UNIVERSIDADE CATÓLICA PORTUGUESA

Data Normalization and The Role of Power BI for Visualization in:

Enhancing Product Sales Decisions

Alexandre Branco de Brito

Católica Porto Business School

April 2025

UNIVERSIDADE CATÓLICA PORTUGUESA

Data Normalization and The Role of Power BI for Visualization in:

Enhancing Product Sales Decisions

Final Work in Organizational Context presented to Universidade Católica Portuguesa in order to obtain the master's degree in Management with a Specialization in Business Analytics

by

Alexandre Branco de Brito

Under supervision of
Maria Alice Trindade

Católica Porto Business School, Universidade Católica Portuguesa
April 2025

Acknowledgments

I would like to express my deepest gratitude to everyone who has supported me throughout the development of this dissertation.

I am especially thankful to my family for their unwavering support, continuous encouragement, and constant belief in me.

I am genuinely grateful to my supervisor, Maria Alice Trindade, for her guidance, availability, and insightful feedback, which were essential to completing this work.

I am fortunate to have had the opportunity to carry out my internship at Salvador Caetano Africa, where I was able to apply and expand my knowledge in a real and challenging environment.

I am particularly indebted to Daniela and Rita for their patience, constant support, and willingness to help me grow personally and professionally throughout the internship.

Finally, I am grateful to my friends for always being by my side, whether with words of motivation, shared laughs, or just being there.

Resumo

Na era digital atual, os dados tornaram-se um dos ativos mais valiosos para as organizações, sendo a capacidade de os transformar em insights acionáveis uma vantagem competitiva crucial. A Business Intelligence tem vindo a assumir um papel cada vez mais relevante no apoio à tomada de decisão, ao permitir que as empresas interpretem grandes volumes de dados através de processos estruturados e suportados por tecnologia.

Esta dissertação de mestrado tem como objetivo desenvolver um dashboard de apoio à decisão que facilite a análise de dados e promova a decisões estratégicas mais eficazes. Para isso, foi adotada a metodologia de *Action-Research,* aplicada ao contexto da Salvador Caetano Africa, que serve como caso de estudo. O trabalho decorreu em dois ciclos iterativos: o primeiro dedicado à normalização dos dados, e o segundo na conceção de um dashboard.

As principais conclusões deste trabalho destacam a importância de um dataset bem estruturado como base para qualquer solução de BI, a interligação entre a preparação dos dados e o desenvolvimento do dashboard, bem como a relevância da utilização de Data Analysis Expressions para criar visualizações dinâmicas e significativas.

Em última análise, este projeto permitiu uma redução significativa do tempo necessário para a análise de posicionamento de preços, aumentou a fiabilidade dos dados e melhorou a produtividade global.

Palavras-chave: Ferramentas de Business Intelligence, Tomada de Decisão, Visualização de dados, Dashboard, Normalização de Dados

Palavras: 7362

Abstract

In the current digital age, data has become one of organizations' most valuable assets, and the ability to transform this data into actionable insights is a key competitive advantage. Business Intelligence is growing in supporting decision-making by helping companies make sense of large volumes of data through structured, technology-driven processes.

This master's dissertation aims to develop a decision support dashboard that facilitates data analysis and promotes more effective strategic decision-making. To achieve this, the Action-Research methodology was adopted and applied to the context of Salvador Caetano Africa, which serves as the case study. The work was carried out in two iterative cycles: the first focused on data normalization, and the second on the design of the dashboard.

The main findings of this work highlight the importance of a well-structured dataset as the foundation for any BI solution, the interconnection between data preparation and dashboard development, and the relevance of using Data Analysis Expressions to create dynamic and meaningful visualizations.

Ultimately, this project significantly reduced time spent on price positioning analysis, enhanced data reliability, and improved overall productivity.

Keywords: Business Intelligence Tools, Decision Making, Data Visualization, Dashboard, Data Normalization

Words: 7362

Contents

ACKNOWLEDGMENTS III

RESUMO V

ABSTRACT VII

CONTENTS IX

TABLE OF FIGURES XII

TABLE OF TABLES XIV

CHAPTER 1 16

1 Introduction 16

1.1 Report Restructuring Goals 16

1.2 Research Methodology 17

1.3 Dissertation Structure 17

CHAPTER 2 19

1 Literature Review: Data Normalization and Visualization and Its Impact on Decision-Making 19

1.1 Business Intelligence 19

1.1.1 BI Benefits in Decision-Making 21

1.2 Data Visualization 23

1.3 Data Normalization 31

1.4 Integrating BI, DV, and DN 34

CHAPTER 3 36

1 Research Methodology 36

1.1 Project Goals 36

1.2 Research Contextualization 37
1.3 Method 38

CHAPTER 4 40

1 Database Normalization 40
1.1 Problem definition 40
1.2 Data Normalization 42

CHAPTER 5 47

1 Power BI Dashboard 47
1.1 General Page 47
1.2 Specific Vehicle Page 49
1.3 Equipment Benchmark Page 51
1.4 Ranking page 54
1.5 Model Releases Page 55
1.6 Information page 55

CHAPTER 6 57

1 Results Discussion 57
2 Final Conclusions 58
3 Limitations and possible future work 59

PROMPTS LIST 60
BIBLIOGRAPHY 61
APPENDIX 66

1 Product Manager Feedback Questionnaire 66

Table of Figures

Figure 1: Magic Quadrant 23

Figure 2: Power BI Components (Microsoft, 2024) 25

Figure 3: Action Research Cycle 38

Figure 4: Action Research Adapted Model 39

Figure 5: Previously Used Tables 41

Figure 6: Examples of created tables 43

Figure 7: Market - Country Table 43

Figure 8: Car_Principal Table 44

Figure 9: Car_Equipment Table 45

Figure 10: Car_Price Table 46

Figure 11: General Page – Table 48

Figure 12: General Page - Price Positioning and Evolution Analysis by Brand 49

Figure 13: Specific Vehicle Page - Table 50

Figure 14: Specific Vehicle Page - Scatter Plot 51

Figure 15: Equipment Benchmark Page - Table and Bar Graph 52

Figure 16: Equipment Benchmark Page - Red Flag Table 53

Figure 17: Ranking Page 54

Figure 18: Model Releases Page 55

Figure 19: Information Page 56

Table of Tables

Table 1: BI Definitions 20

Table 2: BI Benefits 22

Table 3: BI Tools 24

Table 4: Power BI vs. Excel and Tableau: A Performance and Decision-Making Comparison 27

Table 5: DAX functions (Microsoft, 2023) 28

Table 6: Data Normalization Key Concepts 31

Table 7: Before 1NF 32

Table 8: After 1NF 32

Table 9: 2NF Tables 33

Table 10: 3NF Table 33

Table 11: Initial DAX Functions 49

Table 12: Additional DAX Functions for Competitive Equipment Analysis . 53

Chapter 1

1 Introduction

This dissertation was developed as part of an internship at Salvador Caetano Africa (SCA), where the company proposed creating a Power BI report to enhance sales decision-making. The project aimed to address challenges in price positioning by resolving inconsistencies in data entry and formatting, which significantly increased the time required for analysis and decision-making.

This report was designed to be a key tool for the company's product managers, allowing them to analyze vehicle positioning and make better-informed strategic decisions.

SCA is a subsidiary of Grupo Salvador Caetano (GSC), one of the largest and most influential Portuguese automotive companies, with operations in over forty-four countries.

I was part of the Innovation team within the Logistics, Network Development, and Projects department, where this project was developed in collaboration with the IT Department and the Product Team. These teams' cross-functional efforts made the dashboard possible, ensuring alignment with technical requirements and business objectives.

1.1 Report Restructuring Goals

Since the dissertation relies on an internship report, it does not answer a specific research question. The primary purpose of this dissertation is to enhance information integration to enable improved monitoring and analysis, using business intelligence tools. The main goal is to work on tools such as Power BI to evaluate which products should be launched in Africa. For that, some objectives need to be followed:

- Understand the company's key evaluation criteria and methods used
- Develop a plan to normalize the data gathered by the company
- Integrate the data into Power BI
- Create a detailed dashboard that updates every day and influences the daily tasks of product managers

1.2 Research Methodology

Regarding the method, an action research methodology would be the best approach since it is focused on practical, real-world applications rather than just theory. "*In action research, the research problem studied is emerging from the context of investigation (e.g. an organization)*". (Petersen et al., 2014, p.2)

This dissertation was developed during an internship at Salvador Caetano Africa. It will serve as a case study to demonstrate the applicability of normalization, visualization, and the usage of Power BI in assisting product sales decisions.

1.3 Dissertation Structure

This dissertation is structured into six chapters, each detailing different aspects of the project, from its context and theoretical foundations to its implementations and conclusions.

Chapter 1 introduces the context in which this study was developed, outlines the main objectives, the methodology adopted, and the overall structure of the dissertation.

Chapter 2 presents the theoretical framework, covering essential concepts related to Business Intelligence (BI), Data Visualization (DV), and Data Normalization (DN). These areas provide the foundation for understanding the project's relevance and impact.

Chapter 3 contextualizes the internship, detailing the company's challenges. It also details the project goals and provides a theoretical background on the methodology applied.

Chapter 4 focuses on the technical aspects of DN, detailing the issues encountered in the original dataset and the steps taken to optimize it. This chapter describes the restructuring, ensuring data accuracy and consistency for effective visualization in Power BI.

Chapter 5 presents the development of the Power BI dashboard, explaining its design, key visualizations, and functionalities. It demonstrates how the dashboard enhances decision-making by providing actionable insights.

Finally, Chapter 6 summarizes the findings, reflects on the overall outcomes of the project, and evaluates the extent to which the initial objectives were met. It also discusses the main limitation and suggests potential directions for future improvements and research.

Chapter 2

1 Literature Review: Data Normalization and Visualization and Its Impact on Decision-Making

The present chapter provides the theoretical foundation for this master's thesis by reviewing key concepts and previous research relevant to the development of decision-support systems. It aims to comprehensively understand how BI, DV, and DN contribute to more effective decision-making in organizational contexts.

The literature review is organized to progressively analyze these three components: BI, DV, and DN. It starts by defining and discussing BI's evolution, then explores data visualization practices, and concludes with the role of data normalization as a critical step for high-quality data analysis. Each section highlights the connection between these concepts and their application in decision-making.

Additionally, this chapter identifies gaps in existing literature, especially the lack of integrated research on the combined impact of BI, DV, and DN, and positions this dissertation as a contribution toward bridging that gap.

