication, and related equipment
33	Medical equipment, pharmaceuticals, and personal care products
34	Transport equipment and auxiliary products to transportation
35	Security, firefighting, police, and defence equipment
37	Musical instruments, sport goods, games, toys, handicraft, art materials and accessories
38	Laboratory, optical and precision equipment (excl. glasses)
39	Furniture (incl. office furniture), furnishings, domestic appliances (excl. lighting) and cleaning products
41	Collected and purified water
42	Industrial machinery
43	Machinery for mining, quarrying, construction equipment
44	Construction structures and materials; auxiliary products to construction (except electric apparatus)
45	Construction work
48	Software package and information systems
50	Repair and maintenance services
51	Installation services (except software)
55	Hotel, restaurant, and retail trade services
60	Transport services (excl. Waste transport)
63	Supporting and auxiliary transport services; travel agencies services
64	Postal and telecommunications services
65	Public utilities
66	Financial and insurance services
70	Real estate services
71	Architectural, construction, engineering, and inspection services
72	IT services: consulting, software development, Internet, and support
73	Research and development services and related consultancy services
75	Administration, defence, and social security services

- 76 Services related to the oil and gas industry
 - 77 Agricultural, forestry, horticultural, aquacultural and apicultural services
 - 79 Business services: law, marketing, consulting, recruitment, printing, and security
 - 80 Education and training services
 - 85 Health and social work services
 - 90 Sewage, refuse, cleaning, and environmental services
 - 92 Recreational, cultural, and sporting services
 - 98 Other community, social and personal services
-

APPENDIX B

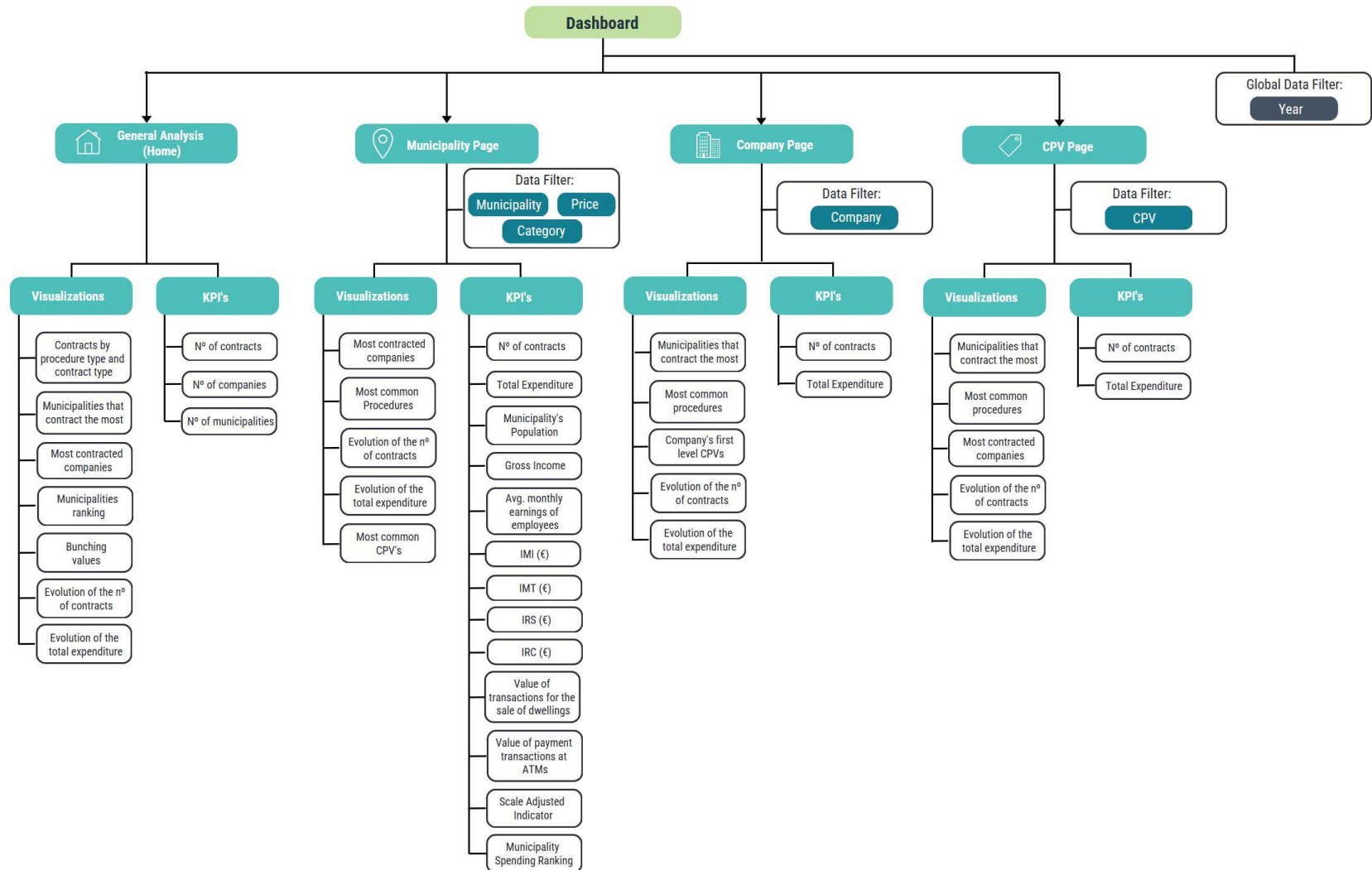


Figure 20 - Illustration of the flow of the pages and visualizations

