Advanced tool for aggregation and visualization of atomized data

Ana Teresa Feliciano da Cruz

U. PORTO

FEUP **FACULDADE DE ENGENHARIA**
UNIVERSIDADE DO PORTO

Mestrado em Engenharia Informática e Computação

Supervisor: Prof. João Correia Lopes

July 27, 2023

Advanced tool for aggregation and visualization of atomized data

Ana Teresa Feliciano da Cruz

Mestrado em Engenharia Informática e Computação

Approved in oral examination by the committee:

Chair: Prof. Rui Camacho
External Examiner: Prof. Benedita Malheiro
Supervisor: Prof. João Correia Lopes

July 27, 2023

Resumo

Hoje em dia, estabelecer uma presença *online* tornou-se essencial para quase todas as organizações. Um aspeto crucial da gestão de um negócio reside na compreensão e avaliação do seu desempenho, o que depende da obtenção de *insights* a partir da análise de dados. Esses servem como base para a tomada de decisões informadas, garantindo a qualidade e eficiência das estratégias empresariais. Consequentemente, a recuperação, agregação, análise e visualização de dados provenientes de várias fontes tornaram-se indispensáveis para empresas de diversas indústrias. No entanto, a abundância de dados heterogéneos provenientes de múltiplas fontes apresenta desafios para as organizações, incluindo a ZOS LDA, no aproveitamento eficaz do poder dos seus dados para a tomada de decisões.

Ferramentas existentes como o PowerBI ou Tableau oferecem recursos de recuperação, análise e visualização de dados. No entanto, frequentemente sofrem com a complexidade e altos custos, tornando-os inacessíveis para muitas empresas que não dispõem de orçamentos adequados ou recursos de análise.

Este trabalho propõe uma ferramenta de fácil utilização para preencher a lacuna entre os dados e os gestores, facilitando a toma de decisões orientada por dados. Um protótipo foi desenvolvido para validar a solução proposta, permitindo que os utilizadores obtenham dados de fontes manuais (*SpreadSheets*) ou fontes automáticas (Application Programming Interface). Os dados são intuitivamente representados por meio de gráficos, tabelas e *dashboards* personalizáveis, fornecendo uma visão geral das métricas selecionadas e *dashboards* dedicadas para cada métrica.

Uma experiência foi realizada para avaliar a usabilidade e a intuitividade do protótipo. Os resultados demonstraram que a ferramenta representou efetivamente as informações e foi fácil de usar, não exigindo conhecimento tecnológico. No entanto, fragilidades identificadas sugerem possíveis melhorias.

Em conclusão, esta pesquisa valida a hipótese de que é viável criar uma ferramenta com uma curva de aprendizagem pequena, que as organizações possam utilizar, a qual é alimentada automaticamente com dados provenientes de múltiplas fontes, e a análise fornecida será útil na toa de decisões empresariais. Os resultados deste estudo destacam o potencial valor da ferramenta proposta ao capacitar as empresas a tomar decisões informadas com base nos seus dados.

Abstract

In these modern business days, establishing an online presence has become essential for nearly every organization. A crucial aspect of managing a business lies in understanding and assessing its performance, which relies on deriving actionable insights from data analysis. These insights serve as the foundation for informed decision-making, ensuring the quality and efficiency of business strategies. Consequently, retrieving, aggregating, analyzing, and visualization of data from various sources have become indispensable for businesses across industries. However, the abundance of heterogeneous data from multiple sources poses challenges for organizations, including ZOS LDA, in effectively harnessing the power of their data for decision-making.

Existing tools like PowerBI or Tableau offer data retrieval, analysis, and visualization capabilities. However, they often suffer from complexity and high costs, rendering them inaccessible to many businesses lacking adequate budgets or analytics resources.

This work proposes a user-friendly tool to bridge the gap between data and managers, facilitating data-driven decision-making. A prototype was developed to validate the proposed solution, enabling users to retrieve data from manual sources (spreadsheets) or automatic sources (API). The data is intuitively represented through charts, tables, and customizable dashboards, providing an overview of selected metrics and dedicated dashboards for each metric.

An experiment was conducted to evaluate the usability and intuitiveness of the prototype. The results demonstrated that the tool effectively represented information and was easy to use, requiring no specialized technological knowledge. However, identified areas for improvement suggest potential enhancements.

In conclusion, this research validates the hypothesis that it is feasible to make a tool with a small learning curve that every organization can use, which is automatically fed with data from multiple sources, and the analysis it provides will be helpful when making business decisions. The findings of this study highlight the potential value of the proposed tool in empowering businesses to make informed decisions based on their data.

Keywords: web app, data visualization, data analysis, data aggregation, business intelligence, data-driven decision-making

Acknowledgments

I am filled with deep gratitude as I reflect on my college journey, and I owe this profound appreciation to all those who have supported me along the way. Their unwavering support and invaluable guidance have played an instrumental role in shaping my path; I am truly and profoundly grateful.

First and foremost, I extend my heartfelt thanks to my fantastic supervisor, João Correia Lopes. Your patience, care, and unwavering belief in my abilities have driven this work. I am sincerely grateful for the countless hours you dedicated to our weekly meetings, providing me with invaluable guidance and sharing your expertise.

I would also like to express my deepest gratitude to the entire team at ZOS. The company not only offered me an exceptional work environment but also provided essential support and access to the technologies that were pivotal to the success of this work. My heartfelt appreciation goes to Pedro Dias, Marco Sousa, and Rui Cruz, who stood by me throughout this journey. In particular, I sincerely thank Rui Cruz, whose willingness to answer my never-ending questions proved immeasurable. Your support, both technically and emotionally, played an irreplaceable role.

To my dear friends I made on this journey, mere words fail to convey my gratitude. Throughout these five years, your unwavering support has been a constant source of strength and inspiration. Thank you for standing by my side, offering your patience, encouragement, and unwavering belief in my abilities. The laughter and smiles we shared provided much-needed joy during challenging times.

I want to express my heartfelt appreciation to my closest friends, André Nascimento and Beatriz Santos. I'm very grateful for our friendship and know that most people do not have the luck that I have to have not one but two amazing best friends.

Por fim, quero agradecer à minha família. Em particular, obrigada aos meus pais pelo apoio incondicional e por me acompanharem neste percurso com o maior orgulho. Obrigada à melhor irmã e amiga do mundo, ao meu cunhado e ao meu irmão emprestado. Obrigada ao meu sobrinho, os momentos contigo eram os que maior alegria me traziam durante esta última etapa desta jornada. Obrigada à minha avó Teresa e ao Duarte por terem sempre lá o meu quartinho pronto, comidinha boa e bons momentos com muita alegria. Obrigada à minha avó Celeste pelo carinho, amor e muitos momentos de ternura.

Dedico este trabalho e todo o meu percurso à minha estrelinha, o meu avô Hernâni. Apesar de não teres cá estado para ver, sei que estás muito orgulhoso. Dedico também à minha tia Conceição que apesar de querer uma médica na família, sei que está muito orgulhosa por ter mais uma engenheira.

Agradeço a todos do fundo do coração. Thank you.

Ana Teresa Feliciano da Cruz

"If you let lesser people determine your self-worth, you'll never reach higher than their limited imagination."

Ana Huang

Contents

List of Figures

List of Tables

Abbreviations

ACID   Atomicity Consistency Isolation Durability
API   Application Programming Interface
BSON   Binary JavaScript Object Notation
BI   Business Intelligence
DDDM   Data-Driven Decision-Making
DOM   Document Object Model
ETL   Extract Transform Load
HTTP   Hypertext Transfer Protocol
JSON   JavaScript Object Notation
JWT   JavaScript Object Notation Web Tokens
NPM   Node Package Manager
OLAP   Online Analytical Processing
RAM   Random Access Memory
SUS   System Usability Scale
UML   Unified Modeling Language
URL   Uniform Resource Locator

Chapter 1

Introduction

This chapter introduces the work, clarifying its context, motivation, and goals.

1.1 Context

This work explores the need for improved business performance assessment in ZOS LDA[1]. This well-known company specializes in football and other sports data and information services. With most businesses having an online presence, the retrieval, aggregation, and storage of large volumes of data from various sources have become essential for effective decision-making.

Nowadays, the amount of data is exponentially growing. This data includes structured and unstructured data [27]. This enormous amount of heterogeneous data makes assessing business performance complex. If not approached efficiently, this process can be time-consuming and inefficient.

1.2 Motivation

The motivation behind this work arises from the realization that an automated tool capable of retrieving data from multiple sources, conducting extensive analysis, and presenting the findings intuitively and user-friendly can significantly assist managers in assessing business performance and making informed decisions.

While existing tools offer similar functionalities to this work, they often come with limitations such as complexity and high costs. This creates an opportunity to develop a more suitable and superior solution that caters specifically to the needs of managers.

---

[1] https://www.zos.pt/

1.3 Goals and Results

The main goal of this work is to develop an automated tool that streamlines the data retrieval process from multiple sources and facilitates its storage. The tool will delegate managers to define and select relevant data inputs and metrics for assessing business performance. Once the data is retrieved, the tool will perform data aggregation and analysis to assign values to each metric. The conclusion of this process will be the presentation of the aggregated information, including data sources, metrics, and their corresponding values, through an intuitive data visualization tool in the form of a dashboard.

The expected outcome of this research is to create a user-friendly and intuitive tool that requires minimal technical knowledge to operate effectively. The tool will adapt to various businesses' needs, accommodating different data inputs and metrics. By providing managers with a straightforward and accessible means to analyze and interpret data, the tool will facilitate informed decision-making based on accurate and up-to-date information.

1.4 Dissertation Structure

Besides this Introduction chapter, this document has seven other chapters.

In Chapter 2, the fundamental concepts of this work are explained. It explores the core principles of Business Intelligence (BI), including its various definitions and essential processes. Furthermore, it highlights the significance of data-driven decision-making and the crucial role of BI tools in facilitating this process.

In Chapter 3, there is the State of The Art investigation with existing BI tools and their comparison. It focuses on prominent tools such as PowerBI, Tableau, and Qlik Sense, selected for widespread adoption and recognition. A detailed comparative analysis is presented to evaluate these tools' features, capabilities, and limitations.

Chapter 4 describes the problem, including a testimony of the company ZOS LDA on the problem, the requirements specification including actors, user stories, and technical requirements. The chapter ends with the formulation of hypotheses to guide the research.

Chapter 5 describes the design and development process of the prototype. It discusses the chosen technologies, the architectural framework, and the data modeling considerations. Furthermore, the chapter presents the rationale behind design decisions, followed by a detailed breakdown of the frontend implementation, including user interface design and implemented pages, and the backend implementation, covering aspects such as authentication, authorization, API specifications, and data handling routines.

Chapter 6 describes the methodology used for conducting experiments. It describes the experimental setup, procedures, data collection methods, and the criteria for evaluating the results. Additionally, the chapter provides insights into the form utilized. Lastly, the limitations that impacted the experiment are discussed.

Chapter 7 analyzes the experimental results and their discussion. The findings are examined in light of the research objectives (testing usability). The chapter also includes a comprehensive discussion of the obtained results, providing valuable insights and addressing any discrepancies or unexpected outcomes.

Lastly, Chapter 8 presents the conclusions and future work.

Chapter 2

Background

This chapter explains some fundamental concepts used in this dissertation, including business intelligence fundamentals and data-driven decision-making.

2.1 Business Intelligence

The term Business Intelligence (BI) was first used by Luhn in an IBM journal article back in 1958. He described it as:

> "*Business is a collection of activities carried on for whatever purpose, be it science, technology, commerce, industry, law, government, defense, et cetera. The communication facility serving the conduct of a business (in the broad sense) may be referred to as an intelligence system. The notion of intelligence is also defined here, in a more general sense, as "the ability to apprehend the interrelationships of presented facts in such a way as to guide action towards a desired goal. [15]*"

Later, in 1989, Howard Dresner of Gartner Research, considered to be the father of BI, defined BI as:

> "*a broad category of software and solutions for gathering, consolidating, analyzing and providing access to data in a way that lets enterprise users make better business decisions [5]*"

The definition of BI has yet to reach a consensus, as seen in Chee et al. article, that presents many definitions for this term [5].

For this work, the definition most appropriate is of Dresner, with the understanding of Negash as well, about BI systems:

> "*BI systems combine data gathering, data storage, and knowledge management with analytical tools to present complex internal and competitive information to planners and decision-makers. [18]*"

This means that the purpose of business intelligence systems is to provide actionable information at the right time and in the proper form, therefore facilitating the managers' performance analysis [18] and all other levels of the business hierarchy, ultimately leading to a competitive advantage [29]. BI systems transform data into information and knowledge, leading to more effective decision-making. The BI system's role in decision-making can be seen in Figure 2.1 [19].
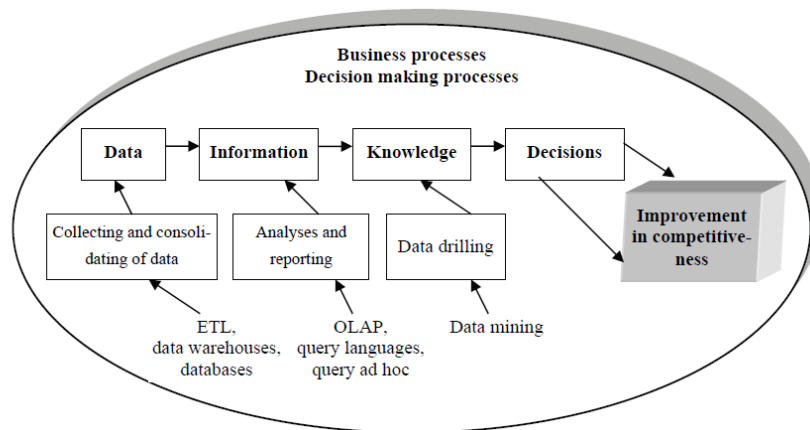
Figure 2.1: BI systems' role in decision making
Source: "Approach to building and implementing business intelligence systems" [19]

A BI system comprises three complementary elements: data management, analytical processing, and knowledge discovery [6]. These elements can be further broken down into more detailed modules [19]:

- Data Management — Data management involves the processes and tools for collecting, storing, and organizing data. The key modules within data management include:

    - ETL (Extract, Transform, Load) Tools — These tools extract data from various source systems, transform it into a consistent format, and load it into the data warehouse.

    - Data Warehouses — Data warehouses are central repositories for structured and organized data from different sources. They provide a unified view of the organization's data, facilitating efficient data retrieval, analysis, and reporting.

- Analytical Processing — Focuses on accessing and analyzing data to derive insights and support decision-making. The main modules within analytical processing include:

    - OLAP (Online Analytical Processing) Tools — OLAP tools enable users to access, analyze, and model business problems based on the data stored in the data warehouse. They provide multidimensional analysis capabilities, allowing users to explore data from different perspectives and levels of granularity.

– Data Mining Tools — Data mining tools utilize statistical and machine learning techniques to discover patterns, relationships, and insights within large datasets. These tools help uncover hidden knowledge and support predictive analytics.

- Knowledge Discovery — Knowledge discovery involves the extraction of valuable insights and actionable knowledge from data. The primary modules within knowledge discovery include:

  – Reporting and *Ad Hoc* Inquiry Tools — These tools enable the creation and utilization of different reports, including predefined and *ad hoc* reports. They allow users to query and explore data in a structured manner.

  – Presentation Layers — Presentation layers provide users with easy and comfortable access to information through graphical and multimedia interfaces. They facilitate data visualization, interactive dashboards, and user-friendly reporting interfaces.

2.2 Data-Driven Decision-Making

Data-driven decision-making (DDDM) is a crucial approach in business to mitigate the risks associated with decisions based solely on intuition or subjective feelings. Instead, DDDM emphasizes utilizing factual information, metrics, and data to inform decision-making processes that align with an organization's goals and objectives [28].

The foundation of DDDM lies in recognizing that intuitive decision-making may not always yield desirable outcomes in the business context. By relying on facts, metrics, and data, decision-makers can access valuable insights that enable informed and objective decision-making. This necessitates collecting relevant data, its comprehensive analysis to unveil patterns and trends, and utilizing these insights as the basis for decision-making.

Adopting a data-driven approach in businesses offers many advantages that contribute to improved operational efficiency, strategic decision-making, and enhanced customer satisfaction [2]. According to Forbes [7], the following benefits have been identified:

- Cost savings through data automation — By leveraging automated data processes, businesses can reduce manual efforts and associated costs in data collection, analysis, and reporting. Automation streamlines data workflows, enabling employees to focus on higher-value tasks.

- Establishment of a common language across teams — Embracing a data-driven culture fosters a shared understanding and promotes a common language around data within an organization. This facilitates effective communication and collaboration across teams.

- Enhanced knowledge sharing and collaboration — With data as a central point of reference, individuals within the organization can readily share their insights, analyses, and findings.

It is essential to acknowledge that while adopting a data-driven approach brings significant benefits, organizations often encounter challenges in their journey toward becoming data-driven. In today's competitive landscape, many organizations recognize the value of leveraging data to gain a competitive advantage [2]. However, achieving this transformation is difficult, mainly due to the sheer volume and complexity of the data. Implementing DDDM requires more than just technological solutions: it necessitates a fundamental shift in organizational culture. Building a data-driven culture entails fostering trust in data and recognizing its critical role in decision-making. This cultural change often requires organizational restructuring and a shift in mindset. Unfortunately, many organizations prioritize technology implementation without adequately addressing the underlying data challenges [13].

Organizations may encounter various hurdles when embarking on the journey toward data-driven decision-making. These challenges include ensuring data quality, acquiring and retaining skilled data professionals, and addressing data privacy and security concerns [2]. Overcoming these challenges requires the implementation of robust data governance practices, investment in data infrastructure, and a commitment to continuous skill development among employees.

After addressing the challenges of adopting a data-driven approach, it is essential to consider the steps involved in making effective data-driven decisions. Tableau, a leading business intelligence tool provider described in Section 3.3, proposes a framework of six key steps to guide organizations in their data-driven decision-making process [28].

1. Identify business objectives — The first step is understanding the organization's goals, like increasing sales. This is essential for choosing key performance indicators (KPI) and metrics later. Organizations can identify relevant data sources and formulate insightful questions by aligning the analysis with key business objectives.

2. Survey business teams for key data sources — The second step is to gather input from various stakeholders across the organization to ensure the success of data-driven decision-making. This step involves engaging individuals from different departments to understand short-term and long-term goals. The inputs collected inform the analysis questions and help prioritize certified data sources. By considering these inputs, organizations can shape their analytics deployment, define roles and responsibilities, establish data architecture and processes, and establish success metrics.

3. Collect and prepare the needed data — This step involves identifying and preparing relevant data sources. Organizations should prioritize data sources with a broad audience impact to deliver immediate value. By centralizing data sources and leveraging analytics platforms, organizations can streamline data preparation, enabling the creation of high-impact dashboards.

4. View and explore data — Data visualization is crucial in DDDM. Organizations can effectively influence stakeholders by visually representing insights. Various visual elements, such as charts, graphs, and maps, enable users to observe trends, outliers, and patterns within

the data. Selecting appropriate visualization types helps convey information accurately and engage users.

5. Develop insights — Critical thinking and data analysis go hand in hand when deriving valuable insights. By leveraging visual analytics, organizations can ask relevant questions and discover opportunities or risks that impact decision-making.

6. Act on and share insights — Once insights are obtained, taking action or sharing them with relevant stakeholders for collaboration is crucial. One effective method is sharing dashboards that highlight vital insights through informative text and interactive visualizations. This approach empowers decision-makers to make well-informed choices in their daily work and influences their decision-making processes positively.

Organizations can effectively leverage data-driven decision-making to improve business outcomes by following these six steps. Business intelligence tools like Tableau are vital in facilitating each process step, from data gathering and exploration to data visualization and monitoring. These tools provide intuitive interfaces, advanced analytics capabilities, and interactive visualizations, enabling users to make sense of complex data and derive actionable insights.

In summary, DDDM offers a systematic and objective approach to decision-making in the business domain. By leveraging data and insights, organizations can make informed decisions that drive superior outcomes and propel them toward their strategic objectives. The adoption of DDDM represents a crucial step for businesses seeking to capitalize on the potential of data in driving success and sustainability. Additionally, business intelligence tools like Tableau or PowerBI can facilitate data-driven processes.

Chapter 3

State of The Art

This chapter introduces the tools market-leading companies provide, namely Power BI, Tableau, and Qlik Sense. Each tool was examined regarding its key features, advantages, and disadvantages. Furthermore, a comparative analysis of these tools is presented, enabling a better understanding of their similarities and differences.

3.1 Introduction

Nowadays, many business intelligence and data visualization tools are available, catering to different needs and scenarios. However, due to the vast number of tools available, it was not feasible to analyze each one in the scope of this work. Therefore, this study examined some of the field's most popular and widely used tools.

Gartner, Inc.[1] is a renowned global research and advisory company that offers valuable insights, advice, and analysis on a wide range of technology-related topics to assist organizations in making informed business decisions. One of Gartner's well-known methodologies is the Magic Quadrants. These quadrants serve as a research framework for assessing and visually representing the relative positions of technology vendors within specific markets. By plotting vendors on a two-dimensional chart, the Magic Quadrants enable organizations to comprehend the competitive landscape, evaluate strengths and weaknesses, and make informed decisions while selecting technology solutions. The vertical axis of the Magic Quadrant represents the company's ability to execute, while the horizontal axis represents the completeness of its vision. Market leaders are positioned favorably in both aspects.

In the report dated 5th April 2023 on analytics and Business Intelligence Platforms, Gartner identified Microsoft, Salesforce (Tableau), and Qlik as market leaders in their Magic Quadrant [8]. This information is visualized in Figure 3.1, providing an overview of the market leaders in the field.

---

[1] https://www.gartner.com/en

Figure 3.1: Gartner's magic quadrant for analytics and business intelligence platforms
Source: Gartner [8]

3.2 Power BI

Power BI[2] is a robust business intelligence tool developed by Microsoft in 2014. It enables users to transform diverse and unrelated data from various sources into actionable insights. Initially introduced as an add-on to Microsoft Excel, Power BI has evolved into a standalone platform that provides a comprehensive suite of tools for data extraction, transformation, modeling, and visualization [24].

The main elements of Power BI include [16]:

- Power BI Desktop — Power BI Desktop is a Windows application that allows users, typically analysts or data professionals, to create rich and interactive reports and visualizations. It provides a range of features such as Power Query for data extraction and transformation, Power Pivot for data modeling and analysis, and Power View and Power Map for data visualization. Power BI Desktop offers advanced capabilities and flexibility in designing and customizing reports.

- Power BI Service — Power BI Service, also known as Power BI Online or Power BI Cloud, is a cloud-based platform that hosts and shares Power BI reports, dashboards, and datasets. Users can publish their reports created in Power BI Desktop to the Power BI Service, enabling easy access and collaboration across teams. The Power BI Service provides additional features like data refresh, scheduled data updates, and sharing and collaboration

---
[2]https://powerplatform.microsoft.com/en-us/

functionalities, allowing users to interact with reports and dashboards from anywhere using a web browser.