1.1 Business Intelligence

We are currently experiencing an era of rapid technological advancement. Digital innovation has significantly transformed our daily lives, with the business sector being particularly affected (Stark, n.d.). To stay competitive in today's fast-paced market, companies must now embrace innovation and adaptability (Alzghoul et al., 2024). As a result, many organizations have implemented technologies such as BI to improve their decision-making processes and sustain a competitive advantage.

A wide variety of definitions characterizes the concept of BI. The following table provides some definitions, their key focus, and their author(s).

Table 1: BI Definitions

Author(s) and Year	Definition of BI	Key Focus
Chen et al. (2012)	Comprehensive framework that leans on data collection, extraction, and analytical methods to improve decision-making, only becoming popular in the 1990s.	Framework for decision-making improvement.
Wanda & Stian (2015)	An "*umbrella term*" that includes various technologies and processes used to collect, store, and analyze data to enhance decision-making.	Technologies and processes for data management.
Babić & Zron (2024)	Set of tools and technologies used to collect, process, and analyze data to support decision-making.	Emphasis on tools for practical application.
Sari (2021)	BI tools are technologies that are crucial in helping organizations analyze historical data, understand current trends, and predict future outcomes.	Trend analysis and outcome predictive

Regardless of the author's specific approach, there is a consensus that BI plays a crucial role in enhancing decision-making within companies. In this context, Alzghoul et al. (2024, p.221) defines BI capability as "*a firm's ability to scan, absorb,* *transform, and analyze internal and external data to minimize uncertainty and detect business opportunities and actionable insights to support business decisions.*" Given its relevance and up-to-date perspective, this definition will serve as the guiding framework for my thesis.

1.1.1 BI Benefits in Decision-Making

To better understand the significance of employing BI, it is essential to explore its concrete advantages, mainly concerning decision-making. This subsection explores some benefits of BI, demonstrating its transformative impact across various industries through real-world case studies.

Juare (2019) points out that companies have recognized the importance of BI, citing key benefits such as improved marketing effectiveness, customer insights and loyalty, and fast problem detection. Tunowski (2015) supports the idea by also emphasizing that BI reduces the time spent on data collection and processing, enabling decisions to be based on reliable data rather than intuition.

An illustrative example of what was stated is Datafortune's analysis of Spotify (Datafortune, 2024), showcasing the application of BI in the music streaming industry. Spotify uses BI to create personalized playlists like Discover Weekly, which relies on users' listening data to generate its accurate recommendation system. These strategies boost user engagement and foster customer loyalty, helping Spotify maintain its strong position in the market.

Gaspar & Pereira Da Silva (2020) presents another example, showing how BI improved the management control of an economic group in the health sector. The results include increased transparency, quicker and easier access to information, and increased autonomy in using a unified data source, ultimately resulting in more informed and efficient decision-making.

Finally, Reginato & Nascimento (2007) focus on implementing BI in several areas across a manufacturer of energy systems, namely the Sales departments, where the adoption of BI gave users the independence to access critical sales information, allowing for a complete, real-time picture of global sales activities. That adoption sped up the analysis of specific results for each client and made it easier to understand whether negotiations were headed in a good direction before their conclusion.

These cases highlight how BI improves decision-making quality, speed, and accuracy. Whether optimizing customer engagement, managing operations, or analyzing costs, BI ensures decisions are informed by data rather than assumptions, driving better outcomes across industries.

Table 2: BI Benefits

Author(s)	**Key BI Benefits**
Juare (2019)	• Improved marketing effectiveness • Customer insights and loyalty
Tunowski (2015)	• Reduced time on data collection and processing
Datafortune (2024)	• Personalized user experiences • Enhanced user engagement
Gaspar & Pereira Da Silva (2020)	• Increased transparency • Faster access to information
Reginato & Nascimento (2007)	• Real-time access to sales data • Client-specific result analysis

1.2 Data Visualization

BI serves as an effective tool for DV, connecting raw data with actionable insights. Building on this, Vyas (2024, p.52) highlights that "*Data Visualization is more than charts and graphs, it is about presenting relevant information in a clear and engaging manner*".

DV is not just a representation of data but a powerful ally in decision-making. It simplifies complex information into clear and engaging visuals, making it easier to spot patterns and trends (Sharma et al., 2024).

Figure 1: Magic Quadrant

As illustrated in Figure 1, there are various BI tools for DV, each positioned based on its ability to execute and completeness of vision. (Richardson et al., 2021). According to Richardson et al. (2021), certain tools are better suited for data visualization. Among them, Microsoft Power BI, Domo, Tableau, Qlik, and SAP are some of the distinguished tools. The table below briefly introduces the tools and their strengths, leaving out Microsoft Power BI, which will be discussed in greater detail later.

Table 3: BI Tools

BI Tool	Brief Overview	Strengths
Domo	A cloud-based Analytics and Business Intelligence (ABI) platform that offers user-friendly data visualizations and dashboards. (Richardson et al., 2021)	Quick implementation using Application Programming Interface (API)-style connectors that adjust to changes in the schema. (Richardson et al., 2021)
Tableau	A visual-based tool that enables users to access, manipulate, analyze and report out information in their data. (Richardson et al., 2021)	Enhanced capabilities for preparing and processing data (Richardson et al., 2021)
Qlik	Comprehensive data analytics and integration platform that utilizes Artificial Intelligence (AI) and Machine Learning (ML). It enables organizations to collect, integrate, analyze, and respond to real-time data. (Qlik, n.d.; Richardson et al., 2021)	Augmented capabilities, the platform offers advanced AI-drive features, such as automatic visualizations. (Richardson et al., 2021)
SAP	A cloud-native analytics tool platform that combines BI,	Advanced augmented analytics, the platform differentiates itself with

	planning, and predictive analysis. (Richardson et al., 2021)	integrated planning, predictive analysis, and "what-if" scenario modeling. (Richardson et al., 2021)

Microsoft Power BI is a user-friendly selection of software devices that work together to help turn various data sources into clear, engaging, and interactive insights (Gowthami & Pavan Kumar, 2017; Microsoft, 2024). It has established itself as a leading tool for data visualization, positioned in the "*Leaders*" quadrant according to the Figure 1. This ranking reflects its exceptional performance making it one of the most robust and versatile platforms available (Richardson et al., 2021).

This powerful tool incorporates various components, illustrated in the figure below.

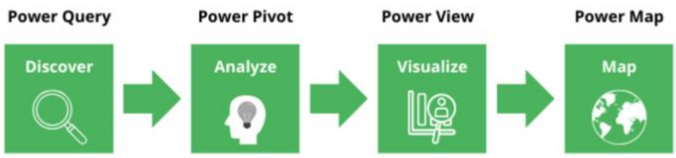


Figure 2: Power BI Components (Microsoft, 2024)

The different Power BI components are detailed below:

- **Power Query** is a self-service ETL (Extract, Transform, Load) tool that acts as an add-in for Excel. It enables users to import data from various sources, transform it into the desired format, and load it into Excel for further analysis (Gowthami & Pavan Kumar, 2017).

- **Power Pivot** is an in-memory data modeling component allowing rapid data aggregation, complex calculations, and highly compressed data storage. Users can create a data model within an Excel workbook, and data can be loaded into Power Pivot directly or via Power Query. (Gowthami & Pavan Kumar, 2017).
- **Power View** is an interactive data visualization tool featuring a drag-and-drop interface. It lets users quickly create dynamic and intuitive visualizations within their Excel workbooks (Gowthami & Pavan Kumar, 2017).
- **Power Map** is an Excel 3D data visualization tool that provides new ways to explore and analyze data. This tool allows users to uncover insights that may not be visible in traditional 2D charts and tables (Gowthami & Pavan Kumar, 2017).
- **Power Q&A** is a helpful tool that "lets you explore your data in your own words", meaning that it understands your questions and delivers accurate answers using natural language processing (Gowthami & Pavan Kumar, 2017).
- **Power BI Desktop** is a comprehensive analytics tool that offers intuitive report authoring. It features a drag-and-drop interface for creating flexible and interactive reports, allowing users to explore data patterns and gain insights through a unified view of connected visualizations (Gowthami & Pavan Kumar, 2017).

These components work together to make it easier to collect, process, and present data, empowering employees to make informed data-drive decisions.

The following table summarizes the aspects of Power BI that outperform traditional tools such as Excel and Tableau to clearly show why it was the main tool chosen for the dissertation.

Table 4: Power BI vs. Excel and Tableau: A Performance and Decision-Making Comparison

Aspect	Key Findings	Decision-Making
Time Efficiency	It demonstrates a faster task completion (12.5 minutes on average) compared to Tableau (17.0 minutes). (Tirupati et al., 2023)	It gives organizations a competitive advantage since it is possible to respond rapidly. (Tirupati et al., 2023)
Accuracy of Insights	It achieved higher accuracy (90%) than Tableau (85%). (Tirupati et al., 2023)	It enhances the precision of insights, resulting in a substantial decrease in errors. (Tirupati et al., 2023)
Adaptability to Real-Time Data	It scored 9.0 for adaptability compared to Tableau (7.5) and Excel (6.0). (Tirupati et al., 2023)	Real-time adaptability allows for faster modifications to effectively respond to emerging information. (Tirupati et al., 2023)
User Satisfaction	Compared to the other tools, it had the highest user satisfaction (4.7 out of 5) in terms of Ease of usage, Visualization Clarity, and Customization Flexibility. (Tirupati et al., 2023)	High user satisfaction promotes adoption, improving data-driven decision-making at every level. (Tirupati et al., 2023)

In summary, Power BI has a better score in each aspect tested, highlighting that it's *"an ideal solution for businesses seeking to improve their data analysis and decision-making processes."* (Tirupati et al., 2023, p.701)

These strengths are further supported by the capabilities of Data Analysis Expressions (DAX), which enhance Power BI's data manipulation and calculation processes. According to (Microsoft, 2023), DAX is a formula-based expression language utilized in Analysis Services, Power BI, and Power Pivot in Excel. DAX formulas consist of functions, operators, and values, enabling advanced calculations and queries on data across related tables and columns within tabular data models. They are used in measures, calculated columns, calculated tables, and row-level security (Microsoft, 2023). As measures play a key role in this dissertation, it is important to introduce their purpose and explore some commonly used DAX functions.

Measures are dynamic calculation formulas that adjust their results based on the given context (Microsoft, 2023), meaning that in Power BI, they operate on each row of data while also responding to filter selections.

The table below provides a summary of key DAX functions and their usage.