- Power BI Mobile — Power BI Mobile enables users to access Power BI reports and dashboards on mobile devices like smartphones and tablets. It provides a responsive and intuitive interface, allowing users to view, explore, and interact with data while on the go.

- Power BI Report Builder — Power BI Report Builder is a standalone Windows application that allows users to create paginated reports. These reports are typically used for structured, printable, and pixel-perfect reporting requirements. Power BI Report Builder provides a rich set of design features and capabilities, such as table-based layouts, support for charts, and precise control over report elements.

- Power BI Report Server — Power BI Report Server is an on-premises solution that allows organizations to deploy Power BI reports within their infrastructure. It provides secure and controlled access to reports behind the organization's firewall. Power BI Report Server enables users to create, publish, and manage Power BI reports within their environment while leveraging the capabilities of Power BI Desktop.

A typical usage scenario involves an analyst using Power BI Desktop to create and design reports by connecting to various data sources, performing data transformations, and creating visualizations. Once the report is ready, it can be published to the Power BI Service, where it can be shared with other users within the organization. Users can then access and interact with the reports on their devices through the Power BI Service or the Power BI Mobile. While this scenario is common, it is worth noting that reports can also be created and edited directly within the Power BI Service itself, providing additional convenience and accessibility [16].

Power BI offers several key features that make it a popular choice for business intelligence [24]:

- Data Connectivity — Power BI supports many data sources, both on-premises and cloud-based. Users can connect to various data sources such as Excel, Google Analytics, and Web API, enabling them to gather data from multiple systems and consolidate them for analysis and visualization.

- Data Visualization — Power BI provides rich visualization options, including charts, scattergrams, maps, and more. Users can create visually appealing and interactive reports and dashboards to represent their data effectively.

- Pre-built Dashboards/Reports — Power BI offers a collection of pre-built templates and sample dashboards, enabling users to start their data analysis quickly. These templates cover various business scenarios and industries, providing a head start in creating informative dashboards and reports.

- Collaboration — Power BI supports collaboration within organizations. Users can share reports and dashboards with colleagues, allowing collaboration and knowledge sharing. Features such as commenting and alerts enhance communication and enable users to stay updated on the latest insights.

- Natural Language Processing — Power BI incorporates a built-in Q&A feature that utilizes natural language processing. Users can interact with the tool using plain language queries, allowing for a more intuitive and conversational approach to data exploration and analysis.

Despite its strengths, Power BI does have some limitations and disadvantages to consider [24]:

- Complexity — While Power BI offers a user-friendly interface, some advanced features and capabilities may require a learning curve. Users may need to invest time in understanding and learning the more complex functionalities of the tool.

- Limited Visualizations — Although Power BI provides a wide range of visualizations, it may be more limited than other specialized data visualization tools. Users seeking highly specific or advanced visualizations may find that Power BI's options are less extensive than those offered by dedicated visualization tools.

- Limited Capacity for Big Datasets — Power BI was not designed explicitly for handling big data. While it can handle large datasets to a certain extent, performance issues may arise when dealing with extremely large or complex data volumes.

On the Gartner website [11], Power BI has received a rating of 4.4 out of 5, indicating high user satisfaction. Approximately 83 % of users reviewing Power BI would recommend the tool to others. User reviews on Gartner provide valuable insights into the strengths and weaknesses of Power BI. In the "Likes" section, many users praise the tool's robust capabilities and ability to handle multiple data sources. They appreciate the wide range of visualizations available, allowing them to create compelling and informative reports and dashboards. However, users have highlighted some common issues and challenges in the "Dislikes" section. Complexity, performance, and cost are among the main concerns mentioned. Some users find Power BI to have a steep learning curve, particularly individuals new to data analysis and visualization. Performance issues can arise when dealing with large datasets or complex queries, impacting the responsiveness of the tool. Additionally, a few users have raised concerns about the cost of Power BI, particularly for organizations with many users or when additional premium features are required.

In conclusion, Power BI is a powerful business intelligence tool developed by Microsoft that offers a range of features for data extraction, transformation, analysis, and visualization. With its extensive data connectivity options, diverse visualization capabilities, and collaboration features, Power BI enables users to gain valuable insights from their data and share them with colleagues. However, it is essential to consider the tool's complexity, potential performance limitations with large datasets, and the associated costs. Despite these challenges, Power BI remains a popular choice among users, as evidenced by its positive user ratings and recommendations.

3.3 Tableau

Tableau[3] is a powerful and widely used business intelligence and data visualization tool that was founded in 2003. It empowers users to explore, analyze, and visualize data from multiple sources to gain valuable insights and make data-driven decisions. Tableau offers a user-friendly and intuitive interface, making it accessible to technical and non-technical users [23].

Tableau provides a range of products and services designed to cater to different user needs. Some of these are [22]:

- Tableau Desktop — Tableau Desktop is the authoring and development tool that allows users to create interactive visualizations, reports, and dashboards. It provides a drag-and-drop interface, making it easy to connect to data sources, perform data blending and transformation, and create visualizations using a variety of charts, graphs, and maps. Tableau Desktop offers advanced analytics features, such as calculations, statistical functions, and forecasting capabilities, allowing users to uncover insights from their data.

- Tableau Server — Tableau Server is a platform that allows users to publish and share Tableau workbooks, dashboards, and visualizations with others in the organization. It provides a centralized repository for storing and managing Tableau content, ensuring users can access the most up-to-date and relevant data. Tableau Server enables collaboration, data governance, and security controls, making it easier for teams to collaborate and share insights across the organization.

- Tableau Cloud — Tableau Cloud, the new Tableau Online, is a cloud-based version of Tableau Server. It offers similar functionality to Tableau Server, allowing users to publish, share, and collaborate on Tableau content through a web browser. Tableau Online eliminates the need for organizations to manage infrastructure and server maintenance, providing a scalable and convenient solution for deploying Tableau dashboards and reports.

- Tableau Mobile — Tableau Mobile is a free mobile app to access Tableau Server and Cloud. It allows users to explore their visualizations, either connected or offline.

Like Power BI, a typical usage scenario is to create dashboards with Tableau Desktop and then publish them to Tableau Server. Users can access the dashboards on their devices with Tableau Server, Tableau Cloud, or Tableau Mobile.

Tableau's key features that contribute to its popularity in the business intelligence and data visualization space [26]:

- Drag-and-Drop Visualizations — Tableau's intuitive drag-and-drop interface allows users to build visualizations without complex coding or scripting. Users can simply drag data fields onto the canvas, choose visual encodings, and interact with the visualizations to explore

---

[3]https://www.tableau.com/

different aspects of their data. This feature enables users, both technical and non-technical, to create compelling visualizations quickly.

- Data Connectivity — Tableau supports various data sources, including relational databases, semi-structured data, and raw data files.

- Recommended Visualizations — Tableau includes a "Show Me" feature that suggests appropriate visualizations based on the data selected for analysis. This feature helps users quickly identify suitable chart types, saving time and assisting in creating visually compelling and meaningful visualizations.

- Interactive Visualizations — Tableau enables the creation of highly interactive visualizations, enhancing the data exploration and analysis process. Users can apply filters, drill down into specific data points, create dynamic dashboards with linked views, and leverage various interaction techniques to gain deeper insights from their data.

Tableau, like any tool, has some limitations that users should be aware of:

- Limited Restoration Options — Tableau Server does not allow for the restoration of individual customer workbooks or data lost due to events other than system failure. This limitation can be problematic in cases where specific workbooks or data need to be recovered.

- Customization of Predictive Analytics — While Tableau provides predictive analytics capabilities, it does not allow users to customize the predictive analytics algorithms. This limitation may restrict users requiring more advanced customization and control over the predictive models.

- Dependence on IT Support — Tableau, especially for integrations and high-end tasks, often requires IT support. Users may need technical assistance or expertise for complex configurations, integrations with other systems, and advanced functionality. This dependence on IT support may limit users' ability to perform these tasks independently through tutorials or self-help resources.

According to Gartner, Tableau has a rating of 4.4 out of 5, indicating high user satisfaction. Approximately 81 % of users reviewing Tableau would recommend the tool to others [10]. In the "Likes" section, users highlight several positive aspects of Tableau. They appreciate the wide range of data sources that Tableau can connect to, allowing users to work with diverse datasets. The ability to work locally and online provides flexibility and accessibility for users. The drag-and-drop visualizations are praised for their intuitiveness, making it easy for users to create visualizations. However, some users have expressed concerns in the "Dislikes" section. One common concern is the cost of Tableau licenses. Some users feel that the price of licenses is high compared to the available features and alternative solutions on the market. Additionally, some users have reported encountering bugs or technical issues with Tableau. The lack of support in addressing these issues is also a drawback.

In conclusion, Tableau is a powerful and widely used business intelligence and data visualization tool that allows users to explore, analyze, and visualize data from various sources. With its intuitive drag-and-drop interface, Tableau makes it easy for users to create interactive visualizations. The tool offers a wide range of data connectivity options and provides recommended visualizations based on the data selected for analysis. Tableau's interactive visualizations enable users to drill down into specific data points, apply filters, and create dynamic dashboards for in-depth exploration. Nevertheless, it is essential to consider the limitations of Tableau, such as the need for IT support for high-end tasks and potential concerns about cost and technical issues reported by some users. Despite these considerations, Tableau has gained positive user ratings and recommendations, highlighting its value in helping organizations derive insights from their data and make data-driven decisions.

3.4 Qlik Sense

Qlik Sense[4] is a modern and intuitive business intelligence and data visualization tool developed by Qlik, expanded from QlikView in 2014. It allows users to explore and analyze data through interactive visualizations, making it easier to uncover insights and make data-driven decisions. Qlik Sense is known for its user-friendly interface and powerful associative engine, which enables users to dynamically navigate data and explore relationships across multiple data sources [25].

The main components of Qlik Sense include [20]:

- Qlik Sense Desktop — Qlik Sense Desktop is a Windows-based application that allows users to create, develop, and prototype visualizations and apps on their local machine. It provides a standalone environment for building visualizations without needing a server or Internet connection. Qlik Sense Desktop is often used for individual or offline analysis.

- Qlik Sense Cloud — Qlik Sense Cloud is the core web-based platform that enables users to create, explore, and share interactive visualizations and dashboards. It offers a drag-and-drop interface and a wide range of data visualization options, empowering users to uncover insights and make data-driven decisions. Qlik Sense provides powerful associative analytics capabilities, allowing users to explore data relationships.

- Qlik Sense Mobile — Qlik Sense Mobile is a mobile application that provides access to Qlik Sense apps and visualizations on smartphones and tablets. It allows users to explore and interact with their data on the go, providing a responsive and intuitive mobile experience.

Qlik Sense offers various features and capabilities to enhance data analysis and visualization. These features include [25]:

- Data Connectivity — Qlik Sense supports connectivity to various data sources, including databases, spreadsheets, cloud services, and web API.

---

[4]https://www.qlik.com/us/products/qlik-sense

- Responsiveness — Qlik Sense allows building visualizations for all platforms (browser, laptop, desktop, tablet, or smartphone).

- Drag and Drop Interface — Qlik Sense provides a drag-and-drop interface that makes it easy to build visualizations and create personalized dashboards and reports.

- Advanced Analytics Integration — Qlik Sense integrates with advanced analytics tools, allowing users to incorporate statistical analysis, predictive modeling, and machine learning algorithms into their data analysis workflows.

Despite being a popular tool, some limitations should be noted [25]:

- Cost — Qlik Sense offers a variety of features, but some may require additional purchases as add-ons. This can increase the overall cost of implementing and maintaining Qlik Sense.

- Comparison with QlikView[5] — Despite being a legacy tool, QlikView still has certain features that are not yet available in Qlik Sense. Organizations using QlikView may find that some functionalities they relied on are not fully replicated in Qlik Sense.

- Version Control — Qlik Sense does not provide built-in version control functionality. This means that managing different versions of apps or dashboards and tracking changes over time can be challenging.

According to Gartner, Qlik Sense has received a rating of 4.4 out of 5, with 81 % of users recommending the tool [9]. In the "Likes" section, users appreciate that Qlik Sense can accommodate various scenarios and handle diverse data types. The continued updates to the tool are seen as a positive aspect, ensuring users access the latest features and improvements. Users also value the support provided through the Qlik Community, which allows for knowledge sharing and problem-solving among the user community. However, there are some concerns highlighted in the "Dislikes" section. Users have mentioned that the collaboration and sharing features in Qlik Sense may not be as robust as desired. Some users feel it is more suitable for data analysis on a smaller scale due to potential performance issues when dealing with large datasets. The steep learning curve is also challenging, indicating that users may require significant time and effort to become experienced with the tool. Lastly, users have expressed dissatisfaction with the level of support received when encountering problems or difficulties.

In conclusion, Qlik Sense is a powerful business intelligence and data visualization tool that offers a range of features to help users analyze and explore their data. It has received positive ratings from users, with many highlighting its flexibility and ability to handle various data types. Qlik Sense's compatibility with different scenarios and its continuous updates are seen as strong points, along with the support provided by the Qlik Community. However, some users have mentioned limitations regarding collaboration and sharing features, scalability for large datasets, and the learning curve associated with the tool.

---

[5] https://www.qlik.com/us/products/qlikview

3.5 Tools comparison

Table 3.1 compares the described tools. The comparison fields were chosen based on their relevance to this work.

Table 3.1: BI tools comparison, adapted from [12]

Feature/Tool	Power BI	Tableau	Qlik Sense
Data Connectivity	Broad range of data sources supported	Broad range of data sources supported	Supports many popular data sources
Data Visualization	Interactive dashboards and reports	Advanced visualization capabilities	Interactive dashboards and reports
Analytics Features	Range of analytics features, including real-time insights, predictive analytics, collaboration, and data storytelling capabilities	Range of analytics features, including real-time insights, advanced calculations, collaboration, and data storytelling capabilities	Real-time insights, advanced calculations, collaboration, and data storytelling capabilities
Complexity	User-friendly, but can have a learning curve for more complex features	Steep learning curve, but highly customizable	User-friendly, but can have a learning curve for more complex features
Support for big data	Native support	Strong support	Supports
Cost	Low to moderate	High	Moderate to high

3.6 Summary

This chapter studied three popular business intelligence tools, Power BI, Tableau, and Qlik Sense.

Power BI, Tableau, and Qlik Sense are all powerful business intelligence and data visualization tools offering various features and capabilities. Each tool has strengths and weaknesses, making it suitable for different use cases and user preferences.

Power BI stands out with its seamless integration with the Microsoft ecosystem, user-friendly interface, and affordability. It offers a comprehensive suite of services enabling users to create, share, and access visualizations across various platforms.

Tableau is widely recognized for its advanced analytics capabilities, interactive visualizations, and extensive data connectivity options. While it may have a steeper learning curve and higher cost than other tools, Tableau provides great flexibility and customization options, making it a preferred choice for users with specific needs.

Qlik Sense offers a user-friendly interface and the ability to handle diverse data types. It is praised for its versatility and compatibility with various scenarios. However, users have noted limitations regarding collaboration and sharing features, scalability for large datasets, and the learning curve associated with the tool.

Gartner reviews indicate that all three tools have received positive ratings, with many users recommending them. However, users have also expressed concerns about each tool's cost, complexity, support, and specific features.

In conclusion, analyzing the reviews of the users, it becomes clear that complexity and handling large datasets are common problems.

Chapter 4

Problem Statement

This chapter provides an overview of how ZOS LDA conducts business performance evaluation and highlights the challenges associated with the current approach. Furthermore, the chapter presents the requirements specification, encompassing actors, user stories, and technical requirements, and concludes with the hypothesis.

4.1 ZOS LDA

In the current reporting process at ZOS, LDA, Excel, and PDF formats are predominantly used. Initially, there was an attempt to automate the Excel reports using the Google Analytics add-on. However, due to the lack of technical expertise among the responsible individuals, they encountered significant difficulties in configuring the automation. Furthermore, even after the configuration, the generated reports did not meet the required specifications, requiring additional manual calculations[1].

Currently, at the end of each month, an employee manually retrieves data from various sources such as Google Analytics, Facebook, Instagram, and Twitter. This data includes information about players, teams, competitions, and news. The employee then copies and pastes this data into an Excel spreadsheet. In the Excel spreadsheet, statistics calculations such as monthly sums are performed to analyze the data. Once the initial analysis is completed, the results are transferred to a PDF format for presentation within the company. This manual process takes approximately four days to complete. Some screenshots of the May report can be seen in Appendix A as an example of the current reporting.

In conclusion, the current reporting process at ZOS LDA, relying on manual data retrieval and analysis using Excel and PDF formats, is time-consuming, cumbersome, and prone to errors. To address these issues, developing a tool that automates data retrieval and streamlines the analysis

---

[1]Personal Communication

19

process would greatly benefit ZOS LDA. Such a tool would eliminate the need for manual data entry, reduce the potential for human errors, and significantly expedite the reporting process.

4.2 Problem Definition

The main goal of this work is to develop a user-friendly tool that simplifies the process of business performance evaluation and decision-making for managers. The tool entrusts users to define the relevant data for evaluation. It offers multiple options for data input, including manual entry or automatic retrieval from platforms like Facebook and Google Analytics through specially developed widgets.

The tool allows users to create custom business metrics or automatically generate metrics based on the data sources. It takes the data provided, performs data cleaning and preparation, and aggregates it to derive meaningful insights. The tool also stores the daily, weekly, and monthly metric values, enabling data display in various visualization formats. Users can access a dedicated dashboard that consolidates metrics providing a high-level overview of the project's performance. Users can also explore individual metric pages that provide comprehensive details, including historical data and interactive charts and graphs.

To ensure usability and accessibility, the tool must require minimal technical knowledge from the user. The interface should be intuitive, allowing users with varying technical expertise to navigate and utilize the tool effectively. Additionally, the tool should be adaptable and customizable to suit the needs of businesses across various industries and sectors.

4.3 Requirements Specification

This subsection presents the requirements specification that guided the development of the tool to overcome the problem effectively. It includes the actors, user stories, and technical requirements.

4.3.1 Actors

The system's actors are represented in Figure 4.1 and described in Table 4.1. Although the data sources do not interact with the system directly, they are involved in them. Therefore, they were considered an actor.

Table 4.1: System actors' description

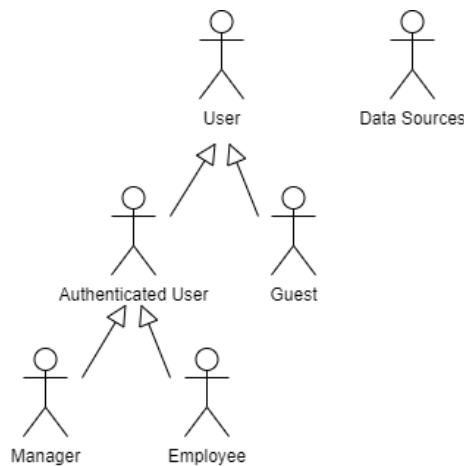| Identifier	Description	Example
Guest	The unauthenticated user who can sign up and sign in	n/a
Manager	The authenticated user who can create and manage projects	n/a
Employee	The authenticated user who can see the projects	n/a
Data Source	The data sources that can be used to retrieve data	Meta API

Figure 4.1: System actors in unified modeling language diagram

4.3.2 User Stories

In agile software development, user stories express user requirements in a semi-structured and natural language format. Each user story follows the form: "As a [who], I want to [requirement], so that [reason]." User stories capture the user's needs, desires, and expected outcomes [21].

The user stories are presented in Tables 4.2 to 4.4 to describe the solution's features. These user stories outline the specific requirements and functionalities from the perspectives of different user roles involved in the system.

Table 4.2: Guest's user stories

Identifier	Name	Priority	Description
US01	Sign in	High	As a Guest, I want to authenticate into the system, so that I can access and utilize the full functionality of the tool
US02	Sign up	High	As a Guest, I want to register an account, so that I can create and manage my projects and metrics within the tool

Table 4.3: Authenticated User's user stories

Identifier	Name	Priority	Description
US03	Logout	High	As an Authenticated User, I want to logout of the system, so that the device can be used by other users or to secure my session
US04	See the list of projects	High	As an Authenticated User, I want to view all my projects, so that I can easily access and manage the projects I am a part of

Table 4.3 – continued from previous page

Identifier	Name	Priority	Description
US05	See a project's dashboard	High	As an Authenticated User, I want to access the project's dashboard, so that I can quickly and conveniently view a high-level overview of the metrics
US06	See a project's metrics	High	As an Authenticated User, I want to view the list of metrics for a project, so that I can see all the available metrics and their details
US07	See a metric's page	High	As an Authenticated User, I want to view the details of a specific metric, so that I can understand its historical data, and visualization
US08	See data in multiple formats	High	As an Authenticated User, I want to visualize my data in various formats (charts, graphs, tables), so that I can gain insights and make data-driven decisions
US09	See the project's alerts	Medium	As an Authenticated User, I want to have a centralized dashboard to view all project alerts, so that I can quickly assess the overall project health
US10	Export data	Medium	As an Authenticated User, I want to export my data or reports to common formats (CSV, PDF, Excel), so that I can further analyze or share them outside of the tool
US11	Notifications	Medium	As an Authenticated User, I want to receive notifications when certain metrics or thresholds are met, so that I can proactively address potential issues or opportunities
US12	Edit profile	Medium	As an Authenticated User, I want to edit my profile information, so that I can keep my profile up-to-date with accurate information
US13	Delete account	Medium	As an Authenticated User, I want to delete my account, so that I can permanently remove my personal information and data from the system

Table 4.4: Manager's user stories

Identifier	Name	Priority	Description
US14	Create Project	High	As a Manager, I want to create a new project, so that I can organize and manage my data and metrics within a dedicated workspace

Table 4.4 – continued from previous page

Identifier	Name	Priority	Description
US15	Customize the dashboard	High	As a Manager, I want to define and customize the dashboard layout and visualization options for each project, so that I can tailor the view to the specific needs of the project
US16	Customize the metric's page	High	As a Manager, I want to customize the layout and configuration of a metric's page, so that I can tailor the presentation and analysis of the metric to meet specific requirements or preferences
US17	Add manual source	High	As a Manager, I want to add a manual source to a project, so that I can incorporate additional data that is not available through automated integrations or existing data sources
US18	Add automatic source	High	As a Manager, I want to add an automatic source to a project, so that I can integrate external data sources and automate the data retrieval process
US19	Remove a source	High	As a Manager, I want to remove a source from a project, so that I can manage and maintain the relevance and accuracy of data within the tool
US20	See the sources	High	As a Manager, I want to see the sources associated with a project, so that I can review and analyze the data providers used for the project's metrics
US21	Delete project	High	As a Manager, I want to delete a project, so that I can remove unnecessary projects from the system
US22	Assign employees	Medium	As a Manager, I want to assign team members to specific projects, so that they can collaborate and contribute to the project's objectives
US23	Remove employee	Medium	As a Manager, I want to remove an employee from a project, so that I can manage the project team and their access to project-specific data and functionalities
US24	Set up notifications	Medium	As a Manager, I want to set up notifications for specific metrics or data conditions, so that I can be alerted of any significant changes or anomalies
US25	Add metric	Medium	As a Manager, I want to add metrics to a project, so that I can track and analyze specific data relevant to the project

Table 4.4 – continued from previous page

Identifier	Name	Priority	Description
US26	Remove metric	Medium	As a Manager, I want to remove metrics from a project, so that I can maintain a relevant and focused set of metrics for analysis
US27	Edit the project	Small	As a Manager, I want to edit project details, so that I can update project information and settings as needed

4.3.3 Technical Requirements

The technical requirements for the system are described in Table 4.5.

Table 4.5: Technical requirements

Identifier	Name	Description
TR01	Usability	The system should be simple and easy to use
TR02	Security	The system should have appropriate security measures, including encryption for sensitive data and secure authentication for access
TR03	Scalability	The system should scale up as the amount of data grows and the number of users increases
TR04	Integration	The system should be able to integrate with other tools and systems, such as API and other data sources, as specified by the user

4.4 Hypothesis

This dissertation was accomplished with the following hypothesis as a guide:

> "It is feasible to make a tool with a **small learning curve** that every organization can use. This tool is fed automatically with data from multiple sources, and the analysis it provides will be helpful when making **business decisions**."

The developed tool aims to have a minimal learning curve, ensuring that users with limited technical knowledge can efficiently utilize it. While users are expected to understand metrics and relevant data for assessing business performance, they are not required to have data modeling or analysis expertise. By creating a user-friendly interface and simplifying complex technical processes, the tool aims to prevent data neglect and encourage its utilization in making informed business decisions.

Chapter 5

Design and Implementation

The purpose of this chapter is to detail the process involved in the implementation of the prototype. It includes the tools and technologies used, the data modeling, the design of the user interface, and the implementation discriminated by the frontend and backend. It is relevant to note that, because of time constrictions, not all the features defined in Chapter 4 were implemented. Instead, only the user stories of high priority were implemented, and some of medium priority. Furthermore, the only actors involved in the system are the *User*, as either *Authenticated User* or *Guest*, and *Data Source*. Therefore, the user stories associated with the *Manager* were seen as user stories of the *Authenticated User*.

5.1 Tools and Technologies

This subsection presents the choices for technologies and tools for both the frontend and backend and why they were chosen.

5.1.1 Frontend

The choice of technologies for implementing the frontend was made based on their suitability for developing user interfaces in a web platform. Among the multiple *JavaScript* frameworks available, *ReactJS*[1] was selected due to its suitability for building web applications, its user-friendly learning curve enabled by extensive documentation, active online question-and-answer communities like *StackOverflow*, and prior experience with the framework. *ReactJS* is a *JavaScript* library designed for creating user interfaces on the web. This framework offers high application performance described by speed and robustness, thanks to its implementation of the Virtual DOM (Document Object Model). Unlike most frameworks that directly interact with the browser DOM, causing an entire page refresh with every change, *ReactJS* utilizes the Virtual DOM to selectively

---

[1]https://react.dev/

re-render only the affected elements in response to state modifications, minimizing unnecessary page reloading [1].

In addition to *ReactJS*, the *React Router DOM*[2] library was used to handle routing and navigation within the application. This library allows the definition of navigation paths and the rendering of the appropriate components based on the Uniform Resource Locator (URL).

Furthermore, the *Axios*[3] library was utilized for handling requests to the backend. *Axios* is a widely-used *JavaScript* library for making Hypertext Transfer Protocol (HTTP) requests from the browser.

Additionally, the *Bootstrap*[4] framework was incorporated alongside *ReactJS*. *Bootstrap* is widely recognized as one of the most popular frameworks for developing responsive applications.

Moreover, several *NPM* (Node Package Manager)[5] packages were also used. *NPM* serves as a comprehensive software registry containing numerous *JavaScript* packages. These packages can be any resource developers have found helpful in addressing specific challenges while being easily integrable with other projects.

5.1.2 Backend

For implementing the backend, *NodeJS*[6] was selected since it is typically used with *ReactJS*. Besides that, *NodeJS* is known for its easy learning curve and popularity. Previous experience with this technology was also a factor in its selection.

Complementing *NodeJS*, *Axios*[7] was also used. The *Axios* library was selected to handle the HTTP requests the backend might need. It facilitates data transmission between the backend and external API or clients.

In addition to *Axios*, the *Express*[8] framework was also utilized. *Express* simplifies the development of server-side applications and API in *NodeJS*. It provides robust features for handling routing, middleware, and HTTP request/response management.

In order to be able to use the database of choice described in the following subsection, a library that allowed communication between it and the backend was necessary. The chosen one was *Mongoose*[9] due to its popularity and community support. This library enables the backend to send requests and execute many database operations.

5.1.3 Database

A primary objective of this work is to empower the analysis of historical data. As a result, the tool must be highly scalable, as the data involved could be enormous in volume. Additionally, the

---

[2] https://reactrouter.com/en/main
[3] https://axios-http.com/
[4] https://getbootstrap.com/
[5] https://www.npmjs.com/
[6] https://nodejs.org/en
[7] https://axios-http.com/
[8] https://expressjs.com/
[9] https://mongoosejs.com/

data may originate from multiple sources and have an unknown structure, necessitating flexibility in accommodating various data schemas. Considering these requirements, a *NoSQL* database was selected. *NoSQL* databases are well-suited for scalability and flexibility due to their ability to handle large data volumes and schema-less nature [17].

In exploring various popular *NoSQL* databases, research was conducted using their documentation and online tutorials to determine the most suitable choice. *MongoDB* was selected due to its high performance and widespread adoption. *MongoDB* is an open-source document database that offers several key features, such as fault tolerance, persistence, aggregation capabilities, *ad hoc* queries, and indexing. In *MongoDB*, documents are primarily stored in BSON (Binary JSON) format, akin to JSON or XML, but represented in binary form. This binary representation enhances storage efficiency and scan speed. A noteworthy advantage of *MongoDB* for the intended tool is its support for polymorphic documents. This means that fields or attributes can differ from one document to another, even within the same collection. As a result, *MongoDB* does not impose strict schema requirements on the documents, providing the desired flexibility. However, it is essential to acknowledge some disadvantages of *MongoDB*. Firstly, it lacks the ACID (Atomicity, Consistency, Isolation, Durability) properties typically associated with relational databases, which can introduce some unreliability in specific scenarios. *MongoDB's* indexing mechanism also requires a substantial amount of Random Access Memory (RAM), which should be considered when designing and provisioning the system [14].

Despite these drawbacks, *MongoDB* was considered the most suitable choice for implementing the tool, given its high performance, popularity, flexibility in handling diverse data structures, and ability to accommodate the tool's varying and potentially large-scale historical data requirements.

5.2   Architecture

The architecture of the implemented solution is portrayed in Figure 5.1. The solution comprises two main components: the frontend and the backend. The frontend focuses on the client-side logic of the application, including the development of the dashboard, its components, user interaction, and data presentation. On the other hand, the backend is responsible for providing data to the frontend, which can be sourced from the database *MongoDB* or the *zerozero* API.

The communication between the frontend and backend, as well as between the backend and the *zerozero* API, is achieved through HTTP requests. In the case of the backend and *MongoDB*, communication is facilitated using the *Mongoose* library. A more detailed illustration of the components and their relationships can be seen in Figure 5.2.

5.3   Data Modeling

The conceptual model of the implemented solution, depicted in the UML Class Diagram of Figure 5.3, illustrates the entities and their relationships.
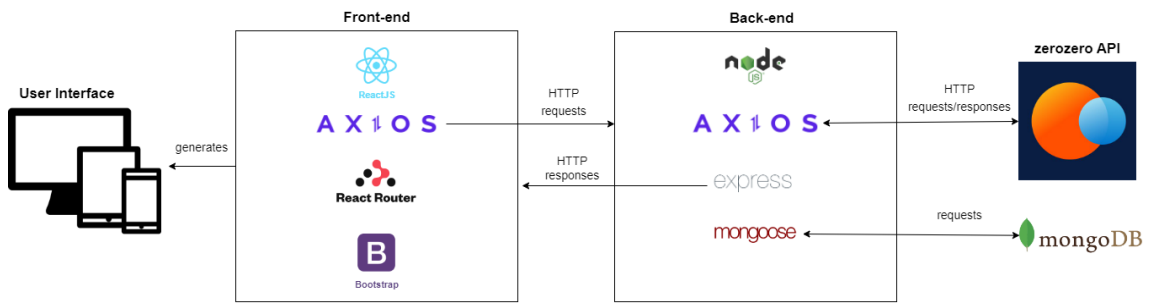
Figure 5.1: Architecture

The *user* entity stores the `username`, `email` (which must be unique), and `password`. Each user can have many projects, while each project is associated with only one user.

A *project* is identified by its `name` and may include an optional `description`. Projects are connected to multiple sources, where each source has a `name`, a `type` (either SpreadSheet or API), a `startRetrievalDate` (indicating when data retrieval from that source should begin), and a `lastRetrieval` (representing the date of the most recent data retrieval from that source). If the source type is API, a URL field is mandatory.

A *source* consists of multiple sourceValues, each containing a `date` and a map with `data` from that source for that date. Additionally, a source is composed of metrics.

A *metric* has an `identifier` (a string retrieved during data collection), a `name` (a user-friendly representation of the metric), an optional `formula`, a `category` (defaulted to 'Others'), an optional `unit`, and an optional `description`. The `betterWhenHigher` attribute is a boolean (defaulted to true) that indicates whether higher values are better for the metric. The `onHome` field is a boolean expressing whether the metric should appear on the homepage. The `dashboard` attribute is an array of strings representing the types of charts that will be displayed on the dashboard to represent that metric. Similarly, the `chart` attribute controls the charts on the metric page, and having at least one element is mandatory. A metric is composed of *metricValues*, which capture specific values for a metric.
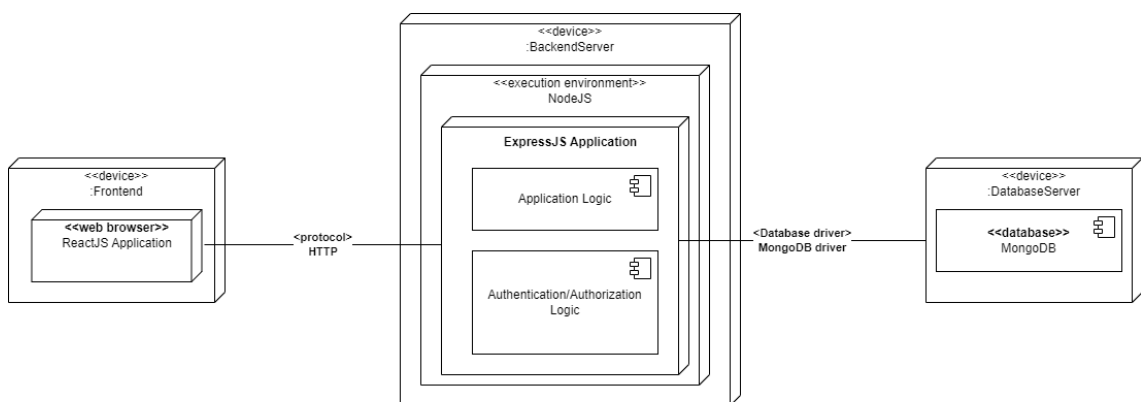

Figure 5.2: UML deployment diagram

A *metricValue* includes the `day`, `month`, `year`, and `week` attributes, which collectively represent the date (separated for easier querying). Other attributes include the `value` itself, a `synced` field (a boolean flag indicating whether it is synchronized with the system), and the `date`.

The *dailyValues* entity is similar to *metricValue*, except it does not have the `date` attribute. Instead, it features a `ranking` field and a `previousDayComparison` field. The `ranking` field represents the ranking of the value for that specific date, while the `previousDayComparison` field indicates the percentage comparison of the value with the previous day.

The *weekValue* and *monthValue* entities are similar to *dailyValues* but represent the week and month, respectively.

During the implementation planning phase, how the data was to be stored was a considerable concern, as the tool needed to be capable of handling substantial amounts of data effectively. The implemented idea was based on the necessity of knowing where the data came from and how to make access easier. The tool operates by retrieving data from a source, and it was deemed crucial to store this data in both *sourceValues* and *metricValues*. This ensures that the origin of each specific value can be traced back accurately using the *sourceValues*. The collections that will experience the most significant growth are *metricValues* and *dailyValues*. The rationale behind having both *metricValues* and *dailyValues* is as follows: *metricValues* only include essential fields, while *dailyValues* contain additional fields such as `ranking` and previousDayComparison, and potentially more. By keeping the *metricValues* documents smaller and focused on the essential fields, they become more easily accessible. Consequently, when new data is retrieved, the immediate action is to update the *metricValues* collection. The *dailyValues*, *weekValues*, and *monthValues* collections, on the other hand, are calculated based on the data stored in the *metricValues* collection. By organizing the data storage in this manner, the tool ensures efficient handling of vast quantities of data. The *metricValues* collection maintains the essential data, while the *dailyValues*, *weekValues*, and *monthValues* collections are derived from the *metricValues* to provide additional analysis and comparisons.

5.4 Front-end

This subsection will describe the process of implementing the frontend of the application, including the design and the implemented features.

5.4.1 User Interface Design

In terms of design, the ZOS LDA company made the decision to avoid the complexity and time investment associated with creating an application design from scratch. Recognizing the availability of existing design solutions, the company purchased a *Bootstrap* template. After thorough research conducted by the design team, the "Good —- Bootstrap 5 Admin Dashboard Template"[10] from Bootstrap Themes was chosen. This template was deemed compatible with the company's

---

[10]https://themes.getbootstrap.com/product/good-bootstrap-5-admin-dashboard-template/
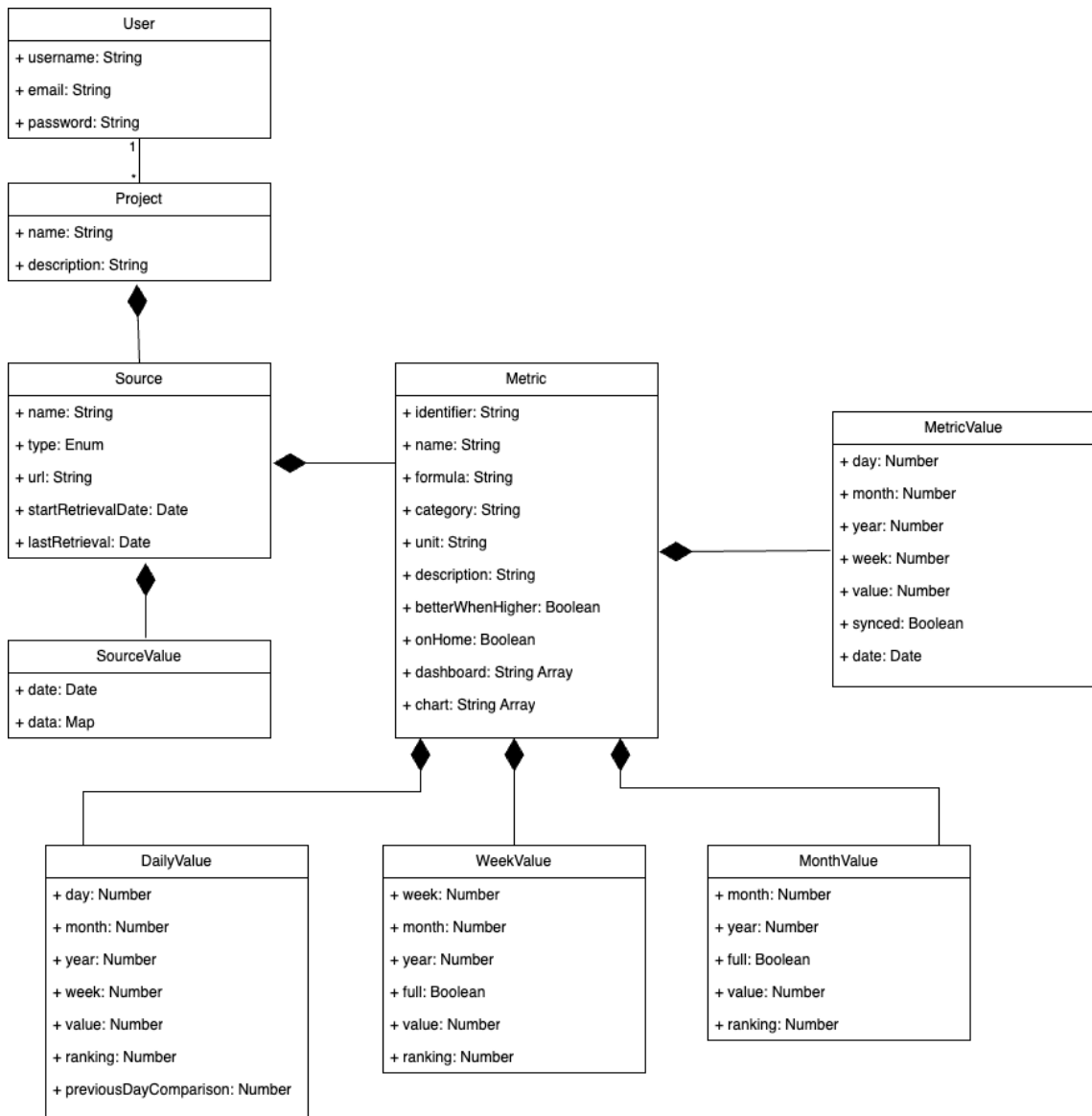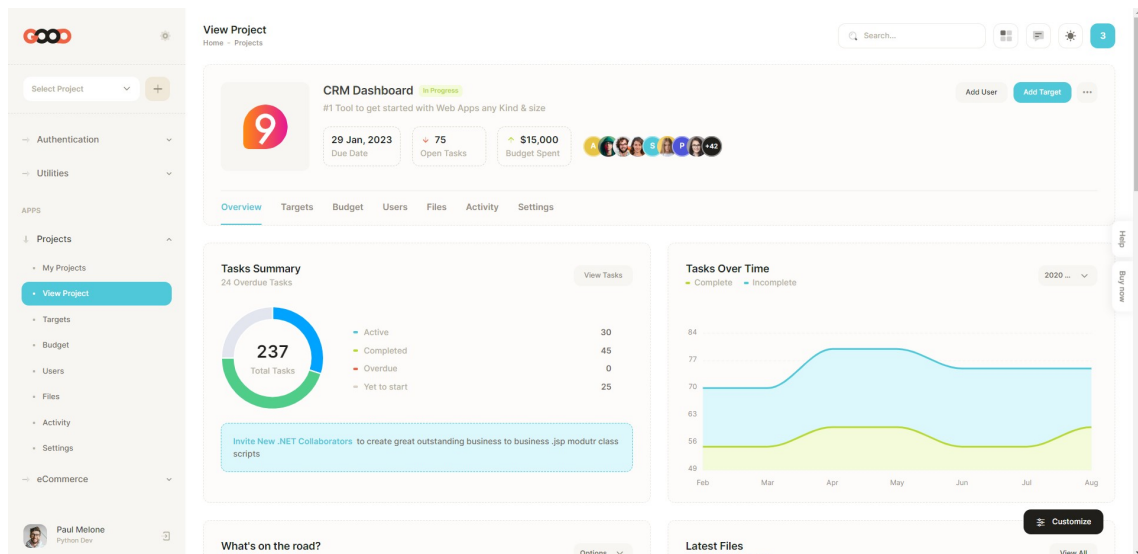
Figure 5.3: Domain model

Figure 5.4: Good template - view project page

chosen technologies, recent and comprehensive, with all the necessary components expected to be needed. Bootstrap Themes provides a range of themes developed explicitly for developers, with each theme serving as an extension of *Bootstrap*, tailored for specific scenarios and equipped with specific components and plugins. These themes also have comprehensive documentation to facilitate their utilization [3]. Figure 5.4 presents the "View Project" from the template.

Upon exploring the purchased Bootstrap theme, several challenges were encountered due to the lack of documentation regarding its adaptation with *ReactJS* or any other framework. Although some online tutorials were available, initially, the process of adapting the theme proved to be complicated. One of the main challenges was the lack of modularity and component-based structure in the theme's code. For instance, common elements like the "navbar" were not implemented as separate components but instead repeated on all pages. As a result, the files became large and challenging to read. Another issue encountered was related to styling. Instead of having separate styling files for each page, only one existed, containing approximately 61 000 lines of code.

Given the objective of developing a modular and component-based codebase, utilizing the template's code as the foundation for the implementation process was not feasible. Moreover, since the frontend was developed using *ReactJS*, the usage of *jQuery* in the theme was unsuitable and needed to be replaced. Therefore, an alternative approach was adopted. During the implementation process, careful consideration was given to the theme, identifying which pages or components could be utilized. Only the necessary code from the theme was selected and copied to achieve the same desired outcome. However, this approach proved to be challenging, particularly when it came to replicating the design aspects. To ensure progress in the implementation process without excessive time investment, the decision was made to copy the styling file entirely from the theme to the prototype's code.

The company's design theme also provided some branding for the prototype, including the *logo* and *favicon*.

5.4.2 User Interface

An unauthenticated user only has access to two pages regarding the authentication. Figure B.1 represents the default page when no user is authenticated and where a user can sign in. This page is accessed through the routes / or /login. Figure B.2 represents the sign-up page that is accessed through the route /register.

If an unauthenticated user tries to access other pages besides these, it is redirected to the sign-in page.

When a user is authenticated, they gain access to various pages related to their projects. These pages share a similar structure and include several common components:

- Sidebar — The sidebar provides the user convenient access to all the primary pages within the application.

- Header — The header appears at the top of each page and contains the page's title or section the user is viewing.

- Breadcrumbs — Breadcrumbs are a navigational aid that displays the user's current location within the application's hierarchy.

By incorporating these common components across the project-related pages, the application maintains consistency and provides users with a familiar interface, enhancing usability.

The sidebar not only provides access to the primary pages but also includes a button to create a new project. When the button next to "Select Project" is clicked, it opens the project creation modal. The project creation process consists of "Project Settings" and "Upload Data".

1. Project Settings (Figure B.3) — In this step, the user is required to choose a name for the project and has the option to provide a description.

2. Upload Data (Figure B.4) — In this step, the user can add sources to the project. Two types of sources are available: Spreadsheets and API.

    - Spreadsheet — The user needs to upload a file with the extension *.xlsx* containing the data for the project. It is required that the file has a column named date that contains the corresponding dates for the data entries. This column is crucial for correct processing and displaying the data within the tool.

    - API — The user must provide the URL for the API, which should accept specific parameters for data retrieval. The required parameters are:

        – DateSince — A date in the format YYYY-MM-DD that defines the start date of retrieval.