Table 5: DAX functions (Microsoft, 2023)

Function	Syntax	Description	Return Value
CALCULATE	CALCULATE (<expression>[, <filter1> [, <filter2> [, …]]])	Applies filters to an expression and recalculates the results based on those filters.	The result of the expression recalculated with the applied filters
TREATAS	TREATAS (table_expression, <column>[, <column>[,…]])	Converts a table into a set of values that can be used as a filter.	A table containing all the rows of the specified

			columns in table_expression.
MAX	MAX (<column>) / MAX (<expression1>, <expression2>)	Returns the largest value in a column or between two expressions.	The largest value found in the column or between the compared expressions.
SELECTEDVALUE	SELECTEDVALUE (<columnName>[, <alternateResult>])	Returns the selected value in a specific context or an alternate value if there is no unique selection.	The unique selected value or an alternate value when multiple selections exist.
SUMMARIZE	SUMMARIZE (<table>, <groupBy_columnName> [,<groupBy_columnName>]… [, <name>, <expression>])	Groups a table by one or more columns and allows the addition of aggregated calculations.	A table containing the grouped columns and the specified aggregated columns.

As previously mentioned, clear and effective visualization is crucial for data-driven decision-making. Dashboards are particularly helpful in this regard. They serve as a single-page interface, often called a canvas, that tells a cohesive story through visualizations (Microsoft, 2024).

"*Data Visualization is a mix of science and art*" (Schwabish, 2021, p.17). Users do not always desire to view data as absolute values for direct comparisons. Sometimes, it is important to be closer to the art side by creating visuals that are more engaging and exciting to the reader, even if they do not allow the most accurate comparisons (Schwabish, 2021).

To achieve that, Sisense (2023) defined some rules to be followed when creating a dashboard.

- "*Your dashboard should provide the relevant information in about 5 seconds*" (Sisense, 2023),
- It should be displayed logically, with the most important and substantial information at the top, accompanied by important information that offers a clearer insight into the overview mentioned earlier, and general and background information at the bottom. In summary, the dashboard should tell a story (Sisense, 2023),
- It should not contain more than nine visualizations; even though some designers must add as many details as possible, the human brain can only understand around seven images simultaneously (Sisense, 2023),
- Selecting the right data visualization. Choosing the correct type of data visualization is crucial for effectively communicating information and insights from data (Sisense, 2023).

1.3 Data Normalization

To enhance the user experience, data must be well-structured and organized, a goal achieved through DN. DN is a crucial process in database management and design. It helps organize data to avoid duplication and inconsistencies, ensuring data accuracy and efficiency (Baxendale & Codd, 1970).

DN acts as a shield by breaking down large tables into smaller ones and establishing relationships, reducing the chances of anomalies. This makes the database more flexible and scalable and instills a sense of security in the data management process (Baxendale & Codd, 1970).

Normalization is usually divided into several normal forms, such as the first normal form (1NF), second normal form (2NF), and third normal form (3NF), each with specific rules for eliminating redundancy and maintaining data integrity(Chris, 2023).

It is important to clarify several key concepts defined in the table below.

Table 6: Data Normalization Key Concepts

Key Concept	Definition
Primary Key	Column that uniquely distinguishes each row of data within a table (Chris, 2023).
Foreign Key	Field related to the primary key in another table (Chris, 2023).
Composite Key	Similar to a primary key but made up of multiple columns (Chris, 2023).
Functional Dependency	A functional dependency exists when one attribute's value determines another's value (Chris, 2023).

Partial Functional Dependency	A functional dependency is considered partial if the determinant consists only of a part of the attributes that form the primary key (Chris, 2023).
Transactive Functional Dependency	This occurs when three distinct attributes, A, B, and C, are related in a specific way. If A → B and B → C, C is considered transitively dependent on A through B (Chris, 2023).

Starting on the 1NF, there are some rules to follow: a column in a table must contain only atomic values, cannot hold multiple values, and must have a primary key for identification. A relation is considered in first normal form only if every attribute contains a single value (Gokila & Balasubramani, 2019).

Table 7: Before 1NF

Student_ID	Student_Name	Course_Name	Course_Teacher	Grade
1	John	Math, Science, English	Anna, Paul, Rachel	16, 19, 12
2	Mary	Math, English, History	Anna, Rachel, Robert	15, 15, 15
3	Peter	English, History	Rachel, Robert	20, 20

Table 8: After 1NF

Student_ID	Student_Name	Course_Name	Course_Teacher	Grade
1	John	Math	Anna	16
1	John	Science	Paul	19

1	John	English	Rachel	12

Building upon this, a table progresses to 2NF if it meets some conditions, such as needing to be in 1NF, and non-key attributes depend on a proper subset of any candidate key of the table. It is known as non-prime attributes, not part of any candidate key (Gokila & Balasubramani, 2019).

Table 9: 2NF Tables

Student_ID	Student_Name
1	John
2	Mary
3	Peter

Course_ID	Course_Name	Course_Teacher
1	Math	Anna
2	Engilsh	Rachel
3	History	Robert

To achieve the final goal, the 3NF, it is necessary to eliminate transitive partial dependencies (Gokila & Balasubramani, 2019).

Table 10: 3NF Table

Student_ID	Course_ID	Grade
1	1	15
2	2	15
3	2	20

In conclusion, DN is a critical process for organizations looking to effectively manage significant volumes of data (Morris, 2024). Implementing DN improves data integrity and facilitates better data analysis, cost reduction, and less chance of errors (Morris, 2024).

1.4 Integrating BI, DV, and DN

Research on BI, DV, and DN has been extensive (Morris, 2024; Chen et al.,2012; Vyas A., 2024), with numerous strategies developed to enhance results across several areas. However, while these areas have been studied individually, there remains a significant gap in research that integrates all three components into a cohesive and practical framework.

In today's data-driven world, organizations are increasingly overloaded with vast amounts of data (Attanapola & Iyer, 2024). This exponential growth in data volume presents a significant challenge (CelerData, 2024). Without proper management and organization, raw data often remains inconsistent or underutilized.

Despite the promise of BI tools in addressing this issue, adoption remains relatively low among employees. According to Gartner, as mentioned in Attanapola & Iyer (2024), although 87% of surveyed organizations have reported an increase in the number of employees using ABI tools, these tools are still used by only 29% of employees on average. This discrepancy suggests that many organizations face barriers in effectively integrating BI into daily decision-making processes, reinforcing the need for solutions that improve data usability.

Existing studies highlight the impact of BI tools in accelerating decision-making processes (Alzghoul et al., 2024) and focus on identifying the most effective BI tools for creating more effective dashboards (Gowthami & Pavan Kumar, 2017). Nevertheless, despite the wealth of research, the fundamental role of data quality, ensuring through proper DN, as a prerequisite of effective BI and DV remains largely underexplored.

This study addresses this gap by investigating how DN enhances the quality and reliability of DV, particularly in Power BI, by leveraging BI capabilities for more effective and data-driven decision-making. By integrating DN as a crucial preprocessing step, this research demonstrates its impact on improving data-driven decision-making, particularly in product sales analysis.

Chapter 3

1 Research Methodology

1.1 Project Goals

This dissertation aims to improve the company's vehicle launch decision-making process through data normalization and the development of a Power BI dashboard. The company was facing challenges in price positioning across various African countries, including Angola, Cape Verde, Kenya, Mozambique, Senegal, and Tunisia. The price positioning process was particularly time-consuming, as it relied entirely on manual data entry and analysis.

As this dissertation is presented in the format of an Internship Report, it does not address a specific research topic based on literature gaps. Instead, it seeks to meet the following core objectives:

1. Normalize the data of the Key Selling Points to ensure data consistency
2. Create a Power BI dashboard that enables more informed and faster decisions regarding vehicle launches
3. Improve the Data Visualization primarily based on the managerial inputs provided

The internship took place at SCA, whose operations and context will be discussed in following section.

1.2 Research Contextualization

GSC, founded in 1946 by Salvador Fernandes Caetano, is the largest and most influential Portuguese company in the automative sector. Initially established as a small workshop specializing in bus body repairs in Vila Nova de Gaia, Portugal, the company has since undergone significant expansion, diversifying its operations and establishing a presence in multiple international markets.

Today, GSC operates across seven key business areas, including:

- Bus manufacturing,
- Automotive distribution and retail,
- Mobility solutions,
- Industrial equipment and workshops,
- Production and automotive assembly,
- Various support services, and
- Aeronautics.

The group comprises over one hundred companies, representing more than forty brands, and maintains operations in over forty-four countries worldwide.

Since this internship was conducted at SCA , the focus will be on its operations within the group. SCA specializes in African markets, providing centralized and robust support across various services. The company ensures the implementation of best practices and drives process optimization and synergies in key areas such as logistics, finance, training, and IT. Additionally, the company plays a crucial role in bridging local operations with original equipment manufacturers (OEM), facilitating efficient interaction and collaboration.

While the core services are based in Portugal, the company actively assesses and support local operations in African markets through on-site visits, ensuring alignment with strategic objectives and market needs.

1.3 Method

To achieve the objectives outlined in subsection 1.1, the chosen methodology was Action Research, originally conceptualized by Lewin (1946). This methodology was chosen due to its iterative nature, which allows for both the analysis of the results and the continuous identification of new improvements.

Action Research is defined as a comparative study of different forms of social action and their effects, leading to actionable insights (Lewin, 1946). Another definition of Action Research is provided by Avison et al. (1999, p.94), who describe it as "*a participatory, democratic process concerned with developing practical knowing in the pursuit of worthwhile human purposes*."

According to Avison et al. (1999), there are four types of Action Research:

- Change and reflection-oriented Action Research;
- Action Science, seeking to resolve discrepancies between espoused and applied theories;
- Participatory Action Research, which prioritizes participation collaboration;
- Action Learning for programmed instruction and experiential learning.

For this dissertation will be used Action Learning, following a cycliclar process of four steps (Dickens & Watkins, 1999):

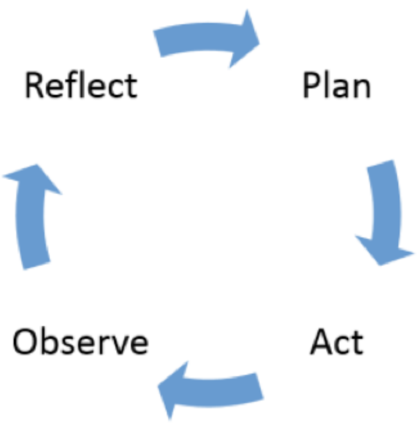


Figure 3: Action Research Cycle

The work was structured around two primary iterative cycles. The first cycle focused on Data Normalization and aimed to ensure that each table adhered to the principles of the 3NF. The second, centered on Dashboard Development, sought to identify and present the key information that product managers need to carry out their tasks more effectively.