        – PerPage — The number of elements returned on each page.
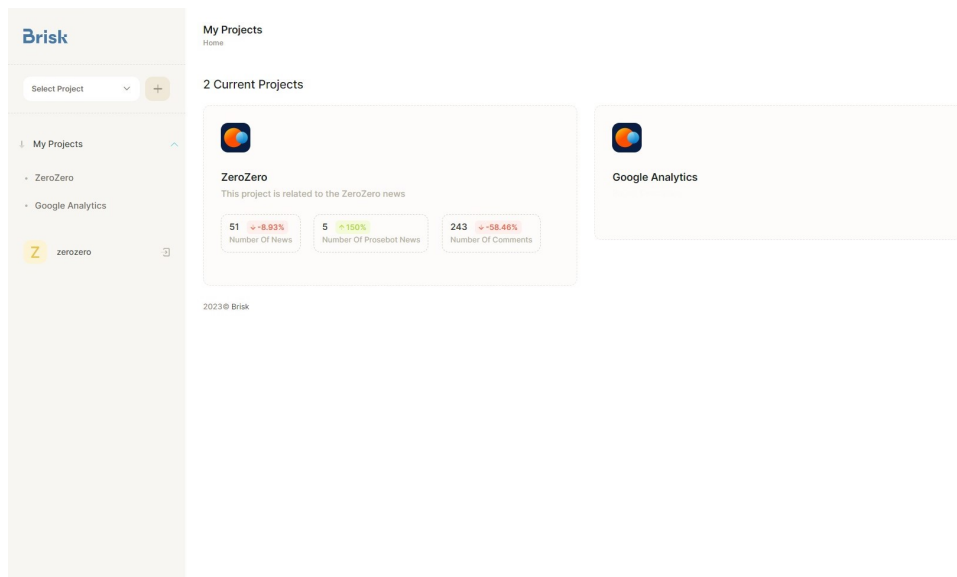
        – Page — The current page.

Figure 5.5: My projects page

In addition to the URL, the user must also provide the start date for data retrieval, which will replace the `DateSince` parameter in the URL. The date picker component from the *rsuite* library, specifically the *DatePicker*[11], was used to implement the date selection feature.

If the project creation process is successful, the user is shown the last tab of the modal called "Completed", which confirms the successful creation of the project.

The authenticated user's homepage, the "My Project" page, is the landing page after sign-in. This page, as depicted in Figure 5.5, provides users with an overview of their projects. Each project is represented by a card that displays the project's name, description (if available), and up to three selected metrics. The metrics shown include the metric name, value from the latest day (excluding today), and the percentage comparison to the previous day. The user can choose which metrics are displayed on the project cards. This page gives users an overview of their projects and allows them to monitor the selected metrics. To access the "My Project" page, users can navigate the root route (`/`) or the `/homepage` route.

Within a project, users have access to its content through five tabs, each offering a unique perspective on the project's information and settings. These tabs are the dashboard, metrics, alerts, sources, and settings.

In the dashboard tab, represented in Figure 5.6, a user is presented with an overview of the project's key metrics and visualizations, which the user can customize. The dashboard includes two fixed elements at the top: "Key Statistics" and "Highlights".

---

[11] https://rsuitejs.com/components/date-picker/

- Key Statistics — This card lists the project's key metrics and displays information about the latest day, including the date, metric value, and a comparison with the previous day. It offers users a quick snapshot of the project's essential metrics and their recent performance.

- Highlights — The Highlights card presents a table format view of the key metrics. It showcases the best and worst recorded values for each metric and their respective dates. This information allows users to identify significant milestones or notable changes in the project's metrics.

Following these fixed elements, individual charts are displayed for each key metric selected by the user. The type of chart used, whether a line or a bar chart, can be customized based on user preferences. The chart's card includes the chart itself and options for aggregation filtering and selecting the date range. To implement the charts, the *Recharts*[12] package was utilized. This package was chosen due to its user-friendly learning curve and flexible capabilities. While *Recharts* offers various chart types, the line and bar charts were deemed most suitable for representing data over time in the prototype. By default, the dashboard tab's charts display the daily data for the current month. However, users can select different periods, such as weeks or months. When the week option is selected, the data is grouped by week, with the x-axis displaying the corresponding week numbers. Similarly, when the month option is chosen, the data is grouped by month. To facilitate the selection of a custom date range, the *DateRangePicker* component from the *rsuite*[13] library was employed.

Another fixed element at the bottom is the data table card provided to present the detailed data for the key metrics. The card with the data table features the *react-data-table-component*[14] package, which provides a highly flexible and well-documented solution for implementing data tables. The data table allows users to explore the project's metrics in a tabular format and includes a filter for selecting the aggregation. The selected filter option determines the first column of the data table. The first column represents the date if the user chooses the "Day" option. If "Week" is selected, the first column represents the weeks; for the "Month" option, the first column represents the months. The remaining columns of the data table correspond to each key metric. Each cell in the table displays the value for the selected aggregation and ranking. If the selected option is "Day", the cells also include the percentage comparison of the value with the previous day. Additionally, the cells for the top five values are distinguished with a medal icon, indicating their significance. Users can sort the columns in the data table, rearranging the data based on their preferences and easily analyzing the metrics according to different criteria.

The dashboard tab provides users with a comprehensive overview of their project's key metrics and insights by combining the "Key Statistics", "Highlights", customizable charts, and a data table. This tab can be accessed through the route `/project/{projectId}`.

---

[12]`https://recharts.org/en-US`
[13]`https://rsuitejs.com/components/date-range-picker/`
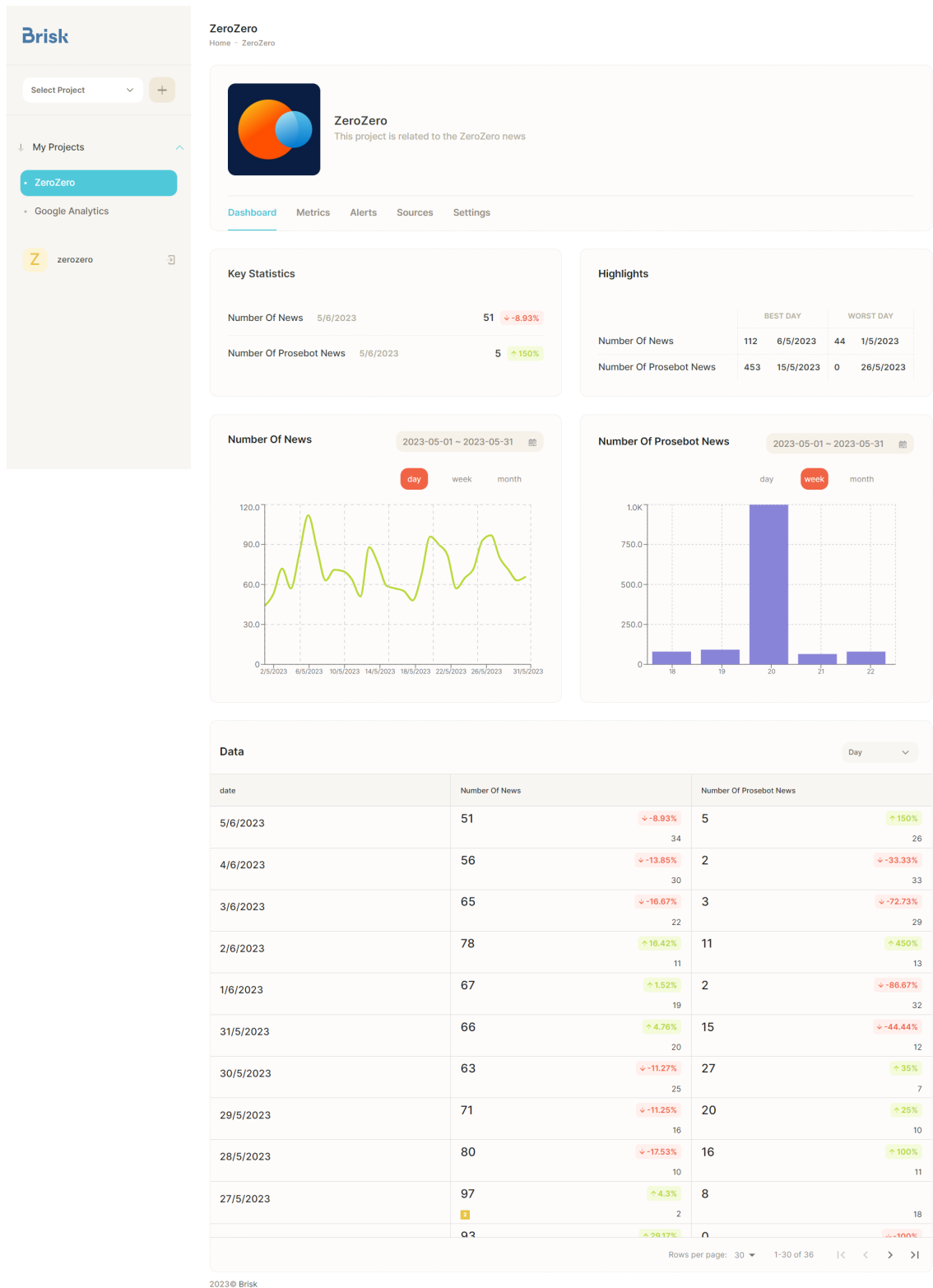[14]`https://react-data-table-component.netlify.app/?path=/story/getting-started-intro--page`

Figure 5.6: Project's dashboard page

In the metrics tab, represented in Figure B.7, users can access all the metrics associated with a project. These metrics are organized by their respective categories. Each metric is presented as a card containing the metric's name and information about the latest day, such as the value and the percentage comparison with the previous day. The metrics tab provides users with a centralized location to explore and analyze the project's metrics in a categorized manner. This page is accessed through the route `/project/{projectId}/metrics`.

By clicking on a metric card, users can access the dedicated metric page, as shown in Figure 5.7. This page provides users with detailed information about the selected metric. It consists of three main components: the charts card, the podium, and the data table. The charts card functions similarly to the charts in the dashboard tab, allowing users to visualize the metric's data over time. The default date range on this page is set to the previous week, providing a more focused view of recent trends and fluctuations. The podium card displays a carousel of the top three days, top three weeks, and top three months in terms of metric value. This allows users to quickly identify the periods with the highest metric values, providing valuable insights into the metric's performance. Additionally, users can explore the metric's data in the data table, which operates similarly to the data table in the dashboard tab. It presents the metric's values, rankings, and comparisons, allowing users to analyze and compare the metric's data across different time periods. To access this page, users can utilize the route `/project/{projectId}/metrics/{metricId}`, providing them with a comprehensive view of the specific metric's information within the project.

The settings icon next to the metric's name allows users to access the settings modal, where they can edit various aspects of the metric. This modal is represented in Figure 5.8. Within the settings modal, users can modify the following information related to the metric:

- Name — The name used to display the metric (mandatory).

- Category — Grouping category for the metric. Users can choose an existing category or create a new one (mandatory).

- Unit — If applicable, the unit of measurement for the metric (e.g., seconds).

- Description — A description of the metric that appears as a tooltip when hovering over the metric's name in the prototype.

- Best Value — Indicates whether the highest or lowest value is considered the best for the metric.

- Show in Metric Page — Users can select which charts should be displayed for the metric on the metric's dedicated page (mandatory to choose at least one).

- Show in Dashboard — Similar to the previous option, users can choose whether to display the metric in the dashboard tab. If none are selected, the metric will not appear in the dashboard.
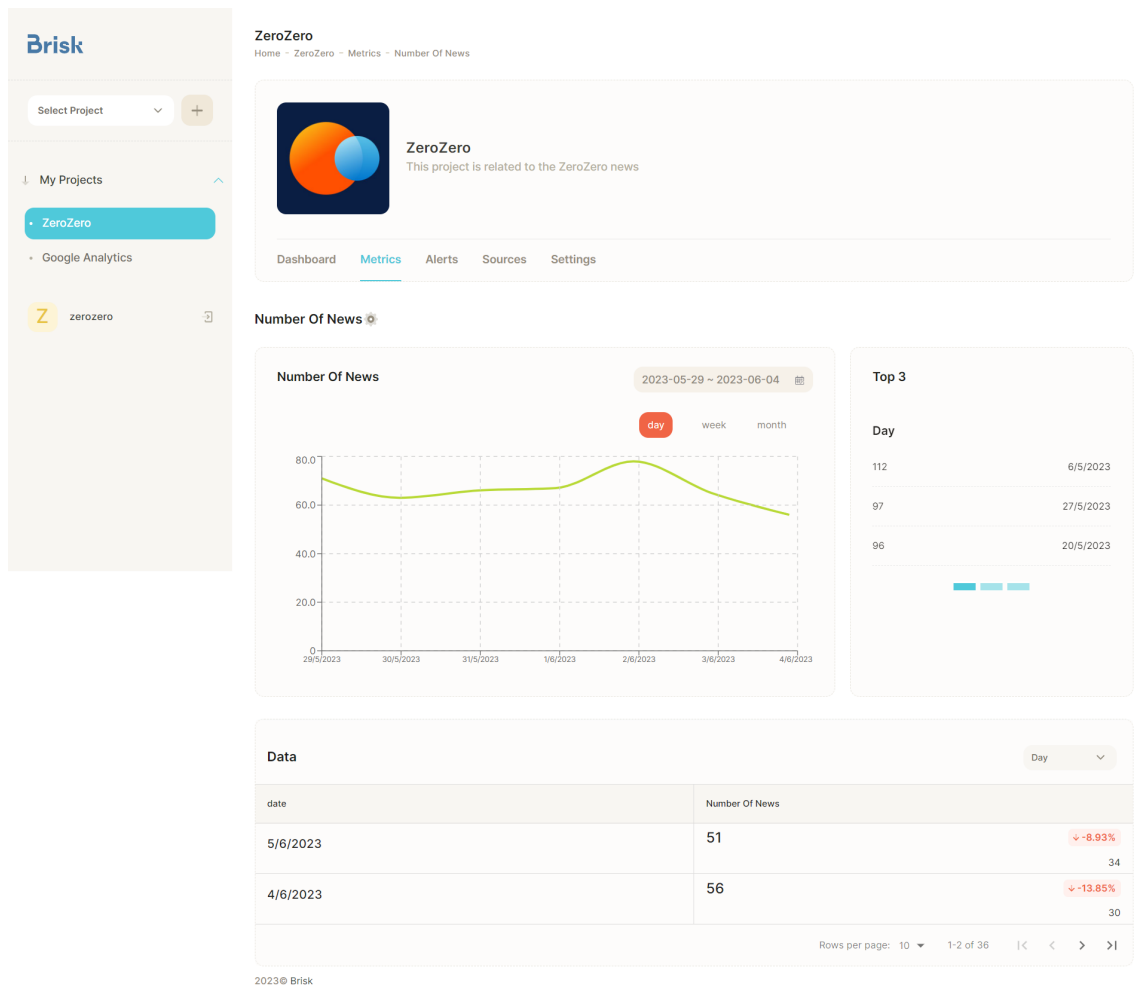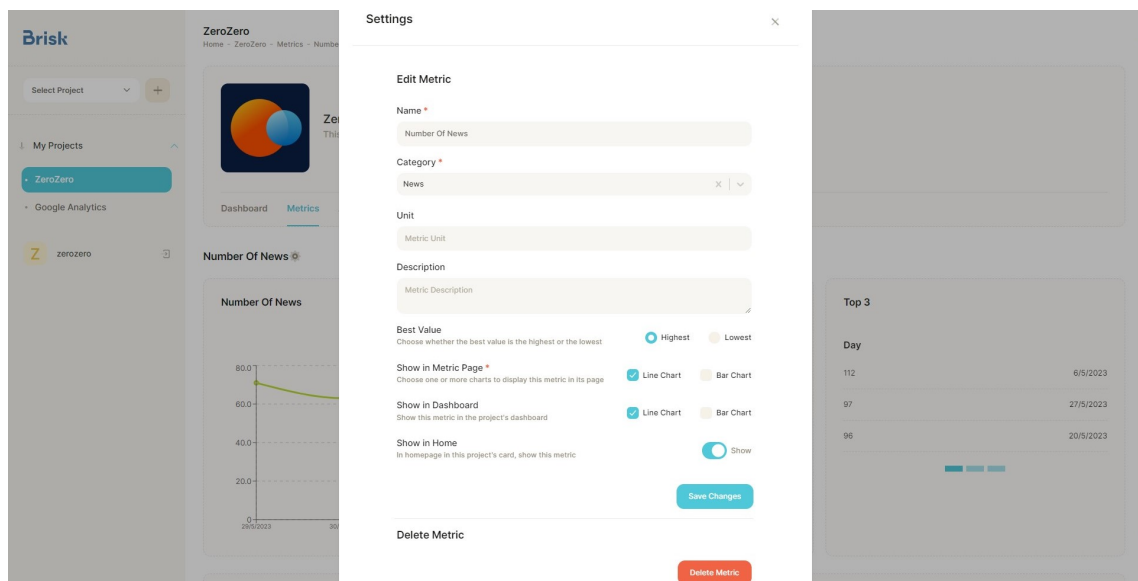
Figure 5.7: Project's metric page

Figure 5.8: Project's metric settings modal

- Show in Home — Users can decide whether the metric should be shown in the project card on the homepage.

Besides editing the metric, the user can also delete the metric on this modal. These settings allow users to customize and manage the metrics according to their specific requirements and preferences.

In the alerts tab, represented in Figure B.11, users can view information about the best day, week, and month for each metric. This page provides valuable insights into the metric's performance and allows users to identify when it achieved its highest value. The alerts tab is accessible through the route `/project/{projectId}/alerts`.

In the sources tab, represented in Figure B.10, users can view the sources associated with the project. Each source is represented by a card displaying an icon to identify the source type (SpreadSheet or API), the source name, and the insertion date. Users can open a settings modal to delete the source by clicking the settings icon within the card. Additionally, the page provides an "Add Source" button, which opens a modal similar to the "Upload Data" section in the project creation modal. This allows users to add a new source to the project. This page is accessed through the route `/project/{projectId}/sources`.

The project's settings tab, shown in Figure B.12, allows users to manage the project's settings. This page provides the option to delete the project. Users can access this page by navigating the route `/project/{projectId}/settings`.

5.5 Backend

This subsection details the backend implementation. It describes the authentication and authorization implementation, the API specification, and how the sources' management routines were

handled.

5.5.1 Authentication and Authorization

MongoDB, as a database system, does not provide built-in authentication functions. However, authentication and authorization can still be implemented effectively when using *MongoDB* with the chosen technologies for the prototype. One common approach is to use JWT[15] (JSON Web Tokens) and cookies to establish a session-based authentication system. JWT are compact, URL-safe tokens that securely carry information between parties. They consist of a header, a payload, and a signature. The header contains metadata, the payload carries the user's claims or attributes, and the signature ensures the token's integrity.

In the authentication process, when a user provides correct credentials and attempts to log in, a random JWT token is generated using the *jwt.sign(payload, secretOrPrivateKey, [options, callback])* function. This function takes three parameters: the payload, the secret or private key, and optional parameters. The payload typically contains information about the user, such as their unique identifier (user ID), roles, or any other relevant data. In this case, the user's ID is used as the payload. The secret or private key is a secret value used to sign the JWT and ensure its integrity. This key is known only to the server and is kept secure. Additionally, options can be provided to customize the token's behavior. One common option is `expiresIn`, which sets an expiration time for the token. In this case, the value was set to 1 d, indicating that the token will expire after 1 day. By setting an expiration time, the token becomes valid only for a specific duration, enhancing security and reducing the risk of unauthorized access if the token is compromised. Upon receiving a successful response from the backend authentication request, the frontend sets a cookie with the JWT token. This cookie is then made accessible on all pages of the application. By storing the token in a cookie, it can be sent automatically in subsequent requests, allowing the server to authenticate and authorize the user.

For the authorization part in the backend, a middleware was implemented in the backend to verify the validity of the JWT token. This middleware is designed to intercept incoming requests and check whether a valid token is present. When a request is made to a protected route or an endpoint that requires user authentication, the middleware retrieves the token from the request headers. It then verifies the token's authenticity and expiration. If the token is valid and has not expired, the middleware allows the request to proceed, and the desired action or endpoint is executed. On the other hand, if the token is invalid or has expired, the middleware rejects the request.

A *React* context was created to persist the authentication state across different routes and components in the frontend. This context stores the current user's state and makes it accessible to various components within the application. When setting the current user, a request is made to retrieve the user's information. Like in the middleware, the token is extracted from the request and verified to ensure validity and integrity. The server returns the user's information if the token is

---

[15]https://www.npmjs.com/package/jsonwebtoken

valid, meaning it has not expired and its signature matches the expected secret key. The frontend then sets the current user in the context, making it available for other components to access and utilize.

5.5.2 Application Programming Interface

The endpoints, excluding authentication and data handling, are organized into three express routers, each serving a specific purpose. The specifications for each route are presented in Tables 5.1 to 5.3. The table focuses on the successful response status for each endpoint while acknowledging that failure to meet the request criteria may result in a 404 Not Found or 500 Server Error response.

The source router, /source, handles operations related to managing sources, such as retrieving all sources and deleting a source. The specification of these endpoints can be seen in Table 5.1.

Table 5.1: Source's endpoints

Endpoint	Method	Parameters	Body	Description	Response
/deleteSource	POST	-	sourceId	Delete the source by id and all the data related to that source	204 No Content: Source successfully deleted
/:projectId	GET	projectId	-	Retrieve all the project's sources	200 OK: Returns an array of sources

The project router, /project, is responsible for handling operations related to project management, including creating, deleting, and retrieving projects. The specification of these endpoints can be seen in Table 5.2.

Table 5.2: Project's endpoints

Endpoint	Method	Parameters	Body	Description	Response
/createProject	POST	-	userId, name, description	Create a project	201 Created: Returns the created project
/deleteProject	POST	-	projectId	Delete the project by id and all the data related to that project	204 No Content: Project successfully deleted

Table 5.2 – continued from previous page

Endpoint	Method	Parameters	Body	Description	Response
/allProjects/:userId/:home	GET	userId, home	-	If home, retrieve all the projects with the metrics information for the latest day, if not just the projects	200 OK: Returns an array of projects
/:id	GET	id	-	Retrieve a project	200 OK: Return the object project

The metric router, /metric, handles operations related to metric management, including updating, deleting, and retrieving metrics. The specification of these endpoints can be seen in Table 5.3.

Table 5.3: Metric's endpoints

Endpoint	Method	Parameters	Body	Description	Response
/updateMetric	POST	-	id, name, category, onHome, dashboard, chart, description, unit, betterWhenHigher	Update the metric	200 OK: Returns the updated metric
/allMetrics/:projectId	GET	projectId	-	Retrieves the project's metrics	200 OK: Returns an array of the metrics, grouped by category

Table 5.3 – continued from previous page

Endpoint	Method	Parameters	Body	Description	Response
/dashboardMetrics/:projectId	GET	projectId	-	Retrieve the metrics for the project's dashboard	200 OK: Returns an array of the metrics with their daily, weekly, and monthly values
/alerts/projectId	GET	projectId	-	Retrieve the project's alerts	200 OK: Return an array of alerts
/:id	GET	id	-	Retrieve a metric	200 OK: Return the object metric with its daily, weekly, and monthly values

5.5.3 Sources and Routines

Data handling is a crucial aspect of the prototype, encompassing three primary scenarios involving data creation and updates. These scenarios include creating a new project and two routines designed to ensure data accuracy and completeness: the database integrity verification routine and the daily data fetching routine.

5.5.3.1 Project Creation

Users provide essential details during the project creation phase, including the project name and an optional project description. Additionally, users can upload data as a spreadsheet or specify API endpoints. Upon project creation, the prototype proceeds to handle the uploaded data. When users upload spreadsheets, the frontend sends the corresponding files, whereas, for API endpoints, an array of endpoints is sent. Although the treatment of each data type differs during retrieval from external sources, the subsequent processing remains essentially the same. To ensure proper organization and management of the data, a new source is created for each uploaded file or specified endpoint.