At the beginning of each cycle, the respective objectives and expected outcomes were clearly defined. Preliminary templates and mockups were then developed to visually represent the intended data structure and dashboard components. These initial drafts served as instruments for discussion and validation, guiding the subsequent development stages.

During the first cycle, meetings were held with product managers to verify whether the proposed data structure was understandable and aligned with their analytical expectations. In parallel, additional meetings were conducted with the IT team to ensure a shared approach to data structuring and transformation across systems.

The second cycle, centered on Dashboard Development, required a deeper engagement with product managers to understand their informational needs better. This process began with observing their day-to-day activities regarding price positioning, focusing on the types of analyses they performed. Initial dashboard prototypes were developed and presented in targeted feedback sessions based on these insights. As the development progressed, these sessions gradually evolved into hands-on trials. Product managers were invited to interact directly with the dashboard and provide feedback based on their user experience.

In parallel, feedback from the project supervisor also played a key role throughout the process.

Chapter 4

1 Database Normalization

This chapter presents the practical implementation of the first phase of the project. While the previous chapters provided the contextual and theoretical foundation for the challenges identified, the focus is on the concrete steps taken to address one of the project's critical issues.

The two subchapters that follow outline the fundamental structure of this initial phase. Subchapter 1.1 defines the initial problem, analyzing the limitations of the data structure previously used by product managers. Subchapter 1.2 then details the normalization process, including steps taken using Power BI tools to consolidate and optimize the datasets.

1.1 Problem definition

Product managers manually selected all relevant competitors for each newly launched vehicle model, analyzed their prices and sales figures for the current and previous years, and updated tables containing detailed model information, such as equipment, engine specifications, prices, and dates. These tables required constant updates whenever new data became available, which was time-consuming for product managers.

Since each product manager manually entered all data, the process was prone to errors, including incorrect vehicle segmentation, model pricing inaccuracies, and inconsistencies in sales data. Additionally, as all available information was stored in a single table per country, these tables needed to be refined to ensure proper functionality in Power BI. The lack of normalization made them susceptible to human errors, affecting data accuracy and analysis.

In the following example, some of the minor errors present in the tables can be observed.

	Brand	TOYOTA	
	Model	HILUX	
	Version	DC BASIC 2.4D 4X4 MT	
	Sales FY 2023		
30	Sales YTD SET 2024		
	Equipment level (base, medium, high)		
Powertrain & Transmission	**Engine size**	2 400	0,007
	Power HP	110	0,011
	Fuel	Diesel	0,020
	Transmission	MT	0,020
	Traction	4x4	0,050
	Ground Clearance	310	0,052
	Limited-slip differential		
Exterior	**LED lights**	-	0,000
	Daytime running lights	-	0,000
	Fog lights	•	0,010
	Wheels	Steel	0,026
	Rim	17	0,024
	Roof bars	-	0,000
	Dimension (Length)	5 325	0,030
	Other		

	Brand	Toyota	
	Model	Hilux	
	Version	Country 2.4L SC MT 4X2	
	Sales 2023		
	Sales YTD August 2024		
	Equipment level (base, medium, high)	Base	
Powertrain & Transmission	**Engine size**	2 393	0,007
	Power HP	101	0,010
	Fuel	Diesel	0,020
	Transmission	MT	0,020
	Traction	4X2	0,025
	Ground Clearance	294	0,049
	Limited-slip differential		
Exterior	LED lights	-	0,000
	Daytime running lights	-	0,000
	Fog lights	-	0,000
	Wheels	Steel	0,026
	Rim	17	0,024
	Roof bars		0,000
	Dimension (Length)	5 325	0,030
	Other		

Figure 4: Previously Used Tables

As shown in Figure 4, some of the errors in the tables used by the product managers become evident. Brands and model names were recorded inconsistently, leading to variations in spelling and formatting. For example, it is possible to identify discrepamcies in the brand name between the two tables, where one lists "TOYOTA" and the other "Toyota". Similarly, inconsistencies apperar in the model name. Additionally, the inexistence of a standardized approach to writing version names resulted in further discrepancies.

These inconsistencies not only complicated data analysis but also increased the risk of misinterpretation and duplication, highlighting the need for a more structured and uniform data entry process.

After establishing the core problem, the implementation of the solution began with the methodology's first phase. This initial stage focused on one of the project's most critical pillars: the organization and structuring of the available data.

In this context, the following section presents the detailed data normalization process, outlining the structure adopted and the technical steps taken to convert disorganized and redundant data into structured and optimized datasets ready for use in Power BI.

1.2 Data Normalization

To maintain the simplicity of the product managers' workflow without significantly altering their existing processes, multiple tables were created to facilitate the normalization of the same data. This structured approach aimed to improve data consistency while ensuring a seamless transition for the team. The created tables were derived from the information contained in the table shown in Figure 4, with the data systematically distributed across the tables presented next. The original table, which consisted of forty rows and was continuously expanding in columns, was divided into sixteen separate tables.

- *Brand_Country_Category*: Divided into three tables containing information about vehicle brands, countries, and categories.
- *Models*: Contains details about each vehicle model, including brand, dimensions, and category.
- *Equipment*: Divided into six tables:
  - Four for engine specifications (fuel type, engine type, transmission, and traction).
  - One for equipment types.
  - One for equipment names.
  - A relational table linking equipment type, vehicle category, score, and available options.
- *Market – Country*: Contains data on vehicles sold or to be sold in each country, with individual files per country. Includes brand, model, active/inactive

status, version, name, equipment, and pricing data stored by quadrants and year for historical analysis.

All Excel files are interconnected to ensure consistency in naming brands, models, equipment, and engine specifications across all datasets. This linkage prevents discrepancies, standardizes data entry, and facilitates a more accurate and efficient analysis.

The following figures show some of the tables.

Category_Eq_ID	Category_Name
1	Wheel
2	Head Lamps
3	Daytime Running Lamps
4	Front Fog Lamps
5	Rim
6	Power Trunk
7	Sunroof

Brand_ID	Brand_Name
1	Acura
2	Alfa Romeo
3	Ashok Leyland
4	Aston Martin
5	Audi
6	BAIC
7	BMW

Equipment_ID	Equipment_Name	Equipment_Type_ID	Category_Eq_ID	Score
1	N/A	1	1	0,0000
2	13	1	1	0,0414
3	14	1	1	0,0445
4	15	1	1	0,0477
5	16	1	1	0,0509
6	17	1	1	0,0541
7	18	1	1	0,0573

Figure 5: Examples of created tables

As mentioned earlier, the tables for each market were designed to be user-friendly, ensuring that product managers could continue their work with minimal disruption. Instead of using unique IDs, data validation procedures, such as dropdown menus, were implemented in each column to standardize data entry.

Brand	Model	Status	Version Não Preencher	Name	Engine	Fuel	Engine Capacity	Engine Power (HP)
Suzuki	Alto	Active	1.0L Petrol MT 4x2 DX 5S	DX	Combustion	Petrol	998	67
Suzuki	S-Presso	Active	1.0L Petrol MT 4x2 GL 5S	GL	Combustion	Petrol	998	68
Suzuki	S-Presso	Active	1.0L Petrol AT 4x2 AGS 5S	AGS	Combustion	Petrol	998	68
Suzuki	Celerio	Active	1.0L Petrol AT 4x2 GL 5S	GL	Combustion	Petrol	998	68
Hyundai	Grand i10	Active	1.0L Petrol AT 4x2 5S		Combustion	Petrol	1000	67
Hyundai	Grand i10	Active	1.0L Petrol MT 4x2 5S		Combustion	Petrol	1000	67
Kia	Picanto	Active	1.2L Petrol AT 4x2 Classic 5S	Classic	Combustion	Petrol	1200	
Fiat	Panda	Active	1.2L Petrol MT 4x2 Easy 5S	Easy	Combustion	Petrol	1200	69
Renault	Kwid	Active	1.0L Petrol MT 4x2 Zen 5S	Zen	Combustion	Petrol	1000	66

Figure 6: Market - Country Table

Since the market table does not adhere to normalization rules, the following explanation outlines the steps required to restructure it using proper normalization principles.

As mentioned earlier, Power Query was used within Power BI to facilitate the normalization process by merging queries and restructuring the market data into three distinct tables.

To ensure a standardized structure, the *Country_ID* was added to each *Market - Country* table, linking each dataset to its respective country. After this step, all individual market tables were consolidated into a single unified table, improving data consistency and simplifying analysis**.**

The first table, *Car_Principal*, is a fundamental component of the database structure. It contains the *Car_ID*, which serves as the primary key, uniquely identifying each vehicle. Additionally, it includes *Model_ID* and *Country_ID*, which function as foreign keys, establishing relationships with the respective *Models* and *Market - Country* tables.

Car_ID	Model_ID	Country_ID	Status	Version Não Preencher	Name	Engine Capacity	Engine Power (HP)
5	454	1	Active	1.0L Petrol AT 4x2 5S		1000	67
386	454	2	Active	1.2L Petrol AT 4x2 Hatchback 5S	Hatchback	1197	82
387	454	2	Active	1.2L Petrol AT 4x2 Sedan 5S	Sedan	1197	82
7	558	1	Active	1.2L Petrol AT 4x2 Classic 5S	Classic	1200	
378	558	2	Active	1.0L Petrol AT 4x2 EX 5S	EX	998	67
174	857	3	Active	1.0L Petrol AT 4x2 M2 5S	M2	999	65
376	857	2	Active	1.0L Petrol AT 4x2 E3 5S	E3	999	68
377	857	2	Active	1.0L Petrol AT 4x2 Climber 5S	Climber	999	68
10	737	1	Active	1.2L Petrol AT 4x2 5S		1193	78
16	441	1	Active	1.4L Petrol AT 4x2 5S		1400	100
17	280	1	Active	1.4L Petrol AT 4x2 Lounge 5S	Lounge	1400	100
18	712	1	Active	1.5L Petrol AT 4x4 5S		1500	120
21	559	1	Active	1.4L Petrol AT 4x2 Executive 5S	Executive	1368	107
22	556	1	Active	2.0L Petrol AT 4x2 GT Line 5S	GT Line	2000	162

Figure 7: Car_Principal Table

The second table, *Car_Equipment*, maps the *Car_ID* to the equipment associated with each vehicle. The *Car_ID* serves as the primary key, while the *Equipment_ID* acts as a foreign key, linking each vehicle to its corresponding equipment. To obtain the *Equipment_ID*, a merge was performed using both the equipment name and its corresponding option.