When retrieving data from API endpoints, the *Axios* library makes the necessary requests to the designated endpoints. For data sourced from spreadsheets, the *read-excel-file*[16] library is

---

[16]https://www.npmjs.com/package/read-excel-file

employed to extract the data from the uploaded spreadsheets. This library enables efficient parsing of spreadsheet files and provides the necessary functions to access and process the data.

Subsequently, the processing of the uploaded data commences. Verification is performed for each column or attribute, excluding the `date` attribute, to determine if the corresponding metric already exists. The existence of a metric is determined by matching the `projectId`, `sourceId`, and `identifier`. If a metric with identical parameters already exists, no additional metric is added. However, a new *metric* is created if no matching metric is found.

Following the metric creation, a new *metricValue* is generated for each day, corresponding to the uploaded data. This step ensures the system retains a comprehensive record of metric values over time. Furthermore, a new *sourceValue* is created to associate the metric values with their respective source.

In case of any issues or errors during this data processing phase, the user is notified of the encountered problem. This ensures that users are informed about any potential discrepancies or anomalies in the uploaded data, enabling them to take appropriate actions or make necessary adjustments.

It is essential to highlight that with the project creation process, the user cannot yet see the data in the prototype. This decision is driven by the potentially large volume of data users may upload during project creation, which could lead to significant waiting times. To optimize user experience, only essential project information is immediately stored, while the processing required for creating project tabs is deferred until the "Synchronize Routine" described in Subsection 5.5.3.3. By deferring the creation of project tabs, the system mitigates potential performance issues associated with data-intensive operations during project creation. This approach ensures a smoother user experience and allows for more efficient handling of large datasets.

5.5.3.2 Retrieve Daily Data

Automatic sources (API) within projects require daily data retrieval and updating to ensure data currency and accuracy. This task is accomplished through a *Cronjob* implemented with the *cron*[17] library, scheduled to run at 4 am, allowing sufficient time for any necessary data updates or processing before daily activities start. It iterates through all projects with automatic sources, triggering a request to retrieve the latest data.

It is worth noting that the data retrieval process is designed to fetch data from the last date it was retrieved, considering any potential issues that may have occurred during previous retrieval attempts. This approach guarantees that the most recent and complete dataset is obtained.

The subsequent steps in processing the retrieved data remain consistent with the previously described procedure outlined in Subsection 5.5.3.1.

By executing the routine and performing the necessary data retrieval and processing, the prototype ensures that users are presented with the most up-to-date and accurate data at the beginning of

---

[17]https://www.npmjs.com/package/cron

each business day. This enables users to make informed decisions based on the latest information available, contributing to the effectiveness and relevance of their decision-making processes.

5.5.3.3   Synchronize Routine

Once the data is retrieved from the sources during project creation or the daily fetch process, it must be processed to be effectively presented in the prototype. To accomplish this, a routine has been implemented using the *cron* library. This routine is scheduled to run every 10 minutes and is responsible for synchronizing the data, ensuring its proper representation in the project's tabs. During synchronization, various data aggregations generate *dailyValues*, *weekValues*, and *monthValues*, which are displayed in the project's tabs.

As described in Section 5.3, each *metricValue* possesses a `synced` field—a boolean flag indicating whether the *metricValue* is synchronized with the system. This flag plays a vital role in the synchronization routine. The *metricValue* is considered synced if there is a corresponding *dailyValue* with an updated value for the same date. Additionally, if the *metricValue* falls on a Sunday, it should have a corresponding *weekValue* with the value and ranking calculated for the respective week; for any other day, there should be a *weekValue* with no value and ranking. Similarly, if the *metricValue* represents the last day of a month, a *monthValue* should exist with the calculated value and ranking for that month; for any other day, there should be a *monthValue* with no value and ranking.

The synchronization routine commences by iterating through the metrics of each project. Within this loop, all *metricValues* that are not yet synced are fetched and sorted by value in descending order. Subsequently, the routine performs the following steps for each *metricValue* regarding the *dailyValues*:

1. The routine attempts to locate the *dailyValue* with an immediately higher value than the *metricValue's* value.

2. It searches for the *dailyValue* from the day before the *metricValue*.

3. A comparison in percentage is calculated between the values of the *dailyValue* from the day before and the *metricValue*.

4. The routine verifies whether a *dailyValue* already exists for the date and metric of the *metricValue*.

   - If no *dailyValue* exists, a new *dailyValue* is created using the data from the *metricValue*, including the previously calculated comparison. If available, the ranking of the closest *dailyValue* is used, and the ranking is set to the value increased by 1; if not, the ranking is set to 1.

   - If a *dailyValue* already exists, the routine checks if its value matches the *metricValue's* value.

– If the values differ, the *dailyValue*, including the value itself, is updated. The ranking and comparison are updated as if it were a newly created *dailyValue*.

– If the values are the same, no further action is taken.

5. Suppose there has been a change in the *dailyValue* (either through creation or update), and the closest *dailyValue* exists. In that case, the routine locates and updates all other *daily-Values* with a ranking greater than the closest *dailyValue*. However, the *dailyValue* created or updated is exempt from this update. The ranking property of each affected *dailyValue* is incremented by 1.

For each *metricValue*, the routine follows the following steps to handle the *weekValues*:

1. The routine verifies whether a *weekValue* already exists for the *metricValue's* week and metric.

   • If no *weekValue* exists, it verifies if the *metricValue's* date corresponds to the end of the week.

      – If it is the end of the week

         (a) It retrieves the *dailyValues* of that week and sums its values.

         (b) It retrieves the *weekValue* with an immediately higher value.

         (c) It creates a new *weekValue* using the data from the *metricValue*. If available, the ranking of the closest *weekValue* is used, and the ranking is set to the value increased by 1; if not, the ranking is set to 1.

      – If it is not the end of the week, a new *weekValue* is created with the data from the *metricValue*, both value and ranking set to 0.

   • If a *weekValue* exists, it verifies if the *metricValue* corresponds to the end of the week.

      – If it does, then the same happens as when no *weekValue* exists, and it corresponds to the end of the week. The only difference is that instead of being created a new document, it is updated.

      – If it does not, then nothing happens.

2. Suppose there has been a change in the *weekValue* (either through creation or update). In that case, the routine locates and updates all other *weekValues* with a ranking greater than the closest *weekValue*. However, the *weekValue* created or updated is exempt from this update. The ranking property of each affected *weekValue* is incremented by 1.

In the routine, the handling of the *monthValues* is the same as the handling of the *weekValues*. The only difference is that instead of verifying if the *metricValue* corresponds to the end of the week, it verifies if it corresponds to the end of the month.

Lastly, if all these operations are successful, the flag `synced` of the *metricValue* is set to true. This way, if something went wrong before since the `synced` flag remains false, the next time the routine runs, the *metricValue* will be analyzed again.

5.6   Summary

This chapter provides an in-depth description of the implementation process of the prototype. It begins by discussing the technologies employed, differentiating between the frontend technologies (*ReactJS*, *Axios*, *React Router*, and *Bootstrap*), backend technologies (*NodeJS*, *Express*, *Axios*, and *Mongoose*), and the database technology used (*MongoDB*). The prototype's architecture illustrates the various components and their communication mechanisms. Additionally, the data modeling aspect is elaborated upon, elucidating the entities and relationships involved in the prototype.

Subsequently, the chapter delves into the implementation details, starting with the rationale behind choosing a Bootstrap template for the design. The frontend implementation is then described, encompassing the different pages implemented and their corresponding routes. Moving to the backend, the authentication process utilizing JWT tokens is explained, highlighting the implementation of middleware and protected routes to enhance authorization mechanisms. The chapter also provides a comprehensive overview of the developed API, elaborating on its functionality and endpoints. Lastly, the chapter focuses on three fundamental functions related to data management within the prototype, offering a more detailed explanation. It delves into the data retrieval, storage, and maintenance process, elucidating the steps involved in each of these aspects.

In essence, this chapter serves as a comprehensive guide to the implementation process of the prototype, covering various aspects ranging from technology choices and architectural considerations to data management functionalities.

Chapter 6

Tests and Validation

This chapter presents the experiments conducted to evaluate the prototype's usability and validate the hypothesis while assessing the achievement of predefined goals. The chapter includes a detailed description of the experiments, the procedure followed, the data collected, and the evaluation criteria employed. Additionally, it presents the form used to conduct the experiment, providing insights into the structure and sections of the form. Finally, the chapter acknowledges the experiment's limitations, highlighting any constraints or factors that may have influenced the study's implementation and the resulting data.

6.1 Overview

The experiment aimed to evaluate the prototype's usability and validate or refute the hypothesis while assessing the achievement of predefined goals. These goals encompassed examining the tool's learning curve and its suitability for users with varying levels of technical knowledge.

The participants consisted of ZOS employees directly involved in reporting processes, representing the target users of the tool. The selection process encompassed individuals with diverse technical expertise, including those well-versed in technology and those with limited or no technical background.

To facilitate the experiment, gathering relevant data was crucial. Collaborative discussions with representatives from ZOS LDA led to the decision that all participants should easily understand the presented data. Consequently, an endpoint at the following URL was developed: `https://www.zerozero.pt/api/v1/exportZZData/AppKey/AgvbeMz0TPBso2Jy`. This endpoint enabled retrieving essential data, such as `NumberOfNews`, `NumberOfProsebotNews`, `NumberOfVisits`, and `NumberOfComments`. The data encompassed the entire month of May, ensuring uniform conditions for all participants.

With the acquired data in hand, the subsequent step involved formulating pertinent questions for the test. These questions sought to measure participants' comprehension of the tool's features

and were designed with objective, correct answers in mind. Throughout the experiment, participants were free to explore and utilize the tool's functionalities to find the correct responses to each question.

6.2   Procedure

The experiment was conducted at the ZOS LDA installations, where two monitors were set up: one for presenting the forms and another for displaying the prototype.

The following procedure was followed during the experiment:

1. Explanation of Experiment — Participants were introduced to the purpose of the experiment, emphasizing the importance of their participation and ensuring their rights as participants. The overall procedure was explained, including the timing of answering questions related to the tool.

2. Demographic Information — Participants were asked to provide demographic information, such as age and gender.

3. Prototype-Related Questions — Participants were presented with questions about the prototype, specifically focusing on information visualization. Participants had to provide their chosen answers and the reasoning behind their selection for each question.

4. Prototype Customization Tasks — Participants were given specific tasks related to customization on the prototype. Participants were asked to describe how they accomplished the task, comment on the ease of customization, and note if changes were immediately visible on the page.

5. System Usability Scale (SUS) Form — Participants were requested to complete the SUS form.

6. Submission of the Form — Once participants completed the tasks and filled out the SUS form, they were instructed to submit the form, signaling the end of their participation in the experiment.

6.3   Collected Data and Evaluation Criteria

The experiment began by gathering demographic data from the participants, including age, gender, prior experience with business intelligence tools (if applicable), and knowledge in the domains of business intelligence and information visualization. Additionally, two specific questions were included to gauge participants' perspectives on the importance of data-driven decision-making within a business and their perception of the usefulness of a tool that enables the collection and visualization of relevant business data.

Afterward, participants were asked questions about the prototype, and their responses were collected. The evaluation of these responses focused on assessing the accuracy of the answers and the validity of the logical reasoning provided. Errors in the questions were analyzed quantitatively and objectively, while errors in logical reasoning were evaluated qualitatively and subjectively. This evaluation aimed to determine participants' comprehension of the tool's features and the accuracy of the information presented.

Furthermore, participants' performance in customizing the prototype tasks was assessed. This evaluation considered the validity of participants' task completion, including whether the outcomes aligned with the expected results. Errors in the tasks were measured objectively and quantitatively based on whether the task was performed correctly or not. Participants' explanations of how they completed the tasks were also considered to evaluate the ease of accomplishing them. This assessment aimed to gauge the tool's ease of use and intuitiveness.

Finally, participants were asked to provide feedback using the SUS form. The responses from the SUS form were subjective and qualitative, allowing participants to express their perceptions and opinions regarding the overall usability of the tool.

6.4 Test

The form used in the experiment is structured into four distinct sections to facilitate data collection and ensure a systematic approach. The complete form can be seen in appendix C.

6.4.1 Consent and Demographic Data

The initial part of the form aims to obtain participants' informed consent for participating in the experiment. The informed consent is divided into four parts:

1. Purpose of the Study — This section introduces the study's purpose and context.

2. Study Procedure — In this section, the form explains the procedure described in Section 6.2.

3. Duration — This part provides participants with an estimated duration of the experiment. It helps them understand the time commitment required for their involvement, allowing them to plan accordingly.

4. Voluntary Participation — The form explicitly states that participation in the study is voluntary. It emphasizes that participants have the right to withdraw from the study at any time without facing any consequences. Furthermore, it assures participants that their data will be deleted if they discontinue participation.

Regarding the demographic data, participants must provide their ages and gender. Participants are also asked about their experience with business intelligence tools, if applicable. This information helps understand participants' familiarity with similar tools and their potential influence on

the experiment's results. Furthermore, participants' business intelligence and information visualization knowledge are assessed using a 5-point Likert scale. This allows to measure participants' self-perceived knowledge and expertise in these areas. By examining the correlation between participants' prior knowledge and experience with the experiment's outcomes, potential relationships or influences on participants' performance and understanding of the prototype might arise.

In addition to the demographic and knowledge-related questions, two questions are included to assess participants' opinions on the core principles of this work. The first question inquires whether participants believe business decisions should be based on data, highlighting the importance of data-driven decision-making. The second question asks participants whether they perceive a tool like the prototype being developed as helpful for businesses. These questions provide insights into participants' attitudes and beliefs regarding the significance of data-driven decision-making and the potential value of the proposed tool in real-world business contexts.

6.4.2 Questionnaire

The questionnaire section of the study focuses on assessing participants' experience with the prototype. Participants are free to use the prototype as they see fit to find the answers to the questions posed. The main objective of these questions is to evaluate the effectiveness of the information display and the overall intuitiveness of the tool. The questionnaire includes a series of questions requiring participants to explore various prototype features. Some questions are straightforward and may only require a quick interaction with the tool, while others may necessitate a more in-depth exploration to uncover the desired information. This approach allows for gathering insights into participants' navigation patterns, the ease with which they find the requested information, and their overall understanding of the prototype's functionality.

The following questions comprise the questionnaire, where excluding the first question, they are related to the project "ZeroZero" (since this was created with the data described in Section 6.1):

1. How many projects can be seen? — Test the homepage's ability to express how many projects exist.

2. How many metrics can be seen in the project's dashboard? — Test the project's dashboard ability to demonstrate what are the metrics present.

3. How many metrics does the project have in total? — Test that the project's tabs have intuitive names and that the metric's categories do not confuse the counting.

4. What was the best day and value for the metric 'Number of News'? (The best day is understood as the day when the metric reached its highest value.) — Test the tool can demonstrate the highest value and its date

5. What was the value for that metric on May 9th, 2023? — Test the tool has the ability to search for the value for a specific date.

6. For that day and the same metric, what was the percentage of growth and ranking? (The percentage of growth can be either positive or negative) — Test the tool has the ability to search for data for a specific date.

7. What was the weekly value, for the same metric, for May 15th, 2023? — Test the tool has the ability to search for data for a specific aggregation.

To ensure the comprehensibility of the questions, a pilot test was conducted prior to the main study. During the pilot test, it became evident that there was a need to provide specific instructions regarding the date format and the start day of the week. Therefore, before presenting the questions to participants, the following information was provided:

- Date Format — Please consider that the dates provided in the questions are in the format DD-MM-YYYY.

- Start of Week — Please note that the week is considered to start on Monday and finish on Sunday.

6.4.3 Tasks

The tasks section of the study focuses on assessing the customization. The tasks are designed to evaluate the ease of use, flexibility, and effectiveness of the customization functionality in the prototype. Participants are given two specific tasks related to the "ZeroZero" project. While the tasks are accomplished in the same space, they serve different purposes.

The tasks are as follows:

1. Change the chart type of the metric 'Number of Comments' on the metric's own page to a line chart. — Test if participants can intuitively identify where to edit the metric.

2. Change the 'Number of Comments' metric so that it appears on the dashboard as a bar chart representation. — Test participants' understanding of the different settings available within the metric settings modal

6.4.4 System Usability Scale

The final section of the form is dedicated to administering the SUS (System Usability Scale) form. The SUS form is a standardized tool for assessing the participants' subjective evaluation of the prototype's usability. It consists of a set of statements that participants are asked to rate on a Likert scale, expressing their level of agreement or disagreement [4].

6.5 Limitations

Despite the careful design and execution of the experiment, several limitations should be acknowledged. These limitations stem from certain constraints that influenced the study's implementation and the resulting data. They are outlined below:

- Prototype Deployment — The prototype was not deployed in a live production environment due to logistical constraints. Consequently, the testing had to be conducted locally, limiting the number of participants and potentially generalising the findings to a larger user population.

- Single Pilot Test — While a pilot test was conducted to refine the form questions, only one pilot test was conducted due to time constraints and the limited number of participants. This limited the opportunity to gather extensive feedback and make further refinements, potentially resulting in unforeseen issues with the questions during the main testing phase.

6.6 Summary

In this chapter, experiments were conducted to evaluate the prototype's usability and validate the hypothesis. The experiments involved participants from ZOS, who represented the target users of the tool. The participants had varying levels of technical knowledge and were requested to answer questions and perform customization tasks related to the prototype. The collected data included demographic information, questionnaire responses, and SUS form feedback. The results indicated participants' comprehension of the tool's features, ability to find information and ease of customization. The study's limitations were also acknowledged, such as the limited number of participants and the lack of live deployment. Overall, the experiments provided valuable insights into the prototype's usability and its potential for supporting data-driven decision-making in a business context.

It is relevant to note that performance tests were not conducted because of the lack of available sources. Since the only implemented sources were SpreadSheets and API, where the API did not provide a large quantity of data at a time, testing the performance would be very limited, therefore, deemed not applicable.

Chapter 7

Results

This chapter presents the experiment's results described in Chapter 6. The chapter begins with presenting the key findings with a detailed account of the participants' performance on the tasks and their responses to the questionnaire. Lastly, the chapter offers an in-depth interpretation and analysis of the findings, aiming to uncover insights, explore potential explanations, and address the research objectives.

7.1   Analysis

The experiment involved nine participants, comprising three women and six men. The participants' ages ranged from 18 to 44 years, with a majority falling within the 18 to 24 age group. The distribution of participants' ages is illustrated in Figure 7.1. It is worth noting that all nine participants responded positively to the questions regarding data-driven decision-making in business and the perceived utility of a tool like the one proposed in this work. This indicates undivided agreement among the participants with the fundamental concepts presented in this study and with the study made of the state of the art, Chapter 3. Furthermore, the data analysis revealed that neither age nor gender significantly affected the participants' responses.

For each question in the questionnaire described in Subsection 6.4.2, it was calculated the percentage of correct answers and the percentage of valid logical reasoning. This evaluation is presented in Table 7.1. According to the data, the evaluation of the questionnaire responses reveals generally high percentages of correct answers and valid logical reasoning across most questions. However, it is notable that questions 4 and 6 obtained relatively lower values for correct answers.

This discrepancy can be attributed to a mismatch between participants' understanding of the questions and the predetermined criteria for evaluating the answers. Both questions required the inclusion of two specific elements: a date and a corresponding value for the best day and the percentage of growth and ranking, respectively. To be considered correct, the answers needed
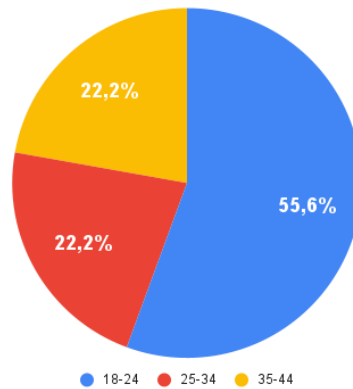
Figure 7.1: Ages distribution

to encompass both components. However, many participants only responded to one part of the question, even if their answer was accurate, leading to an overall incorrect response.

For question 4, only two participants provided entirely incorrect answers, while three correctly answered half of the question, and the remaining four answered both components correctly. This indicates that while only 44.4% of participants answered the question entirely correctly, 77.8% responded accurately to at least half of the question. This discrepancy between the percentages for correct answers and valid reasoning can be attributed to the partial correctness of some participants' responses.

Similarly, for question 6, all participants provided at least half of the correct answer, but only six participants answered the question in its entirety. This discrepancy explains why the percentage of valid logical reasoning is 100% while the percentage of correct answers is not.

Table 7.1: Percentage of correct answers and valid logical reasoning for each question

Question	Correct Answers	Valid Reasoning
1	100%	100%
2	88.9%	88.9%
3	88.9%	77.8%
4	44.4%	77.8%
5	100%	100%
6	66.7%	100%
7	100%	100%

Furthermore, an analysis of the logical reasoning provided by participants for questions 4, 5, and 7 indicates a discernible trend in their preference for using either the charts or the data table to formulate their responses. The distribution of participants' choices can be observed in Figure 7.2. Notably, the six participants who answered question 4 using the chart also chose the chart for question 5. This suggests a consistent approach in utilizing visual representations for

these specific questions. Conversely, out of the three participants who responded to questions 4 and 5 using the data table, two continued to employ the data table for answering question 7.
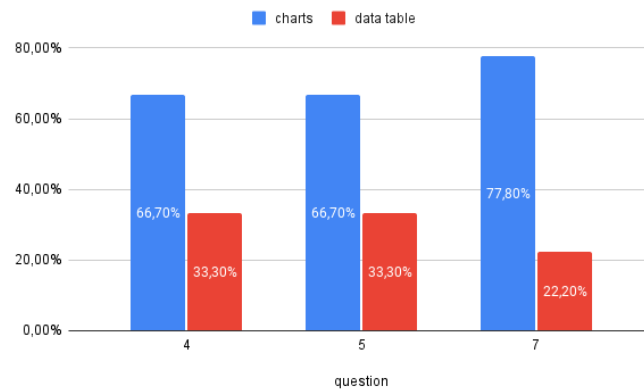
Figure 7.2: Comparison of chart and data table usage for questions 4, 5, and 7

With the timings for each question recorded in the format *minutes:seconds:milliseconds*, a comprehensive analysis was conducted to examine the distribution and variability of completion times. The statistical analysis encompassed the calculation of average, median, minimum, maximum, first quartile, third quartile, and standard deviation for the completion timings of each question. These key summary statistics are presented in Table 7.2. Upon observing the table, notable patterns emerge regarding the completion times for each question. It becomes evident that the first question required the least time to answer, while the final question posed the longest challenge to participants. Furthermore, questions 2, 3, and 5 display comparable average completion times, suggesting similar complexity or effort in addressing these particular items. Conversely, substantial variations in individual response times are evident when examining the disparity between the minimum and maximum completion times for each question.

Table 7.2: Average, median, minimum, maximum, 1st quartile, 3rd quartile, and standard deviation of the timings for each question

Question	Avg.	Med.	Min.	Max.	1st Q.	3rd Q.	Std. Dev.
1	0:08:41	0:05:30	0:02:01	0:20:00	0:04:29	0:14:28	0:06:37
2	0:21:44	0:17:00	0:06:39	0:53:26	0:10:26	0:30:30	0:15:09
3	0:17:01	0:10:08	0:04:00	0:53:19	0:08:27	0:15:49	0:16:13
4	0:41:55	0:35:47	0:17:22	1:27:51	0:24:56	0:39:21	0:26:12
5	0:17:59	0:15:56	0:05:04	0:40:34	0:12:21	0:21:32	0:10:54
6	0:34:59	0:42:15	0:07:03	0:53:30	0:24:12	0:49:11	0:17:04
7	0:58:16	0:57:00	0:20:58	1:33:25	0:30:07	1:20:38	0:27:54

To investigate the potential influence of participants' previous experience with a BI tool and their knowledge in the BI and information visualization domains on the questionnaire results,

correlations were calculated between these variables. Specifically, the following correlations were examined:

1. Correlation between previous experience with a BI tool and the percentage of correct answers and valid reasoning in the questionnaire.

2. Correlation between knowledge in the BI domain and the percentage of correct answers and valid reasoning.

3. Correlation between knowledge in the information visualization domain and the percentage of correct answers and valid reasoning.

The Spearman rank correlation coefficient was employed to determine these correlations, as it is suitable for assessing the relationships between ordinal variables. The Spearman rank correlation coefficient ranges from -1 to 1. A value close to -1 indicates a strong negative correlation, a value close to 1 indicates a strong positive correlation and a value close to 0 indicates no correlation. The correlation matrix is presented in Figure 7.3.

Analyzing the first correlation, there was no significant correlation between previous experience and the percentage of correct answers (r = 0). However, a weak positive correlation was found between previous experience and the percentage of valid reasoning (r = 0.12). Regarding the second correlation, the results revealed a moderate negative correlation between knowledge in the BI domain and the percentage of correct answers (r = -0.35), indicating that a higher level of knowledge in the BI domain was associated with a lower percentage of correct answers. Similarly, knowledge in the BI domain exhibited a moderate negative correlation with the percentage of valid reasoning (r = -0.49), suggesting that higher knowledge in the BI domain was linked to lower levels of valid reasoning in the questionnaire. Lastly, for the third correlation, the information visualization domain demonstrated a weak negative correlation with the percentage of correct answers (r = -0.26), indicating that higher knowledge in the information visualization domain was associated with a slightly lower percentage of correct answers. Moreover, a strong negative correlation was observed between knowledge in the information visualization domain and the percentage of valid reasoning (r = -0.60), indicating that participants with higher knowledge exhibited lower levels of valid reasoning.

Concerning the tasks described in Subsection 6.4.3, it is notable that all participants completed them without any errors. The explanations provided by the participants on how they accomplished each task were consistent, indicating a clear understanding of the steps required to achieve the desired outcomes. However, a slight variation was observed among participants regarding their ease in locating the metric's settings for editing.

Regarding the timing of the tasks, a statistical analysis was conducted to examine the differences in the time taken to complete each task. This analysis is presented in Table 7.3. As anticipated, notable variations were observed between the two tasks, which is expected considering they are carried out within the same area of the prototype. Specifically, the average time required to complete the first task was significantly higher than the average time for the second
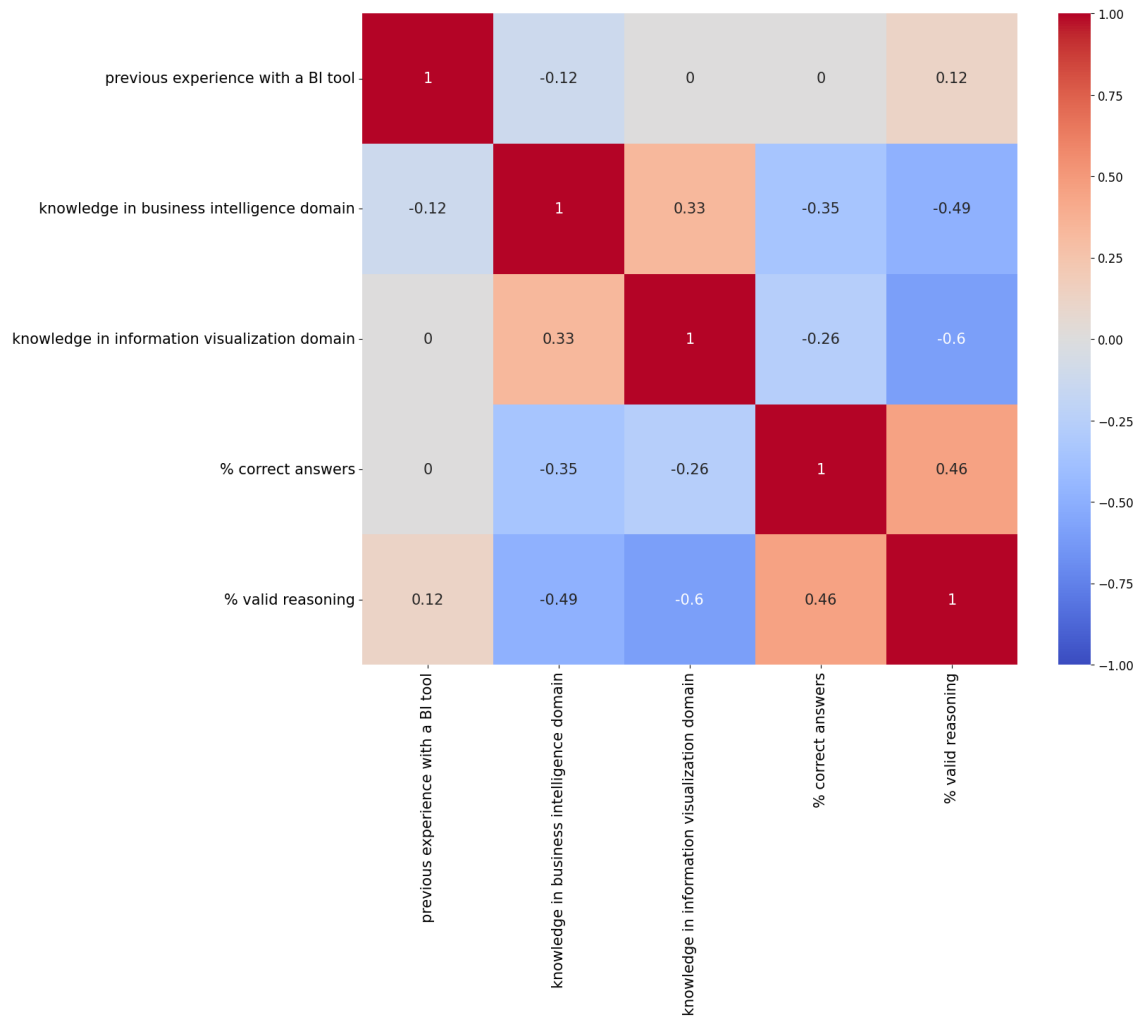
Figure 7.3: Correlation matrix between demographic data and percentage of correct answers and valid reasoning
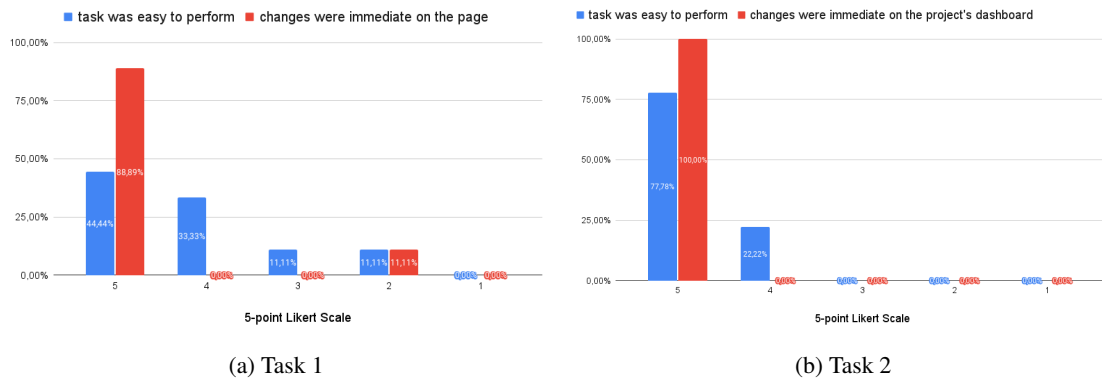
(a) Task 1

(b) Task 2

Figure 7.4: Evaluation of task ease and immediate visibility of changes for task 1 and task 2

task. The minimum and maximum times analysis further highlights the disparity between the two tasks. The first task consistently exhibited longer completion times compared to the second task. These findings indicate that participants generally spent more time on the initial task, suggesting that it may involve more complex or intricate steps that require additional consideration or exploration or that the participants learned how to do what was requested in the first task and for the second task, it was just the execution and almost no thought process was needed.

Table 7.3: Average, median, minimum, maximum, 1st quartile, 3rd quartile, and standard deviation of the timings for each task

Task	Avg.	Med.	Min.	Max.	1st Q.	3rd Q.	Std. Dev.
1	1:01:30	0:37:00	0:24:25	3:05:31	0:31:36	0:59:03	0:52:12
2	0:23:36	0:25:05	0:09:16	0:32:44	0:21:18	0:27:12	0:06:50

Participants were also asked to rate their perceived ease of completing the task and the immediacy of the changes made for each task. These ratings were collected using a 5-point Likert scale. The participants' responses are presented in Figure 7.4. Analyzing the results, it is evident that the first task was perceived as more challenging. Only 44.4% of participants strongly agreed that the task was easy to perform, indicating lower perceived ease. However, for the second task, the percentage of participants who strongly agreed that the task was easy to perform increased significantly to 77.8%. Regarding the immediacy of changes, the first task mainly received positive ratings. Eight out of nine participants strongly agreed that the changes made during the first task were immediately visible. However, one participant disagreed, indicating a slight discrepancy in their perception. In contrast, all participants strongly agreed that the changes made were immediately visible for the second task, demonstrating a unanimous perception of immediacy.

The final section of the form consisted of evaluating the usability of the developed tool using the SUS form. A score was calculated for each participant's response, and various statistical measures were derived from these scores, including the average, median, minimum, maximum, first quartile, third quartile, and standard deviation. The results of this evaluation are presented
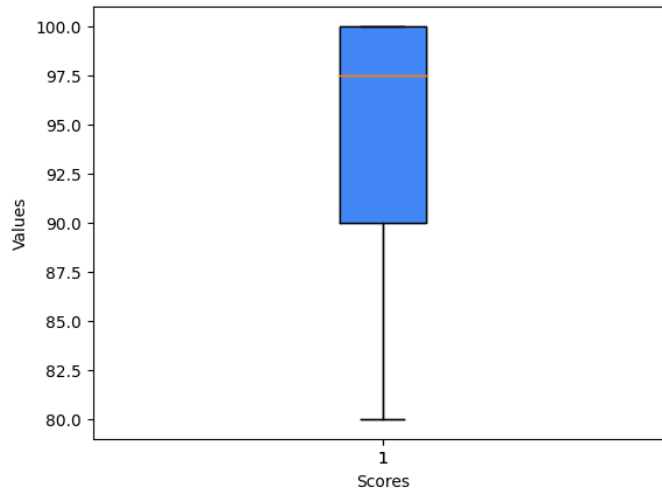
Figure 7.5: Boxplot of the SUS score

in Table 7.4 and Figure 7.5. The usability evaluation with the SUS form revealed an average score of 94.2, indicating a generally positive perception of the developed tool's usability. The median score of 97.5 suggests that most participants expressed a high level of agreement with the usability statements presented in the form. The scores obtained ranged from a minimum of 80 to a maximum of 100, highlighting the variation in the perceived usability among participants. The first quartile score of 90 and the third quartile score of 100 reflect the distribution of responses, indicating that most participants rated the tool's usability above 90. The standard deviation of 6.7 indicates a moderate level of dispersion in the scores, signifying some variation in participants' perceptions of usability.

Table 7.4: Average, median, minimum, maximum, 1st quartile, 3rd quartile, and standard deviation of the SUS score

Avg.	Med.	Min.	Max.	1st Q.	3rd Q.	Std. Dev.
94.2	97.5	90	100	90	100	6.7

7.2 Discussion

In evaluating the responses to the questions, it is evident that participants generally demonstrated a high percentage of correct answers and valid reasoning, indicating the tool's intuitive nature and effective display of information. Questions 1, 5, and 7 notably received perfect scores for both answers and valid reasoning, indicating that the tool successfully facilitated understanding and elicited accurate responses.

However, some areas for improvement became apparent in questions 2, 3, 4, and 6. In question 2, participants encountered confusion in differentiating between the two metrics displayed on the dashboard. The lack of clear differentiation led to some participants mistakenly counting all the

cards. To address this, revisions to the tool's design should provide more precise visual indications to distinguish between the two metrics, ensuring better comprehension for participants. Similarly, question 3 revealed difficulties related to the grouping of categories. Some participants struggled to understand the grouping, resulting in incorrect responses. Enhancements in the visual presentation or labeling of categories could alleviate this confusion and improve participants' comprehension and accuracy. Question 4 posed challenges for participants as it required them to answer two distinct aspects within a single question. Many participants only responded to half of the question, leading to a low percentage of correct answers. To mitigate this issue, separating the question into two distinct parts would allow participants to focus on each aspect individually and provide more accurate responses. Besides, the participants felt they needed clarification regarding the question itself, whether the best was the all-time best or the best of the data presented in the chart. Clear instructions and contextual information should also be provided to ensure participants interpret the question correctly. Regarding question 6, all participants provided correct reasoning, with only a small number failing to answer both parts. This suggests that presenting two distinct questions instead of a single question with multiple parts may yield better outcomes in terms of participant responses.

Participants predominantly relied on the charts or data table to answer the questions, deviating from the intended design where answers were distributed across different tool elements. Participants gravitated towards larger and more visually prominent elements, such as charts and data tables, for information retrieval. Enhancing the visibility and discoverability of relevant information throughout the tool's various sections could improve user engagement and facilitate accurate responses. Furthermore, participants' inclination to rely on charts or data tables highlights their preferred cognitive processing or decision-making styles. This observation underscores the importance of providing both representations of the data within the tool, as different users may have different preferences and utilize the representation that best aligns with their cognitive approach.

Correlations were calculated to examine the potential influence of participants' previous experience with a BI tool and their knowledge in the BI and information visualization domains on the questionnaire results. However, the results were different from what was expected. It was expected that the participants with previous experience or knowledge in the domain would perform better in the questionnaire, which did not agree with the correlations found. When working with small sample size, the variability in the data tends to increase, and the statistical power to detect significant correlations decreases. As a result, it becomes more challenging to identify meaningful relationships and draw accurate conclusions from the data. Therefore, the unexpected results could be because of the small number of participants.

The variations in response times can be attributed to the differing complexities of the questions. Question 1, the most straightforward, had the quickest average response time, while questions 4, 6, and 7, which required more cognitive involvement, exhibited longer response times. The range of response times, from the minimum to the maximum, can be justified by individual differences in reading speed and cognitive processing. These variations underscore the importance of considering cognitive load and processing demands when interpreting participants' responses.

Regarding the tasks performed during the usability evaluation, the overall success in task completion indicates a reasonable level of intuitiveness in the prototype. However, some areas need improvement, as reflected in the participants' ratings of the ease of the first task. The ratings for this task were dispersed, with one participant even expressing disagreement with its ease. This suggests that participants could accomplish the tasks but found the first task more challenging than intended. This is further supported by the maximum completion time for the first task, which was approximately 3 minutes. Upon analyzing the participants' explanations for the tasks, it became evident that a tab labeled 'Settings' was misleading, leading most participants to assume that the task should be performed within that tab. Additionally, one participant remarked that the settings icon next to the metric's name needed to be bigger. These findings highlight the need to revise the placement of metric editing to enhance usability. In contrast, the second task received undivided agreement from all participants regarding its ease of completion. The average time taken to complete the task also decreased, which aligns with expectations, considering that the primary challenge encountered in the first task was understanding where to edit metrics. In summary, the successful completion of the tasks by all participants indicates that, although improvements can be made to the metric's settings, the options for editing were generally clear and understandable.

7.3 Summary

In this chapter, the findings of the study were presented and analyzed. The results indicated that participants achieved a high percentage of correct answers and valid reasoning for most of the questions, suggesting that the prototype was intuitive and effectively conveyed the required information. However, specific questions exhibited lower performance and revealed areas that could be improved, precisely questions requesting two things.

The analysis of task completion times revealed variations based on the complexity of the tasks. The first task proved to be more complex than the second one. Although all participants accomplished the tasks, some areas could be improved, especially where in the prototype to edit a metric.

To conclude, although the prototype proved intuitive and easy to use, improvements can still be made.

Chapter 8

Conclusions

This chapter summarizes the work done and discusses the future work that could be done to improve the prototype developed.

8.1 Summary

In today's digital age, businesses across various industries have established a solid online presence, rapidly increasing the volume of data generated. The evaluation of business performance necessitates the analysis of extensive volumes of data from diverse sources. However, relying on separate reports from each source can be a time-consuming and inefficient process for managers. To enable informed decision-making, extracting valuable insights from data is crucial, rather than relying exclusively on intuition or guesswork.

This work aimed to develop a tool that empowers businesses to make data-driven decisions. The tool aimed to retrieve data from multiple sources and present it intuitively and visually appealingly. Usability was an essential requirement, ensuring that users with minimal or no technical knowledge could easily navigate the tool. This tool was particularly beneficial for ZOS LDA, as their current reporting process heavily relies on manual Excel-based workflows.

The background section established this work's fundamental concepts, including business intelligence and data-driven decision-making. Multiple definitions of BI were explored, with the chosen definition by Dresner stating the combination of data and analytic tools to provide information for decision-makers. The components of a BI system, encompassing data management, analytical processing, and knowledge discovery, were also discussed. Data-driven decision-making, closely tied to BI systems, was defined as making data-based decisions. The importance, benefits, and challenges of adopting a data-driven approach within organizations were explored, highlighting its connection to BI systems.

The state of the art analysis focused on market-leading BI tools, namely Power BI, Tableau, and Qlik Sense, as recognized by Gartner. The study of these tools encompassed their features,

such as available visualizations and collaboration capabilities, as well as their limitations. Although these tools are top-rated, an analysis of the users' reviews showed that these tools are yet to meet the full expectations and needs of the users. The study revealed some common complaints: complexity, cost, and ability to handle big datasets.

After defining the problem, a set of requirements was established to guide the development process. These contained various aspects, including actors, user stories, and technical requirements. The actors involved in the system were categorized into authenticated and unauthenticated users, with the authenticated user further divided into manager and employee roles. Additionally, the data sources themselves were considered actors within the system since they are involved in system interactions. To capture the functional aspects of the solution, user stories were employed. These user stories outlined the desired features and functionalities of the tool. Key features included authentication capabilities, project creation and management, data retrieval from multiple sources, source, and metric management, data visualization in various formats, data export functionality, notifications, and collaboration features to facilitate user interaction. In addition to the functional requirements, technical considerations were also considered. Usability was a fundamental focus, ensuring that the tool was intuitive and easy to navigate for users with varying levels of technical expertise. Security measures were considered to safeguard the data and ensure user privacy. Scalability was considered to accommodate potential growth and handle increasing volumes of data. Integration capabilities were also highlighted to enable interaction with other systems and data sources. Based on these requirements, a prototype was developed to validate the proposed solution.

The developed prototype consisted of two main components: the frontend implemented in ReactJS and the backend implemented in NodeJS with Express. The backend component established a connection to a MongoDB database, serving as the storage mechanism for the application. The primary objective of the prototype was to showcase the core functionalities and provide a visual representation of the planned tool. Due to time constraints, only user stories categorized as a high priority were fully implemented, with some medium priority stories also included. Consequently, not all user roles described in the problem definition, such as manager and employee, were incorporated into the prototype. Instead, the user stories of the manager role were considered to be of the authenticated user. The prototype enabled users to create and manage projects, facilitating the organization and tracking of relevant data. Data upload could be through manual sources, such as spreadsheets, or automatic sources, such as API. Upon data upload, the system automatically generated corresponding metrics based on the provided data. The visual representation of data was presented in an intuitive and visually appealing dashboard, which included a dedicated dashboard for each metric. Users could also manage data sources and projects. Furthermore, a customization feature was implemented, empowering users to modify metrics according to their preferences. This feature provided options to select and display specific metrics on the project's dashboard and choose the desired format for presenting the data.

An experiment was conducted to evaluate the usability and effectiveness of the developed prototype. The experiment focused on gathering participant feedback regarding their experience with

the tool. The experiment included collecting demographic data, such as whether the participants had prior experience with BI tools, their self-rated knowledge in the BI and information visualization domains on a 5-point Likert scale, and if they thought that business decisions should be data-driven and the usefulness of a tool that would retrieve the data and display it. This demographic data aimed to provide insights into the participants' backgrounds and expertise in relevant areas. The experimental process consisted of presenting participants with questions and tasks related to the prototype. The questions varied in terms of the level of interaction required from the user. Additionally, participants were assigned two tasks that focused on exploring and utilizing the customization feature of the prototype. The completion time of the questions and tasks was retrieved. To assess the overall usability of the prototype, participants were asked to fill out the SUS form.

The experiment results indicated that the developed tool was generally perceived as easy to use and did not require extensive prior knowledge or expertise. Participants found it accessible and intuitive, which aligned with the goal of creating a user-friendly solution. However, the results also revealed certain areas where improvements could be made. One significant aspect was the customization of metrics, which proved challenging for some users. Participants expressed difficulties in locating and editing metrics within the tool. Additionally, an attempt was made to explore the correlation between participants' prior knowledge and expertise in BI and information visualization domains and their success in the experiment. However, the results did not produce the expected correlations. It is worth noting that the sample size of participants was small, which could limit the ability to establish definitive correlations.

After the results, validating or refuting the established hypothesis is possible: "It is feasible to make a tool with a **small learning curve** that every organization can use. This tool is fed automatically with data from multiple sources, and the analysis it provides will be helpful when making **business decisions**.". Since the prototype could retrieve data from two sources and intuitively provide the information, its usage required no technical knowledge and had a user-friendly learning curve, the hypothesis is validated.

8.2   Future Work

In terms of future work, several areas can be identified for further exploration and improvement based on the findings and limitations of the current study.

Firstly, the customization feature identified as an area for improvement should be further explored and refined. This includes making the editing and customization of metrics more intuitive and user-friendly.

Secondly, expanding the participant pool for future experiments is essential to increase the accuracy and generalizability of the results. A more comprehensive understanding of the correlations between prior knowledge, expertise, and success in using the tool can be obtained by including a larger and more diverse group of participants.

Furthermore, incorporating sharing and collaboration features into the tool should be a priority based on the study of market leaders in BI tools. Enabling users to collaborate, share insights, and work together on projects can enhance the tool's usefulness and effectiveness for teamwork and decision-making processes.