Car_ID	Equipment_ID
856	161
855	161
854	161
853	161
852	161
851	161
850	161
849	161
848	161
847	161
846	161
845	161

Figure 8: Car_Equipment Table

The third table, *Car_Price*, stores the *Car_ID* alongside pricing details, including the price and date. *Car_ID* is also the primary key, but this table does not contain any foreign keys. From the date column, two additional columns, Quarter and Year, were derived to be used in the dashboard in the future, allowing for better time-based analysis.

Car_ID	Quarter	Year	Price	Date
860	2	2024	0	30/06/2024
859	2	2024	0	30/06/2024
858	2	2024	0	30/06/2024
857	2	2024	0	30/06/2024
856	2	2024	0	30/06/2024
855	2	2024	0	30/06/2024
854	2	2024	0	30/06/2024
853	2	2024	0	30/06/2024
852	2	2024	0	30/06/2024
851	2	2024	0	30/06/2024
850	2	2024	0	30/06/2024
849	2	2024	0	30/06/2024
848	2	2024	0	30/06/2024
847	2	2024	0	30/06/2024
846	2	2024	0	30/06/2024

Figure 9: Car_Price Table

After completing the entire data normalization process, the dataset was fully structured and optimized for further processing in Power BI. By ensuring proper normalization, eliminating redundancies, and establishing clear relationships between tables, the data becomes ready for efficient analysis and visualization.

With the dataset prepared, the next chapter will focus on the development of the dashboard, representing the second cycle of the methodology, detailing its structure, key visualizations, and how it enables better decision-making through improved data insights.

Chapter 5

1 Power BI Dashboard

Previously, all analyses relied solely on the tables presented in Figure 4, requiring an entirely manual process that was time-consuming and prone to human errors. Since no dashboard had been previously developed for this purpose, one was designed to support the product team in price positioning.

As discussed in the previous chapter, sixteen tables were initially added to Power BI to facilitate the dashboard's development. In addition, six more tables created by the company were incorporated. These additional tables contain detailed sales information for each model, each representing data from a specific country.

Initially designed solely for price positioning, the dashboard gradually evolved to incorporate analyses over time. As a result, it was structured into six distinct pages, three of which are dedicated to price positioning : General, Specific Vehicle, and Equipment Benchmark. Additionally, one page focuses on ranking, another on model releases, and the last one provides general information. Throughout this chapter, each group of pages will be explained in detail.

1.1 General Page

The first page, General, enables a wide range of analysis through interactive filters, allowing users to select criteria ranging from brand to price. These filters (Figure 10, section A) offer analytical flexibility in refining data, leaving it up to the manager to determine the type of analysis they wish to conduct.

To ensure consistency in price representation, a currency conversion API was integrated. This API updates daily and converts the local price into the currency chosen by the user (Figure 10, section C) standardizing the analysis without requiring manual conversions.

The results generated based on the selected filters are displayed in a structured table format, as illustrated in the following figure, section B.

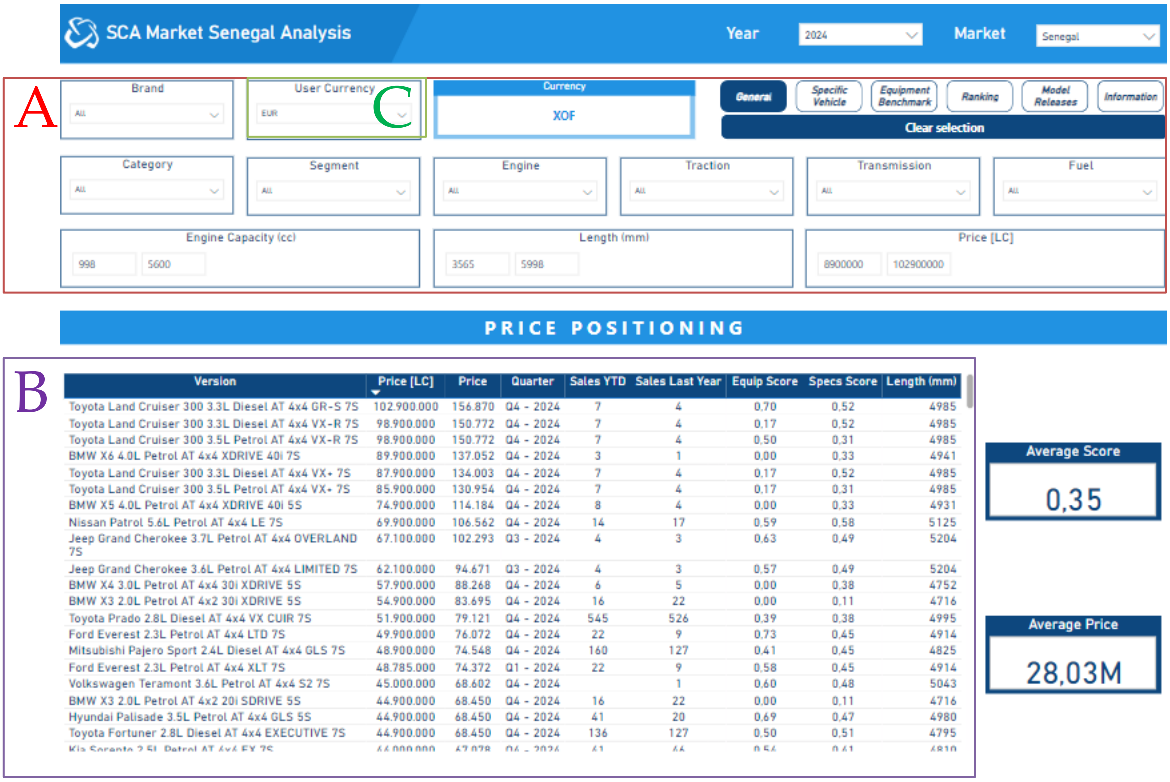


Figure 10: General Page – Table

The page also provides an additional analysis that is independent of the previously selected filters, allowing managers to examine data based on brand, segment, or sub-segment. Conditional formatting was applied to highlight the selected option to enhance readability and facilitate quick interpretation.

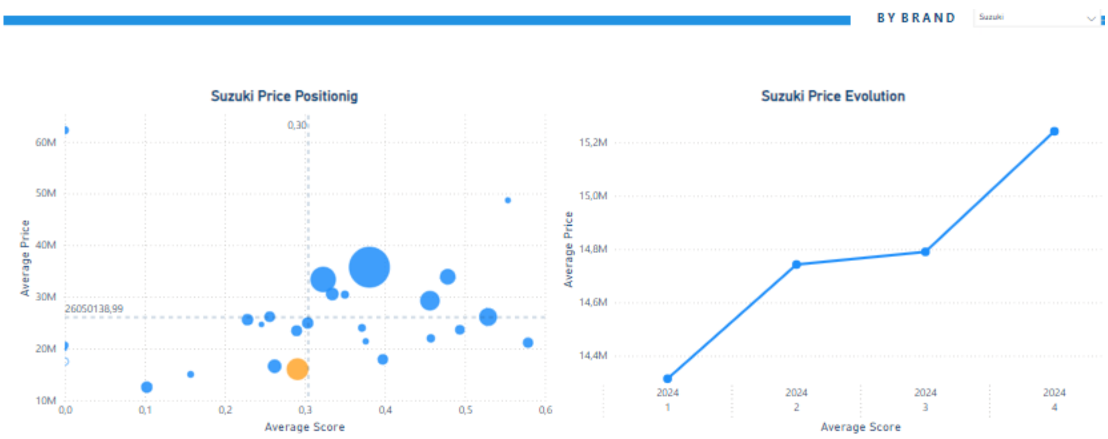


Figure 11: General Page - Price Positioning and Evolution Analysis by Brand

1.2 Specific Vehicle Page

The second and third pages of the dashboard are related, as they rely on a set of three interconnect filters that relate to the vehicle version.

This page focuses on analyzing a specific vehicle version and its competitors. Users can define a version by selecting brand, model, and version (Figure 12, section A). Once selected, a set of DAX functions were applied to generate results. This page displays a table (Figure 12, section C) listing the competing vehicles identified based on specific characteristics. The selected vehicle is also included in the table, and it is highlighted for easier identification.

The following table provides an explanation of the DAX functions used.

Table 11: Initial DAX Functions

DAX Measure	Description	Key Variables and Logic
Same_Carachteristics_Vehicle	Identifies vehicles with similar characteristics based on	Uses SELECTEDVALUE and MAX function

	various attributes such as engine, etc.	to compare characteristics.
Highlighted_Vehicle	Calculates the total score of the highlighted vehicle.	Uses CALCULATE and TREATAS functions to retrieve and sum the relevant scores.
Selected_Car	Applies conditional formatting to highlight the selected vehicle.	Checks if the sum of total and engine scores matches the highlighted vehicle, then assigns a specific highlight color.

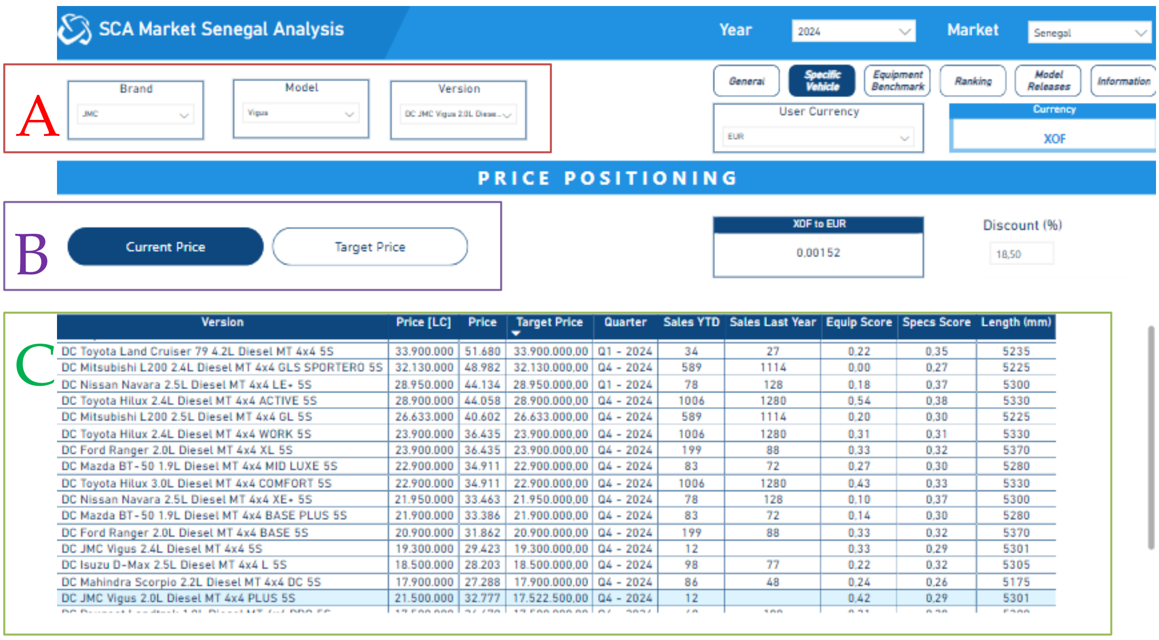


Figure 12: Specific Vehicle Page - Table

The table shown in Figure 12, section B, provides two scenarios: Current Price and Target Price. When selecting the Target Price option, the manager can apply a discount to the vehicle for potential scenario analysis. Based on the target price, the vehicle dynamically shifts within the table, reflecting its new position relative to competitors.