Additionally, conducting performance tests is crucial, especially considering the tool's ability to handle large volumes of data and the importance of data history. Evaluating the tool's performance with sizable datasets can help identify potential bottlenecks and optimize efficiency and prepare for scale.

In conclusion, the developed prototype is a good starting point for addressing the problem. However, further improvements, such as refining the customization feature, expanding the participant pool, incorporating sharing and collaboration capabilities, and conducting performance tests, are necessary to enhance the tool's functionality, usability, and scalability.

References

[1] Sanchit Aggarwal. Modern Web-Development using ReactJS. *International Journal of Recent Research Aspects*, 5(1), 2018.

[2] Richard Berntsson Svensson and Maryam Taghavianfar. Toward Becoming a Data-Driven Organization: Challenges and Benefits. In Fabiano Dalpiaz, Jelena Zdravkovic, and Pericles Loucopoulos, editors, *Research Challenges in Information Science*, pages 3–19, Cham, 2020. Springer International Publishing.

[3] Bootstrap. Official Themes Built by Bootstrap Developers. https://themes.getbootstrap.com/official-themes/, 2023. (Accessed on 10/06/2023).

[4] John Brooke. SUS: A quick and dirty usability scale. *Usability Eval. Ind.*, 189, 11 1995.

[5] Timothy Chee, Lee-Kwun Chan, Min-Hooi Chuah, Chee-Sok Tan, Siew Fan Wong, and William Yeoh. Business Intelligence Systems: State-of-the-art Review and Contemporary Applications. *In Symposium on Progress in Information and Communication Technology*, 2, 01 2009.

[6] CW Fisher, E Lauria, I Chengalur-Smith, and RY Wang. *Introduction to Information Quality*. MITIQ Press, Cambridge, MA, 2006.

[7] Forbes. Beyond the Buzzword: What Does Data-Driven Decision-Making Really Mean? https://www.forbes.com/sites/tableau/2022/09/23/beyond-the-buzzword-what-does-data-driven-decision-making-really-mean/, 2023. (Accessed on 23/06/2023).

[8] Gartner. Magic Quadrant for Analytics and Business Intelligence Platforms. https://www.gartner.com/doc/reprints?id=1-2955ETOT&ct=220215&st=sb?ocid=lp_pg398450_gdc_comm_az, 2023. (Accessed on 30/06/2023).

[9] Gartner. Qlik Sense Likes & Dislikes. https://www.gartner.com/reviews/market/analytics-business-intelligence-platforms/vendor/qlik/product/qlik-sense, 2023. (Accessed on 30/06/2023).

[10] Gartner. Salesforce (Tableau) Likes & Dislikes. https://www.gartner.com/reviews/market/analytics-business-intelligence-platforms/vendor/salesforce-tableau/likes-dislikes?marketSeoName=analytics-business-intelligence-platforms&vendorSeoName=salesforce-tableau, 2023. (Accessed on 30/06/2023).

[11] Gartner. Top Microsoft power BI Likes & Dislikes. https://www.gartner.com/reviews/market/analytics-business-intelligence-platforms/

vendor/microsoft/product/microsoft-power-bi/likes-dislikes?
marketSeoName=analytics-business-intelligence-platforms&
vendorSeoName=microsoft&productSeoName=microsoft-power-bi, 2023.
(Accessed on 30/06/2023).

[12] K Gowthami and M R Pavan Kumar. Study on Business Intelligence Tools for Enterprise Dashboard Development. *International Research Journal of Engineering and Technology (IRJET)*, 04(04), 2017.

[13] Hannu Hannila, Risto Silvola, Janne Harkonen, and Harri Haapasalo. Data-driven Begins with DATA; Potential of Data Assets. *Journal of Computer Information Systems*, 62(1):29–38, 2022.

[14] MongoDB Inc. What Is MongoDB? https://www.mongodb.com/en/what-is-mongodb, 2023. (Accessed on 10/06/2023).

[15] H. P. Luhn. A Business Intelligence System. *IBM Journal of Research and Development*, 2(4):314–319, 1958.

[16] Microsoft. What is Power BI? https://learn.microsoft.com/en-us/power-bi/fundamentals/power-bi-overview, 2023. (Accessed on 30/06/2023).

[17] A. Nayak, A. Poriya, and Dikshay Poojary. Article: Type of nosql databases and its comparison with relational databases. *International Journal of Applied Information Systems*, 5:16–19, 01 2013.

[18] Solomon Negash and Paul Gray. Business Intelligence. In *Commun Assoc Inf Sys*, volume 13, page 423, January 2003.

[19] Celina Olszak and Ewa Ziemba. Approach to Building and Implementing Business Intelligence Systems. *Interdisciplinary Journal of Information, Knowledge, and Management*, 2:135–148, 01 2007.

[20] Qlik. Qlik Sense | Modern Cloud Analytics. https://www.qlik.com/us/products/qlik-sense#broadusecases, 2023. (Accessed on 30/06/2023).

[21] Indra Kharisma Raharjana, Daniel Siahaan, and Chastine Fatichah. User Stories and Natural Language Processing: A Systematic Literature Review. *IEEE Access*, 9:53811–53826, 2021.

[22] Salesforce. Our Products. https://www.tableau.com/products, 2023. (Accessed on 30/06/2023).

[23] Salesforce. What is Tableau? https://www.tableau.com/why-tableau/what-is-tableau, 2023. (Accessed on 30/06/2023).

[24] SelectHub. Power BI Pricing, Demo & Review. https://www.selecthub.com/p/business-intelligence-tools/power-bi/, 2023. (Accessed on 30/06/2023).

[25] SelectHub. Qlik Sense Pricing, Demo & Review. https://www.selecthub.com/p/business-intelligence-tools/qlik-sense/, 2023. (Accessed on 30/06/2023).

[26] SelectHub. Tableau Pricing, Demo & Review. https://www.selecthub.com/p/business-intelligence-tools/tableau/, 2023. (Accessed on 30/06/2023).

[27] Uthayasankar Sivarajah, Muhammad Mustafa Kamal, Zahir Irani, and Vishanth Weerakkody. Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70:263–286, 2017.

[28] Tableau. Data-Driven Decision Making. https://www.tableau.com/learn/articles/data-driven-decision-making, 2023. (Accessed on 23/06/2023).

[29] Jianwen Wang, Abdullah Omar, Fahad Alotaibi, Dr.Yousef Daradkeh PhD., P.Eng, and Sara Althubiti. Business intelligence ability to enhance organizational performance and performance evaluation capabilities by improving data mining systems for competitive advantage. *Information Processing & Management*, 59:103075, 11 2022.

Appendix A

ZOS LDA May report

This appendix contains some screenshots of the May report of ZOS LDA.
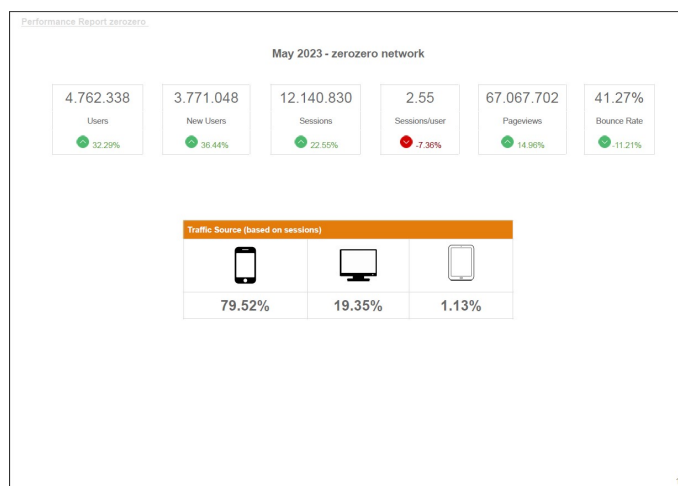
Figure A.1: zerozero network
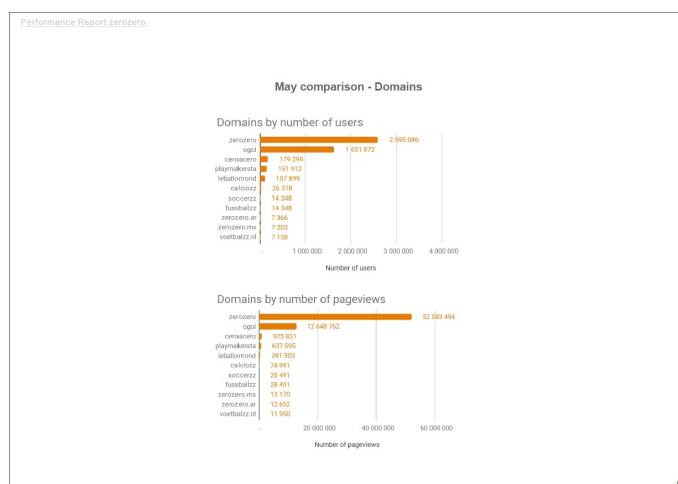
Figure A.2: Domains comparison
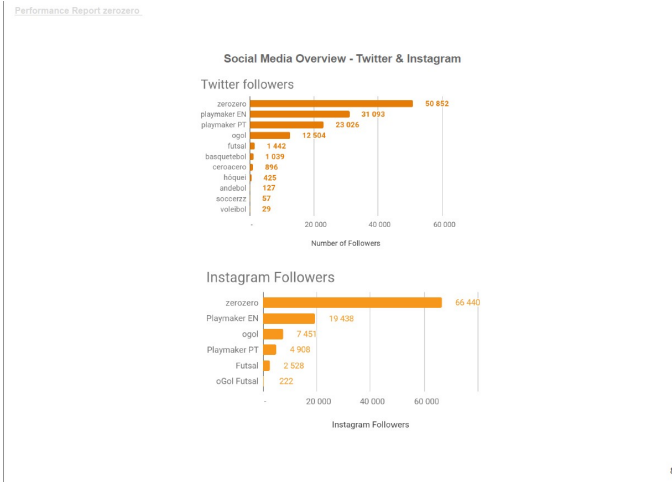
Figure A.3: Social media overview — Twitter & Instagram