A scatter plot, shown in Figure 13, was created to assist product managers in analyzing the competition. It displays all competing vehicles while applying conditional formatting to highlight the selected version. Additionally, two more charts were later incorporated: one illustrating the price evolution and another showcasing their sales performance over time.

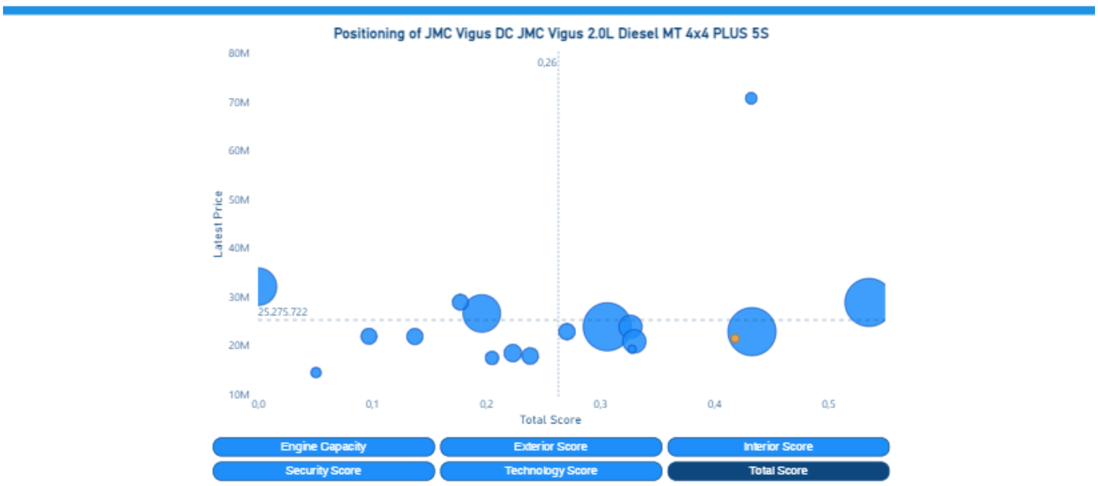
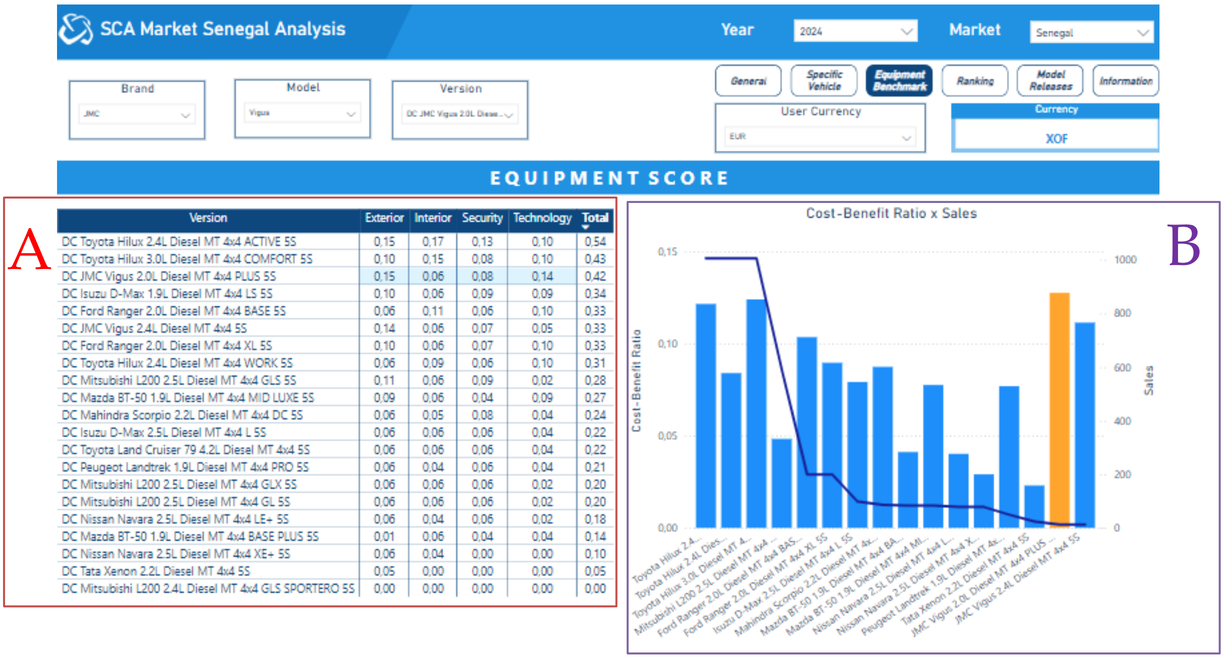


Figure 13: Specific Vehicle Page - Scatter Plot

1.3 Equipment Benchmark Page

This page evaluates the equipment offering of the selected vehicle and its competitors. It includes a table (Figure 14, section A) that presents the score of each version for different equipment categories, such as safety, interior, exterior, and technology.

A complementary bar chart (Figure 14, section B) was also added to illustrate the ratio between the total equipment score and the price of each version. This visualization enables product managers to identify the most competitive versions in terms of features and cost-effectiveness.



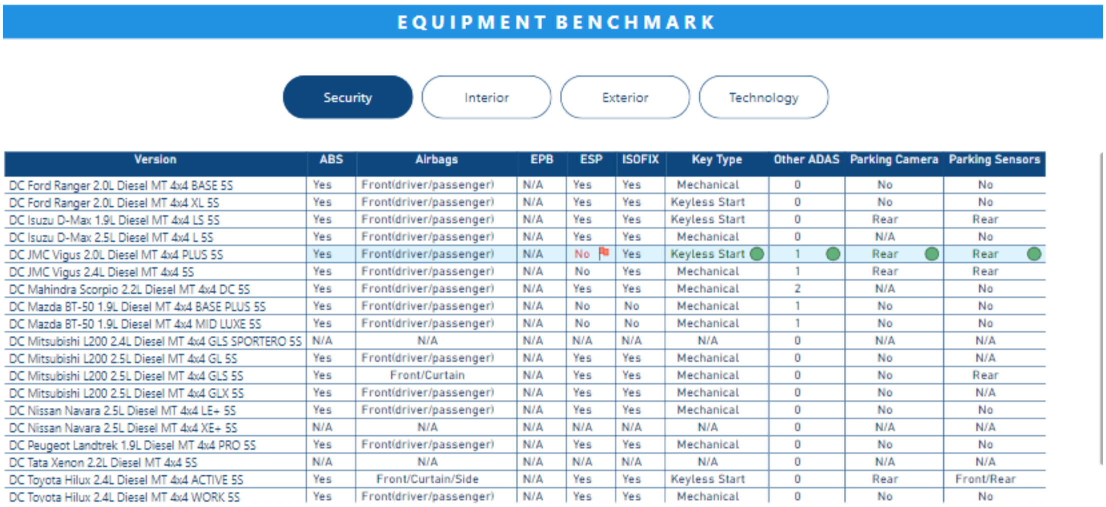
Version	Exterior	Interior	Security	Technology	Total
DC Toyota Hilux 2.4L Diesel MT 4x4 ACTIVE 5S	0,15	0,17	0,13	0,10	0,54
DC Toyota Hilux 3.0L Diesel MT 4x4 COMFORT 5S	0,10	0,15	0,08	0,10	0,43
DC JMC Vigus 2.0L Diesel MT 4x4 PLUS 5S	0,15	0,06	0,08	0,14	0,42
DC Isuzu D-Max 1.9L Diesel MT 4x4 LS 5S	0,10	0,06	0,09	0,09	0,34
DC Ford Ranger 2.0L Diesel MT 4x4 BASE 5S	0,06	0,11	0,06	0,10	0,33
DC JMC Vigus 2.4L Diesel MT 4x4 5S	0,14	0,06	0,07	0,05	0,33
DC Ford Ranger 2.0L Diesel MT 4x4 XL 5S	0,10	0,06	0,07	0,10	0,33
DC Toyota Hilux 2.4L Diesel MT 4x4 WORK 5S	0,06	0,09	0,06	0,10	0,31
DC Mitsubishi L200 2.5L Diesel MT 4x4 GLS 5S	0,11	0,06	0,09	0,02	0,28
DC Mazda BT-50 1.9L Diesel MT 4x4 MID LUXE 5S	0,09	0,06	0,04	0,09	0,27
DC Mahindra Scorpio 2.2L Diesel MT 4x4 DC 5S	0,06	0,05	0,08	0,04	0,24
DC Isuzu D-Max 2.5L Diesel MT 4x4 L 5S	0,06	0,06	0,06	0,04	0,22
DC Toyota Land Cruiser 79 4.2L Diesel MT 4x4 5S	0,06	0,06	0,06	0,04	0,22
DC Peugeot Landtrek 1.9L Diesel MT 4x4 PRO 5S	0,06	0,04	0,06	0,04	0,21
DC Mitsubishi L200 2.5L Diesel MT 4x4 GLX 5S	0,06	0,06	0,06	0,02	0,20
DC Mitsubishi L200 2.5L Diesel MT 4x4 GL 5S	0,06	0,06	0,06	0,02	0,20
DC Nissan Navara 2.5L Diesel MT 4x4 LE+ 5S	0,06	0,04	0,06	0,02	0,18
DC Mazda BT-50 1.9L Diesel MT 4x4 BASE PLUS 5S	0,01	0,06	0,04	0,04	0,14
DC Nissan Navara 2.5L Diesel MT 4x4 XE+ 5S	0,06	0,04	0,00	0,00	0,10
DC Tata Xenon 2.2L Diesel MT 4x4 5S	0,05	0,00	0,00	0,00	0,05
DC Mitsubishi L200 2.4L Diesel MT 4x4 GLS SPORTERO 5S	0,00	0,00	0,00	0,00	0,00

Figure 14: Equipment Benchmark Page - Table and Bar Graph

Additionally, it features another table, illustrated in Figure 15, that shows the version alongside the specific option available for each equipment feature. This table is interactive, incorporating buttons corresponding to the four equipment categories.