Figure A.4: Overall sports performance May 2022 vs May 2023

Appendix B

User Interface

This appendix contains screenshots of the user interface of the prototype developed.
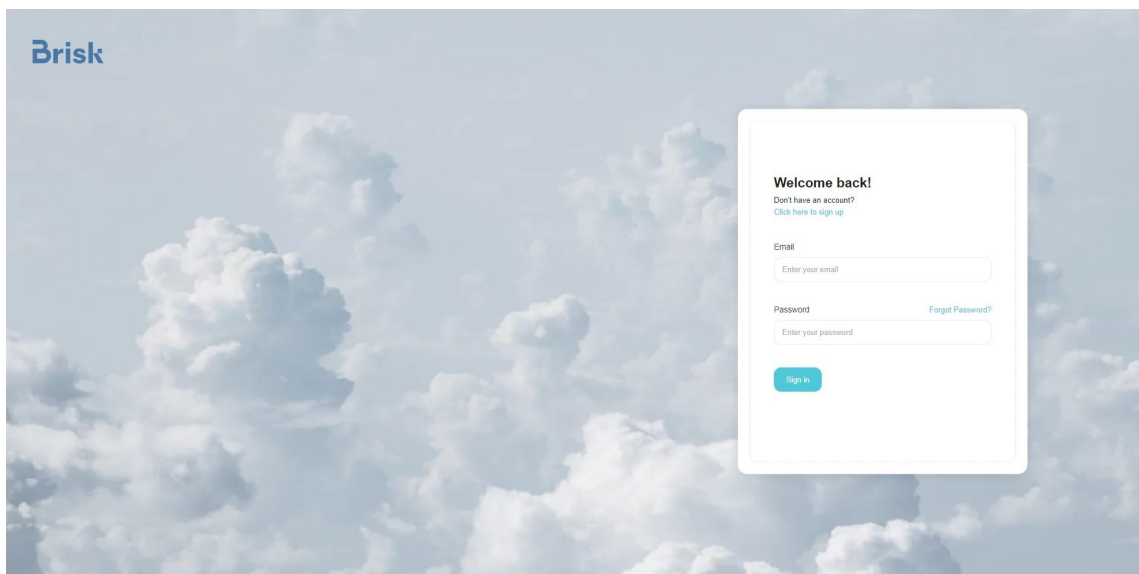
Figure B.1: Sign in page
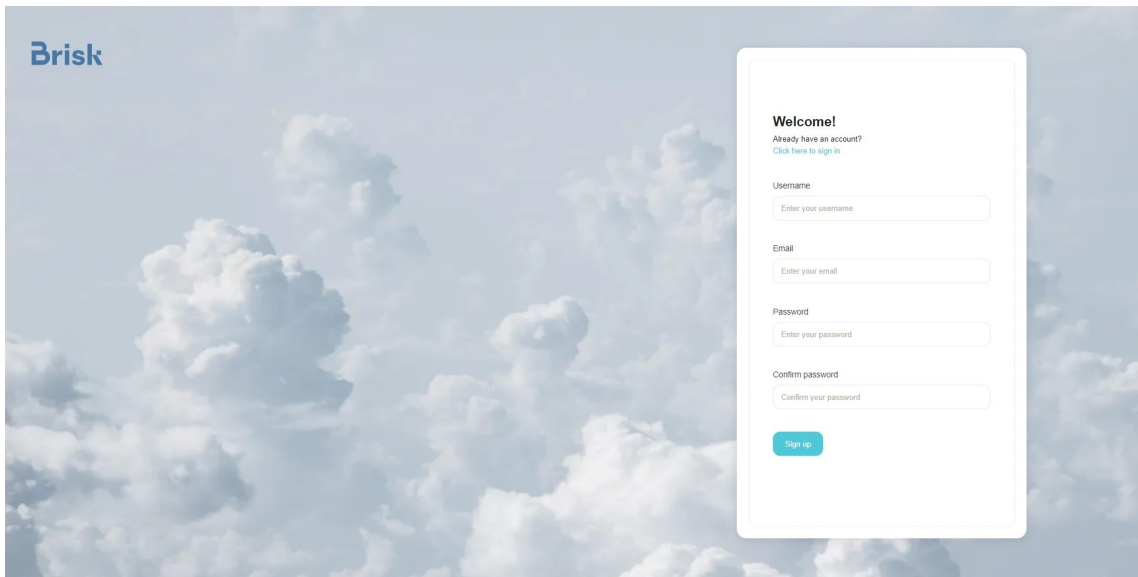
Figure B.2: Sign up page
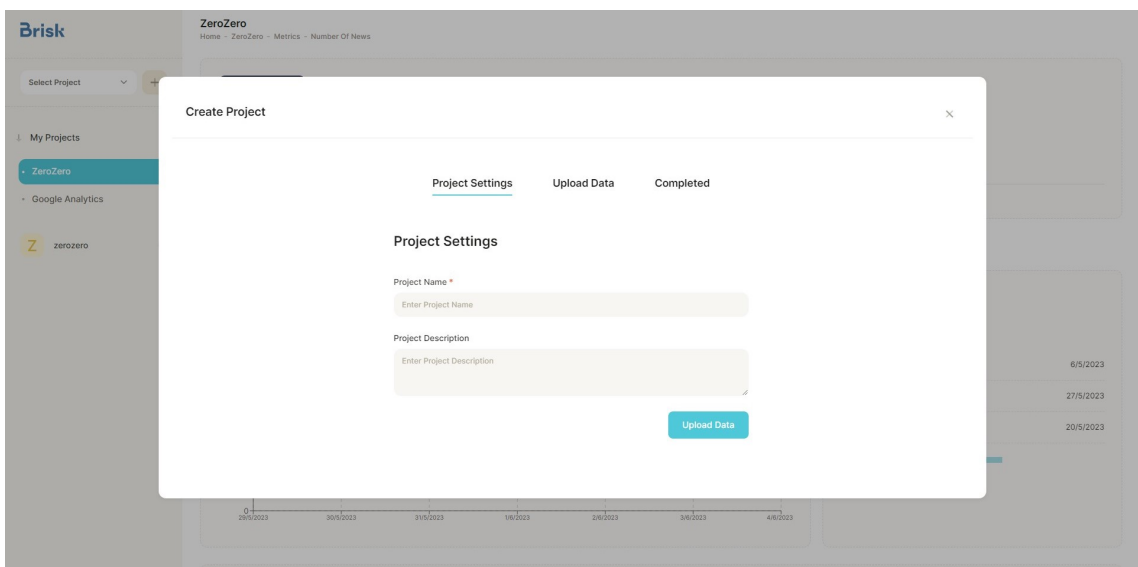
Figure B.3: Create project modal — project settings
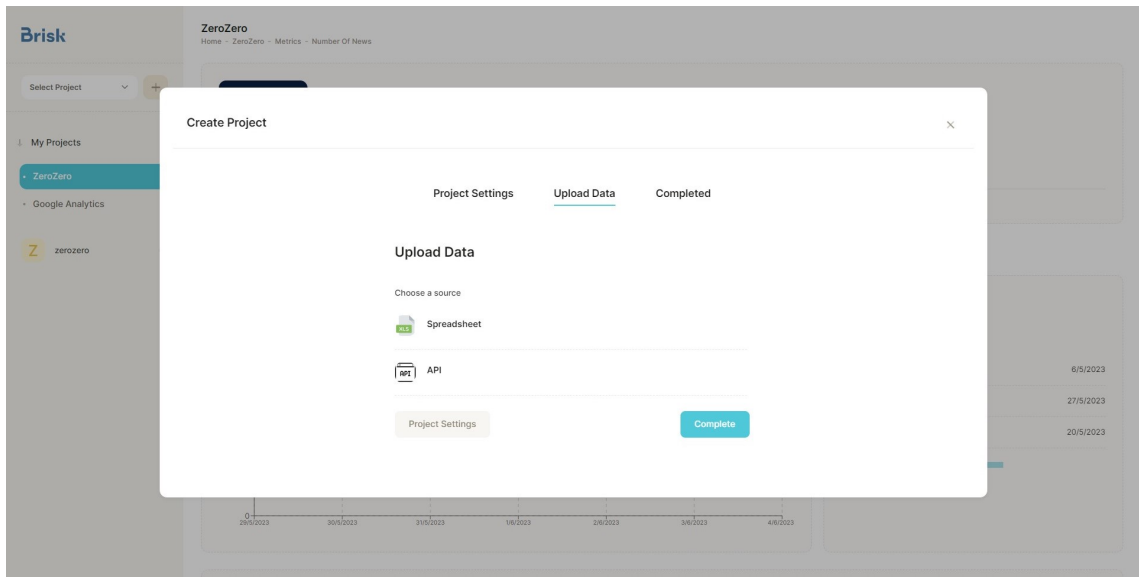
Figure B.4: Create project modal — upload data
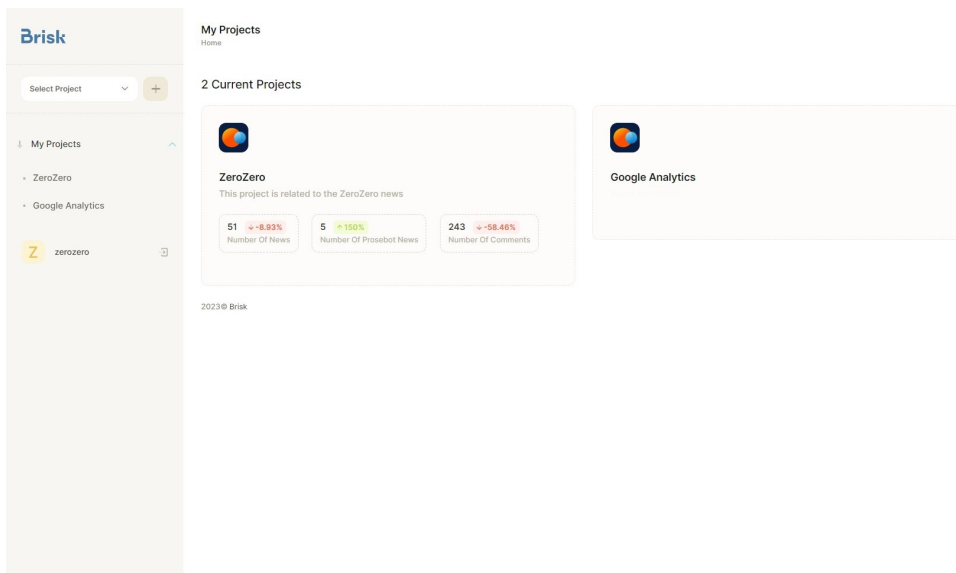
Figure B.5: My projects page
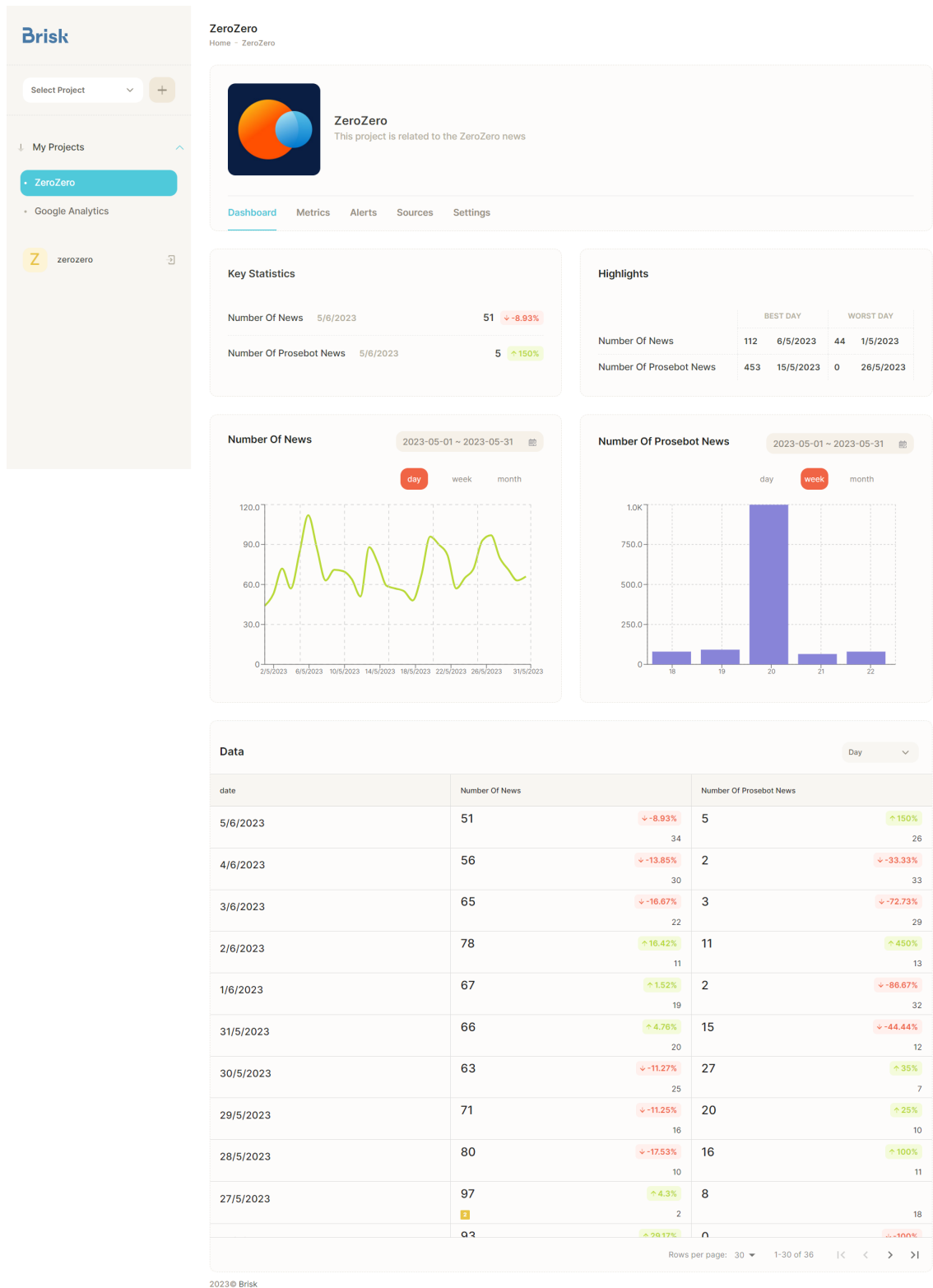
Figure B.6: Project's dashboard page
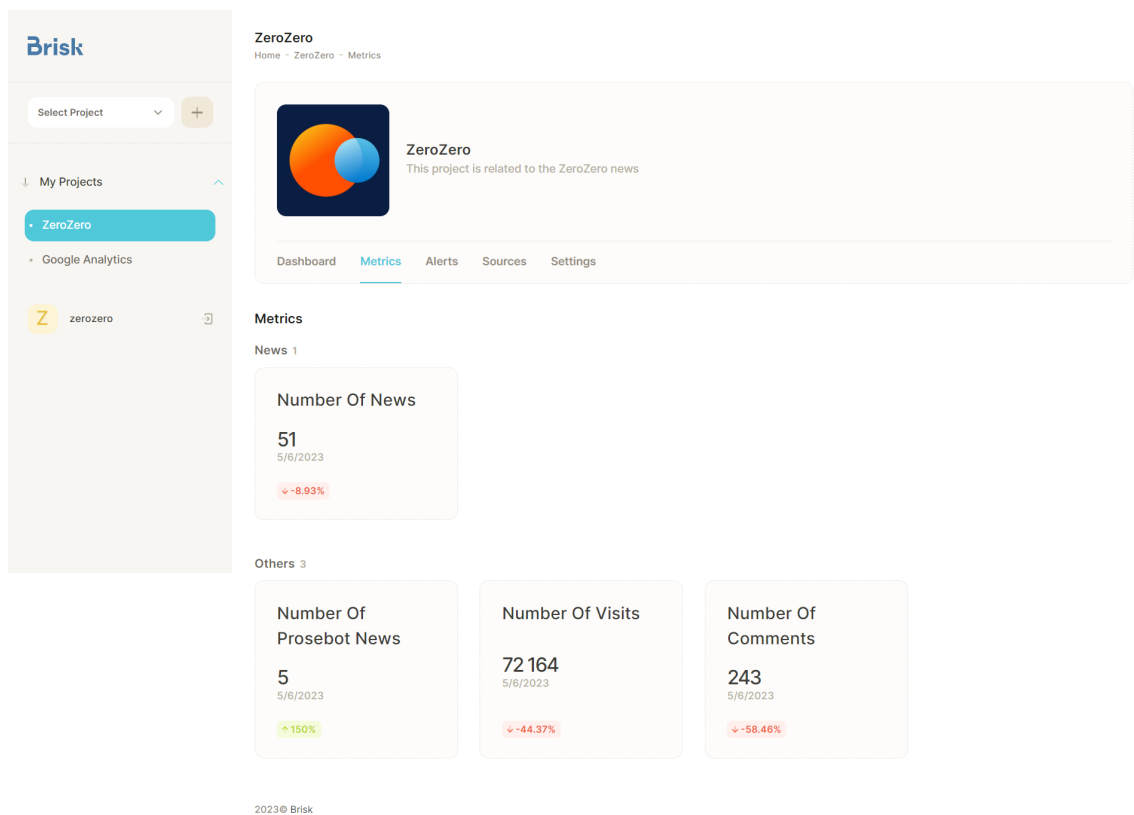
Figure B.7: Project's metrics page
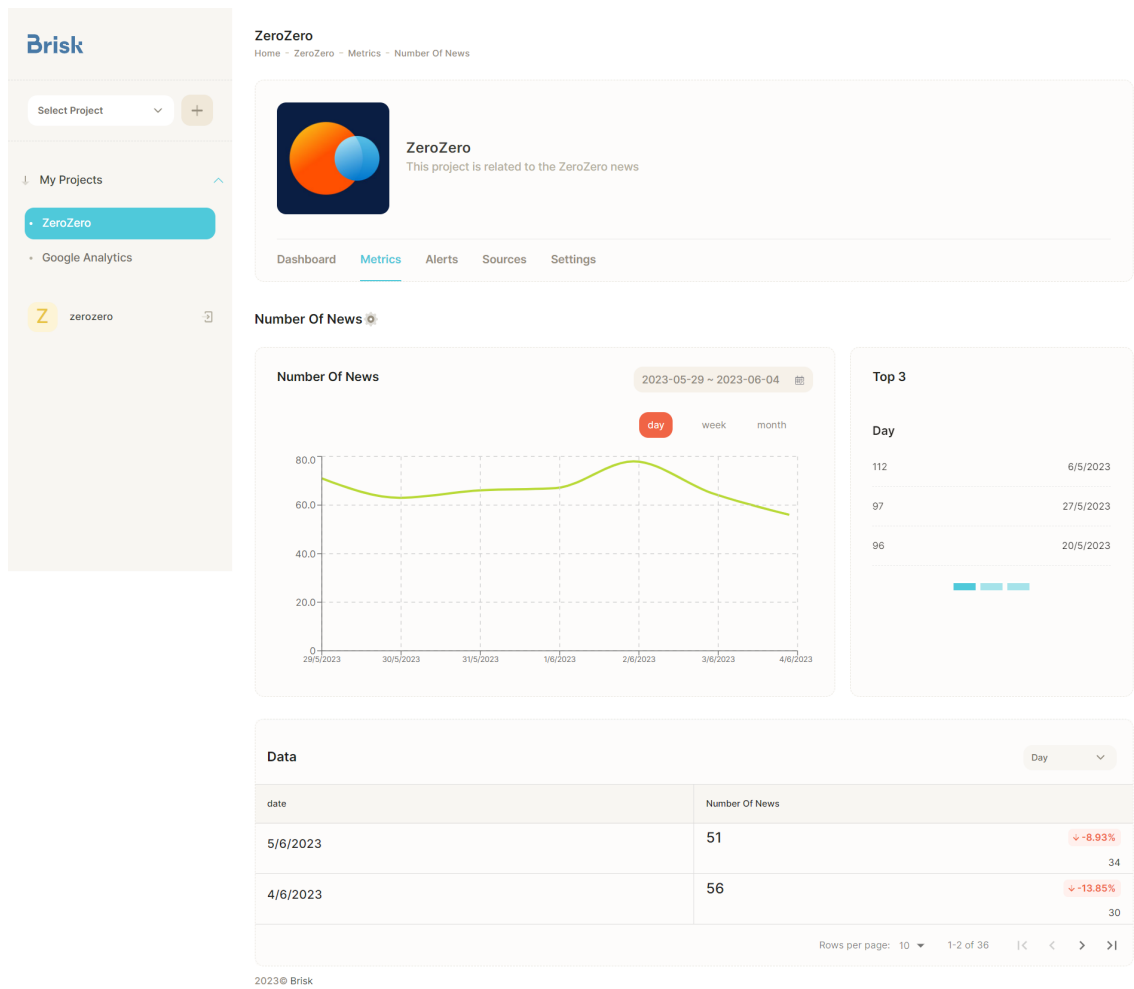
Figure B.8: Project's metric page
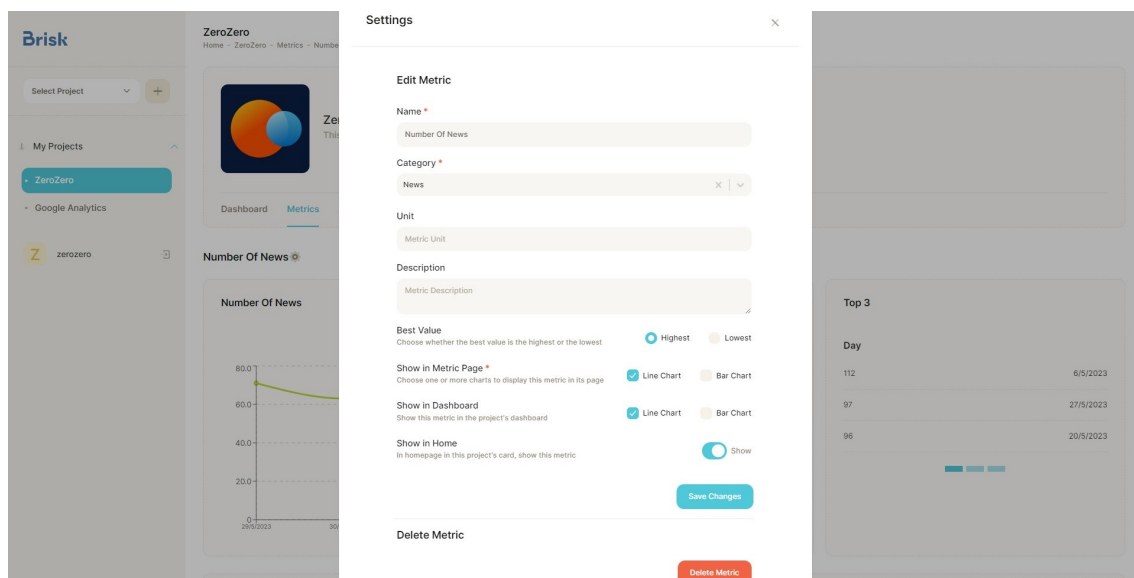
Figure B.9: Project's metric settings modal
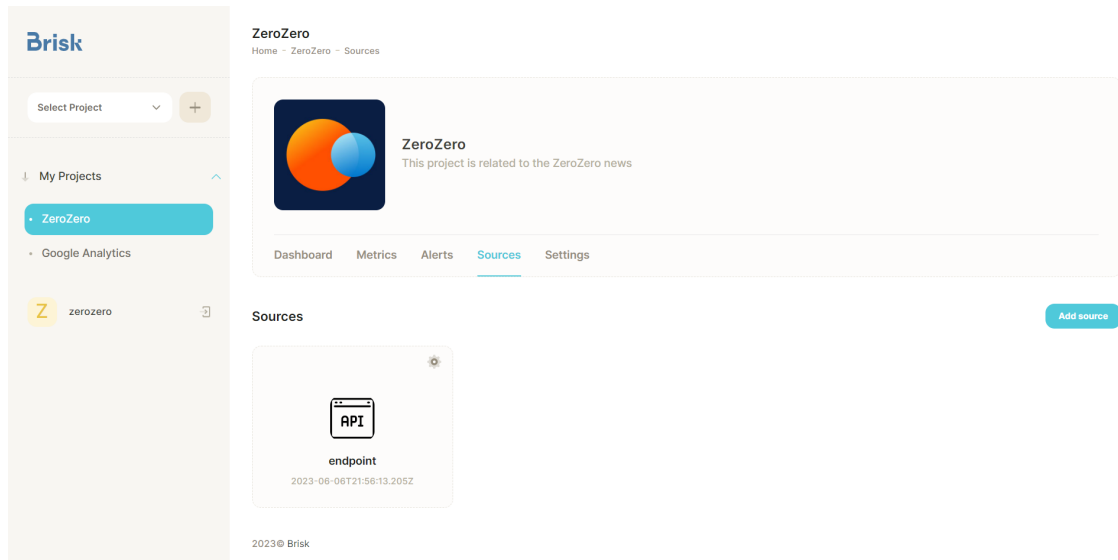
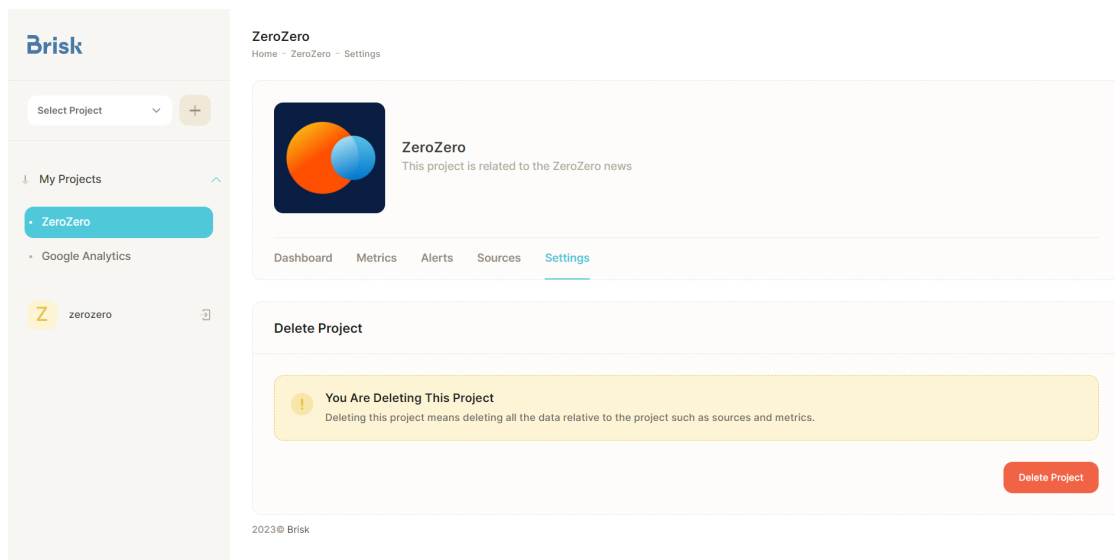Figure B.10: Project's sources page
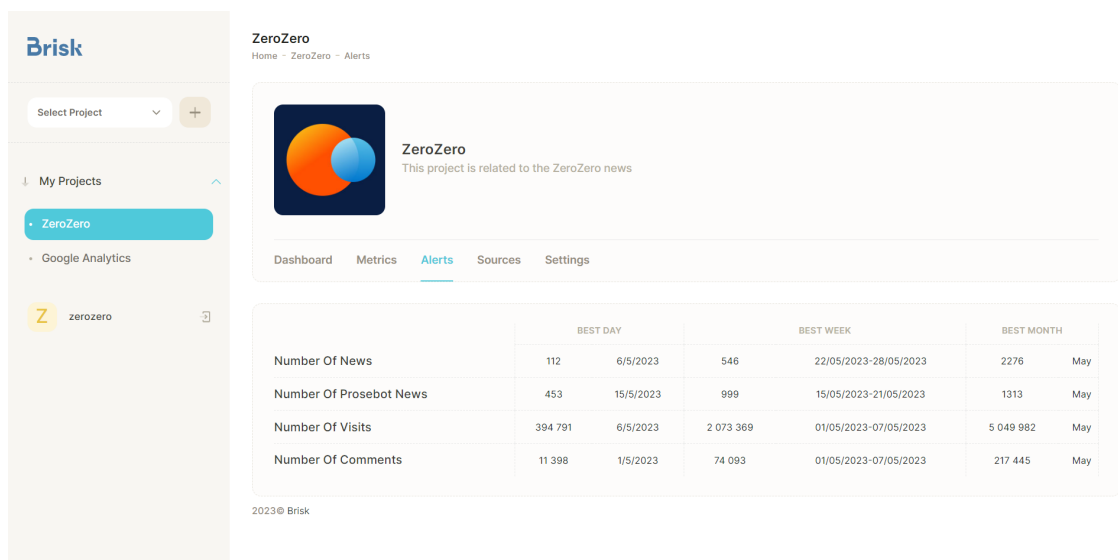
Figure B.11: Project's alerts page

Figure B.12: Project's settings page

Appendix C

Usability Test Form

This appendix contains the Google Forms used in the experiment, which comprises four parts: consent and demographic data, questionnaire, tasks, and SUS.

BRISK - Advanced tool for aggregation and visualization of atomized data

PROPÓSITO DO ESTUDO

Está convidado a participar num estudo de investigação em relação à usabilidade de uma ferramenta de *business intelligence*, que está a ser realizada pela aluna Ana Cruz no âmbito da dissertação de mestrado em Engenharia Informática e Computação na Faculdade de Engenharia do Porto, tendo como supervisor na empresa ZOS LDA o Diretor Técnico Marco Sousa e como orientador o professor João Correia Lopes.

PROCEDIMENTOS DO ESTUDO

Irão ser colocadas algumas questões pessoais como idade, género e se tem algum conhecimento na área de *business intelligence* para estatísticas. De seguida, serão colocadas questões relacionadas com o conteúdo da ferramenta, que terão de ser justificadas. Por fim, são colocadas algumas questões de usabilidade.

As questões que envolvem a utilização da ferramenta desenvolvida serão cronometradas pela aluna Ana Cruz.

O estudo termina com a resposta e a submissão do questionário.

Nenhuma informação recolhida será partilhada, exceto informação pessoal estatística e não identificável.

DURAÇÃO

Este estudo tem uma duração estimada de 10 minutos.

PARTICIPAÇÃO VOLUNTÁRIA

A sua participação neste estudo é totalmente voluntária. Irá decidir se pretende participar ou não. Caso decida participar, irá preencher os campos seguintes que irão confirmar o seu consentimento e a sua participação. Poderá desistir durante o estudo e toda a informação submetida será destruída.

Ao clicar em continuar está a consentir com a recolha de toda a informação mencionada anteriormente.

* Indica uma pergunta obrigatória

Informação Estatística

Por favor preencha a seguinte informação.

1. Idade *

Marcar apenas uma oval.

○ <18

○ 18-24

○ 25-34

○ 35-44

○ 45-54

○ 55-64

○ >=65

2. Género *

Marcar apenas uma oval.

○ Masculino

○ Feminino

○ Prefiro não responder

3. Já alguma vez usou alguma ferramenta de *business intelligence*? *
Exemplos: PowerBI; Tableau

Marcar apenas uma oval.

○ Sim

○ Não

4. Se sim, qual?

Marcar tudo o que for aplicável.

☐ PowerBI

☐ Tableau

☐ QlikSense

☐ Outra: _____

5. De 1 (Não domino) a 5 (Domino) avalie o seu conhecimento em: *

Marcar apenas uma oval por linha.

	1	2	3	4	5
No domínio da área de business intelligence	◯	◯	◯	◯	◯
No domínio de visualização de informação	◯	◯	◯	◯	◯

6. Considera importante que um negócio tome decisões fundadas em dados do próprio   *
negócio?

Marcar apenas uma oval.

◯ Sim

◯ Não

7. Considera que uma ferramenta que permitisse a recolha de dados relevantes ao negócio *
e a sua visualização seria útil?

Marcar apenas uma oval.

◯ Sim

◯ Não

Ferramenta Desenvolvida

A partir desta secção as questões são relacionadas com a ferramenta desenvolvida. Por favor use a ferramenta para procurar as respostas.

Ferramenta Desenvolvida

8. 1.Quantos projetos podem ser vistos? *

_____

9. Por favor explique o seu raciocínio. *

_____

Ferramenta Desenvolvida

As próximas questões são relacionadas com o projeto ZeroZero e focam-se na visualização da informação.
Considere que as datas encontram-se no formato DD-MM-YYYY. Considere também que a semana começa à segunda-feira e acaba ao domingo.

10. 2.Quantas métricas consegue ver na *dashboard* do projeto? *

_____

11. Por favor explique o seu raciocínio. *

_____

12. 3.Quantas métricas tem o projeto no total? *

_____

13. Por favor explique o seu raciocínio. *

_____
_____
_____
_____
_____

14. 4.Qual foi o melhor dia e o melhor valor da métrica *Number Of News?* *
Como melhor dia entende-se o dia em que a métrica atingiu o maior valor.

_____

15. Por favor explique o seu raciocínio. *

_____

_____

_____

_____

16. 5.Qual foi o valor atingido pela mesma métrica no dia 9 de maio de 2023? *

_____

17. Por favor explique o seu raciocínio. *

_____

_____

_____

_____

_____

18. 6.Nesse dia, para a mesma métrica, qual foi a percentagem de crescimento e o *ranking*? *
A percentagem de crescimento pode ser tanto positiva como negativa.

_____

19. Por favor explique o seu raciocínio. *

_____

_____

_____

_____

_____

20. 7.Qual o valor semanal atingido pela mesma métrica na semana de 15 de maio de   *
2023?

_____

21. Por favor explique o seu raciocínio. *

_____

_____

_____

_____

_____

Ferramenta Desenvolvida

As próximas questões são relacionadas com o projeto ZeroZero e focam-se na customização.

Tarefa 1

Edite o tipo de gráfico da métrica *Number of Comments* na página da própria métrica para gráfico de linhas.

22. Por favor explique o seu raciocínio. *

_____

_____

_____

_____

_____

23. Avalie de 1 (Discordo completamente) a 5 (Concordo completamente) as seguintes afirmações relativamente à primeira tarefa. *

Marcar apenas uma oval por linha.

	1	2	3	4	5
A tarefa foi fácil de realizar.	◯	◯	◯	◯	◯
As alterações foram visíveis de imediato na página.	◯	◯	◯	◯	◯

Tarefa 2
Edite a métrica *Number Of Comments* de maneira a que apareça na *dashboard* através da representação gráfico de barras.

24. Por favor explique o seu raciocínio. *

_____

_____

_____

_____

_____

25. Avalie de 1 (Discordo completamente) a 5 (Concordo completamente) as seguintes   *
     afirmações relativamente à segunda tarefa.

Marcar apenas uma oval por linha.

	1	2	3	4	5
A tarefa foi fácil de realizar.	◯	◯	◯	◯	◯
As alterações foram visíveis de imediato na dashboard do projeto.	◯	◯	◯	◯	◯

Usabilidade da ferramenta desenvolvida

26. Avalie de 1 (Discordo completamente) a 5 (Concordo completamente) as seguintes *
afirmações relativamente à usabilidade da ferramenta desenvolvida.

Marcar apenas uma oval por linha.

	1	2	3	4	5
Gostaria de utilizar esta ferramenta com frequência.	◯	◯	◯	◯	◯
Esta ferramenta é desnecessariamente complexa.	◯	◯	◯	◯	◯
Esta ferramenta é fácil de utilizar.	◯	◯	◯	◯	◯
Precisaria de ajuda para poder utilizar esta ferramenta.	◯	◯	◯	◯	◯
As várias funções desta ferramenta estão bem integradas.	◯	◯	◯	◯	◯
Esta ferramenta é demasiado inconsistente.	◯	◯	◯	◯	◯
A maioria das pessoas aprenderia a utilizar esta ferramenta muito rapidamente.	◯	◯	◯	◯	◯
Esta ferramenta é complicada/incómoda de utilizar.	◯	◯	◯	◯	◯
Senti-me confiante a utilizar esta ferramenta.	◯	◯	◯	◯	◯
É preciso adquirir conhecimentos novos antes de poder utilizar esta ferramenta.	◯	◯	◯	◯	◯

27. Comentários/Observações

_____

_____

_____

_____