DAX functions were applied to determine the mode for each equipment feature across competitors, assisting product managers in evaluating whether a vehicle is well-positioned in terms of equipment within its segment.

Additionally, conditional formatting was used to implement both a red flag indicator, highlighting cases where the selected vehicle lacks a common feature, and a green flag indicator, signaling instances where the vehicle offers superior equipment compared to most competitors.


EQUIPMENT BENCHMARK

Security | Interior | Exterior | Technology

Version	ABS	Airbags	EPB	ESP	ISOFIX	Key Type	Other ADAS	Parking Camera	Parking Sensors
DC Ford Ranger 2.0L Diesel MT 4x4 BASE 5S	Yes	Front(driver/passenger)	N/A	Yes	Yes	Mechanical	0	No	No
DC Ford Ranger 2.0L Diesel MT 4x4 XL 5S	Yes	Front(driver/passenger)	N/A	Yes	Yes	Keyless Start	0	No	No
DC Isuzu D-Max 1.9L Diesel MT 4x4 LS 5S	Yes	Front(driver/passenger)	N/A	Yes	Yes	Keyless Start	0	Rear	Rear
DC Isuzu D-Max 2.5L Diesel MT 4x4 L 5S	Yes	Front(driver/passenger)	N/A	Yes	Yes	Mechanical	0	N/A	No
DC JMC Vigus 2.0L Diesel MT 4x4 PLUS 5S	Yes	Front(driver/passenger)	N/A	No	Yes	Keyless Start	1	Rear	Rear
DC JMC Vigus 2.4L Diesel MT 4x4 5S	Yes	Front(driver/passenger)	N/A	No	Yes	Mechanical	1	Rear	Rear
DC Mahindra Scorpio 2.2L Diesel MT 4x4 DC 5S	Yes	Front(driver/passenger)	N/A	Yes	Yes	Mechanical	2	N/A	No
DC Mazda BT-50 1.9L Diesel MT 4x4 BASE PLUS 5S	Yes	Front(driver/passenger)	N/A	No	No	Mechanical	1	No	No
DC Mazda BT-50 1.9L Diesel MT 4x4 MID LUXE 5S	Yes	Front(driver/passenger)	N/A	No	No	Mechanical	1	No	No
DC Mitsubishi L200 2.4L Diesel MT 4x4 GLS SPORTERO 5S	N/A	N/A	N/A	N/A	N/A	N/A	0	N/A	N/A
DC Mitsubishi L200 2.5L Diesel MT 4x4 GL 5S	Yes	Front(driver/passenger)	N/A	Yes	Yes	Mechanical	0	No	N/A
DC Mitsubishi L200 2.5L Diesel MT 4x4 GLS 5S	Yes	Front/Curtain	N/A	Yes	Yes	Mechanical	0	No	Rear
DC Mitsubishi L200 2.5L Diesel MT 4x4 GLX 5S	Yes	Front(driver/passenger)	N/A	Yes	Yes	Mechanical	0	No	N/A
DC Nissan Navara 2.5L Diesel MT 4x4 LE+ 5S	Yes	Front(driver/passenger)	N/A	Yes	Yes	Mechanical	0	No	No
DC Nissan Navara 2.5L Diesel MT 4x4 XE+ 5S	N/A	N/A	N/A	N/A	N/A	N/A	0	N/A	N/A
DC Peugeot Landtrek 1.9L Diesel MT 4x4 PRO 5S	Yes	Front(driver/passenger)	N/A	Yes	Yes	Mechanical	0	No	No
DC Tata Xenon 2.2L Diesel MT 4x4 5S	N/A	N/A	N/A	N/A	N/A	N/A	0	N/A	N/A
DC Toyota Hilux 2.4L Diesel MT 4x4 ACTIVE 5S	Yes	Front/Curtain/Side	N/A	Yes	Yes	Keyless Start	0	Rear	Front/Rear
DC Toyota Hilux 2.4L Diesel MT 4x4 WORK 5S	Yes	Front(driver/passenger)	N/A	Yes	Yes	Mechanical	0	No	No

Figure 15: Equipment Benchmark Page - Red Flag Table

Table 12: Additional DAX Functions for Competitive Equipment Analysis

DAX Measure	Description	Key Variables and Logic
Mode_Equip	Determines the most frequently occurring option (mode) for each equipment feature within a segment, country, and category.	Uses ADDCOLUMNS and SUMMARIZE to create a dataset. Calculates frequency using COUNT while excluding 'N/A' values.
Is_Red_Flag	Identifies whether a selected vehicle lacks a commonly available feature (red flag) or exceeds the standard	Uses CALCULATE and TREATAS to determine the minimum red flag value per category and compares it with the selected vehicle's

	offering (green flag) using conditional logic.	equipment ID to flag deficiencies or advantages.

The following pages were created at a later stage to provide each Product Manager with insights into sales performance and new model launches across different countries and a broader market perspective.

1.4 Ranking page

The fourth page, Ranking, offers an overview of each model's ranking within its segment, based on sales performance in a selected year. Conditional formatting was applied to enhance clarity:

- If the vehicle ranks in the TOP 3 and belongs to the company, it is highlighted in green
- If the vehicle ranks in the TOP 10 and belongs to the company, it is highlighted in blue

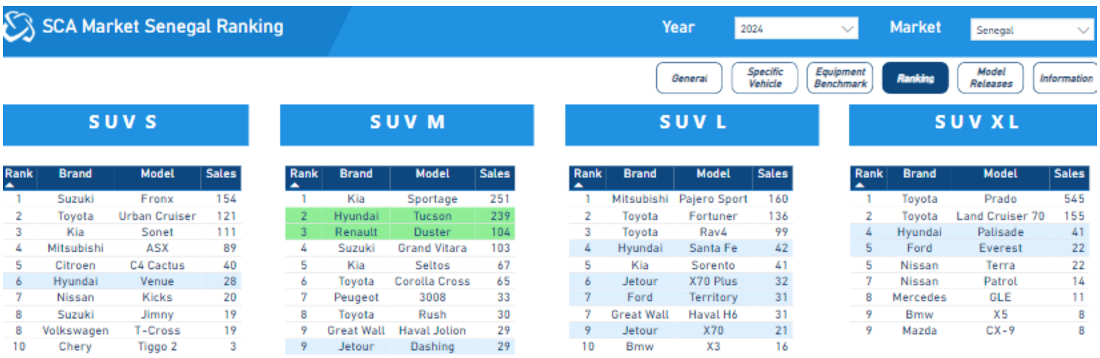


Figure 16: Ranking Page

1.5 Model Releases Page

The fifth page tracks the initial market introduction of each model across different countries. The page features:

- A graph that illustrates the evolution of model launches over time (Figure 17, section A)
- A table displaying the models launched in the year selected in the filter (Figure 17, section B)

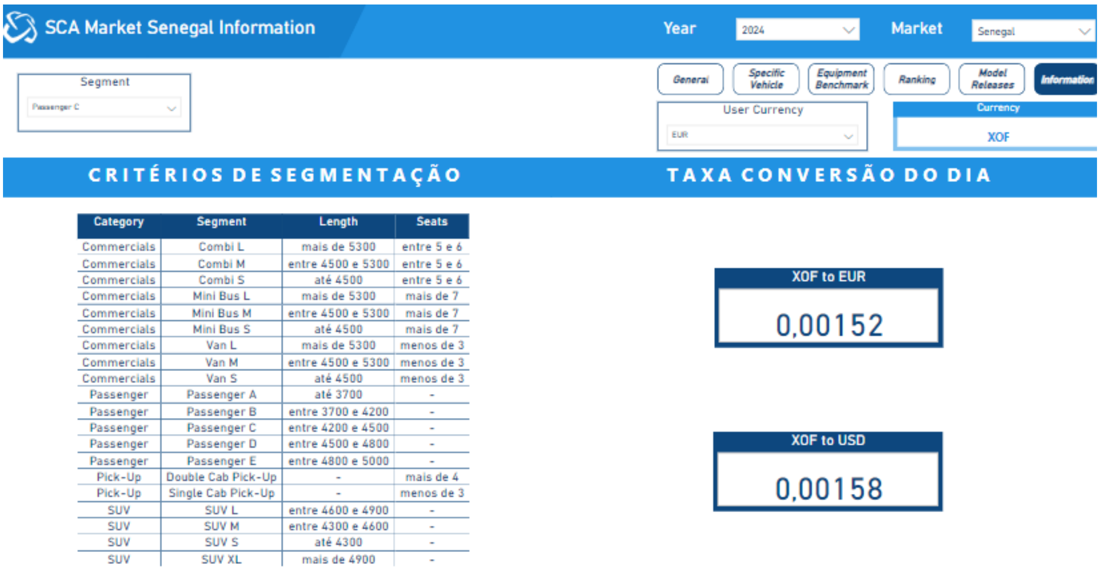
Figure 17: Model Releases Page

1.6 Information page

The dashboard's final page, Information, provides detailed information to support a deeper understating of the analysis framework. The page presents:

- Segmentation criteria
- Daily conversion rate
- Mode of the equipment features
- Evaluation criteria (scores) for both vehicle specifications and equipment

This page serves as a reference hub for product managers by consolidating this essential information.



Category	Segment	Length	Seats
Commercials	Combi L	mais de 5300	entre 5 e 6
Commercials	Combi M	entre 4500 e 5300	entre 5 e 6
Commercials	Combi S	até 4500	entre 5 e 6
Commercials	Mini Bus L	mais de 5300	mais de 7
Commercials	Mini Bus M	entre 4500 e 5300	mais de 7
Commercials	Mini Bus S	até 4500	mais de 7
Commercials	Van L	mais de 5300	menos de 3
Commercials	Van M	entre 4500 e 5300	menos de 3
Commercials	Van S	até 4500	menos de 3
Passenger	Passenger A	até 3700	-
Passenger	Passenger B	entre 3700 e 4200	-
Passenger	Passenger C	entre 4200 e 4500	-
Passenger	Passenger D	entre 4500 e 4800	-
Passenger	Passenger E	entre 4800 e 5000	-
Pick-Up	Double Cab Pick-Up	-	mais de 4
Pick-Up	Single Cab Pick-Up	-	menos de 3
SUV	SUV L	entre 4600 e 4900	-
SUV	SUV M	entre 4300 e 4600	-
SUV	SUV S	até 4300	-
SUV	SUV XL	mais de 4900	-

Figure 18: Information Page

Chapter 6

1 Results Discussion

To evaluate the dashboard's effectiveness and usability, a survey (see Appendix) was conducted among product managers who regularly use the tool in their decision-making process.

The responses obtained reinforce the benefits of BI tools described in the literature. As highlighted by Tunowski (2015), BI improves the speed of data handling and enhances the quality of decision-making. The dashboard developed in this project, supported by a normalized and structured dataset, aligns with these findings and demonstrates the real impact of integrating BI into operational routines.

According to the results, 80% of product managers considered the dashboard "very easy to use," while the remaining 20% rated it as "relatively easy to use." This positive evaluation confirms the design choice, particularly regarding layout, logical flow, and user interface. Moreover, 100% of respondents confirmed that the dashboard presents the most relevant information at the top, according to Sisense's best practices, which recommend prioritizing essential context to be accessible within five seconds.

Regarding logical storytelling, which Sisense (2023) emphasizes as crucial in dashboard design, 100% of users agreed that the dashboard effectively guides them through the information logically.

Regarding visual density, 80% of respondents considered the number of visual elements per page balanced, while 20% felt that it significantly facilitated analysis. This indicates that limiting the number of charts per page was effective and well-received.

The perceived effectiveness of the dashboard in supporting pricing decisions was also significant. Eighty percent of respondents stated that the dashboard fully supports their pricing decision, with the remaining twenty percent agreeing that it does so in most situations. Similarly, eighty percent reported that the dashboard has significantly reduced the time required for analysis and decision-making. This finding mirrors Tirupati et al. (2023), who demonstrated that Power BI improves time efficiency and user satisfaction.

2 Final Conclusions

This dissertation aimed to improve the vehicle launch decision-making process at Salvador Caetano by normalizing the available datasets and developing a Power BI dashboard. The results confirm that all proposed objectives were successfully met.

The work demonstrated that transforming unstructured and inconsistent datasets into a normalized relational model significantly enhances data quality, consistency, and usability. This structured foundation enabled the development of a Power BI dashboard tailored to the actual needs of product managers.

The methodology adopted, based on iterative Action Research cycles, proved effective in aligning technical improvements with business needs. Each iteration allowed for reflection, adjustment, and continuous user feedback.

Feedback collected through a survey confirmed the dashboard's effectiveness. Users reported that the tool was easy to use, logically structured, and visually balanced. Most importantly, it significantly reduced the time required to perform price positioning analyses.

Beyond the technical results, this work revealed the importance of integrating data quality, visualization principles, and user engagement in the design of BI systems.

In conclusion, this dissertation contributes both academically and practically to BI. It provides evidence that structured, well-visualized, and user-oriented BI tools can transform the way organizations manage information and make decisions.

3 Limitations and possible future work

Despite the positive outcomes, one limitation of this project is the continued dependence on manual data entry. Although data normalization significantly improves structure and accuracy, the initial data input process still depends on manual work by product managers, which may introduce errors.

Future improvements could focus on automating data collection and integration to overcome this limitation. A promising approach would be the development of AI - based systems capable of automatically retrieving updated information directly from each brand.

In addition to automation, implementing real-time validation mechanisms, such as automatic error detection and correction suggestions, would enhance data reliability and ensure consistent, high-quality inputs across all datasets.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of my written work/thesis, "Data Normalization and The Role of Power BI for Visualization in: Enhancing Product Sales Decisions", ChatGPT (OpenAI) was used for the following tasks:

- Refining the structure and language of written sections
- Reformulating and rephrasing paragraphs to improve clarity and coherence
- Supporting the organization and articulation of ideas based on research and methodology

The prompts used are listed at the end of the document in the Prompts List section.

After using this tool, I reviewed and edited the content as necessary, and I take full responsibility for the content of the work presented.

I also declare that I am aware of and respect the Artificial Intelligence Rules of Conduct of Católica Porto Business School.

Prompts list

- Help me rewrite this paragraph with a more academic tone.
- Rephrase this text avoiding gerunds and making it more fluid.
- Suggest a section title based on this content.

Bibliography

Alzghoul, A., Khaddam, A. A., Abousweilem, F., Irtaimeh, H. J., & Alshaar, Q. (2024). *How business intelligence capability impacts decision-making speed, comprehensiveness, and firm performance.* Information Development, 40(2), 220–233. https://doi.org/10.1177/02666669221108438

Attanapola, K., & Iyer, A. (2024). A new era in BI: Overcoming low adoption to make smart decisions accessible for all. IBM. https://www.ibm.com/think/insights/business-intelligence-adoption

Avison, D., Lau, F., Myers, M., & Nielsen, P. A. (1999). *Action Research.* In Communications of the ACM (Vol. 42, Issue 1). https://doi.org/10.1145/291469.291479

Babić, A., & Zron, A. (2024*). Business Intelligence tools in the interpretation of the ranking of smart cities.* 2024 47th ICT and Electronics Convention, MIPRO 2024 - Proceedings, 211–217. https://doi.org/10.1109/MIPRO60963.2024.10569892

Baxendale, P., & Codd, E. F. (1970). *A relational model of data for large shared data banks.* https://doi.org/10.1145/362384.362685

CelerData. (2024). What is a data overload. https://celerdata.com/glossary/data-overload

Chen, H., Chiang, R. H. L., Storey, V. C., & Robinson, J. M. (2012). *Business Intelligence and analytics: From big data to big impact.* MIS Quarterly, 36(4), 1165–1188. https://doi.org/10.2307/41703503

Chris, K. (2023). Database Normalization – Normal Forms 1nf 2nf 3nf table examples. FreeCodeCamp. https://www.freecodecamp.org/news/database-normalization-1nf-2nf-3nf-table-examples/

Datafortune. (2024). BI in Real-Life: *Case studies showcasing the impact of business intelligence services.* https://datafortune.com/bi-in-real-life-case-studies-showcasing-the-impact-of-business-intelligence-services/

Dickens, L., & Watkins, K. (1999). *Action research: Rethinking Lewin.* Management Learning, 30(2), 127–140. https://doi.org/10.1177/1350507699302002

Gaspar, C., & Pereira Da Silva, V. (2020). *Influência do business intelligence no controlo de gestão e no desempenho organizacional: um estudo de caso.*

Gokila, D., & Balasubramani, S. (2019*). Impact of normalization in future.* International Journal of trend in scientific research and development, 3(5), 153-156.

Gowthami, K., & Pavan Kumar, M. R. (2017). *Study on business intelligence tools for enterprise dashboard development.* International Research Journal of Engineering and Technology, 4(6), 2987-2992 .

Juare, S. (2019). *Business Intelligence: concepts, components, techniques and benefits.* IOSR Journal of Engineering (IOSR JEN), 4(4), 48–53.

Lewin, K. (1946). *Action research and minority problems.*

Microsoft. (2023). DAX overview. Microsoft Learn. https://learn.microsoft.com/en-us/dax/dax-overview

Microsoft. (2024). Intro to dashboards for Power BI designers. Microsoft Learn. https://learn.microsoft.com/en-us/power-bi/create-reports/service-dashboards

Morris, S. (2024). Data normalization: Definition, importance, and advantages. Coresignal. https://coresignal.com/blog/data-normalization/

Petersen, K., Gencel, C., Asghari, N., Baca, D., & Betz, S. (2014*). Action research as a model for industry-academia collaboration in the software engineering context.* In Proceedings of the 2014 International Workshop on Long-Term Industrial Collaboration on Software Engineering, 55–62. https://doi.org/10.1145/2647648.2647656

Qlik. (n.d.). Why Qlik. https://www.qlik.com/us/why-qlik-is-different

Reginato, L., & Nascimento, A. (2007). *A case study of business intelligence as an instrument to support controllership.*

Richardson, J., Schlegel, K., Sallam, R., Kronz, A., & Sun, J. (2021). *Magic quadrant for analytics and business intelligence platforms*. Gartner. https://www.gartner.com/doc/reprints?id=1-253W9DLN&ct=210129&st=sb

Sari, A. (2021). *Utilisation Of business intelligence for data mining to influence decision-making in business organisations.* Turkish Online Journal of Qualitative Inquiry, 12(7), 13043-13051.

Schwabish, J. (2021). *Better Data Visualizations: A guide for scholars, researcher, and wonks.* Columbia University Press.

Sharma, M., Puri, C., & Kumar, P. (2024). *A review: Transformative impact of data visualization across various industries.* In Proceedings of the 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), 799–804. https://doi.org/10.1109/I-SMAC61858.2024.10714609

Sisense. (2023). Dashboard design best practices – 4 key principles. Sisense. https://www.sisense.com/blog/4-design-principles-creating-better-dashboards/

Stark, M. (n.d.). Why business intelligence is important. Premier International. https://www.premier-international.com/articles/why-business-intelligence-is-important

Tirupati, K. K., Singh Nagar, A., Joshi, A., Singh, S. P., Chhapola, A., Jain, S., & Gupta, A. (2023). *Leveraging Power BI for enhanced data visualization and business intelligence.* Universal Research Reports, *10*(2), 1–6. 676-711. https://doi.org/10.36676/urr.v10.i2.1375

Tunowski, R. (2015). *Business Intelligence in organization. Benefits, risks and developments.* Przedsiebiorczosc i Zarzadzanie, 16(2), 133–144. https://doi.org/10.1515/eam-2015-0022

Vyas, A. (2024). *Unlocking the Power of Data Visualization.* The CPA Journal.

Wanda, P., & Stian, S. (2015). The secret of my success: An exploratory study of business intelligence management in the Norwegian industry. Procedia Computer Science, 64, 240–247. https://doi.org/10.1016/j.procs.2015.08.486

Appendix

1 Product Manager Feedback Questionnaire

This questionnaire was created to gather feedback from product managers who used the Power BI dashboard developed during this project. The goal was to evaluate the dashboard's usability, structure, visual clarity, and impact on the decision-making process related to product positioning.

- Is the dashboard easy to use?
    - Very easy
    - Relatively ease
    - Neither easy nor difficult
    - Difficult
    - Very difficult

- Does the dashboard structure clearly present the most relevant information at the top of the page?
    - Yes, completely
    - Partially
    - No

- Do you think the dashboard logically guides the user through the information?
    - Yes
    - Partially
    - No

- Does the number of visual elements per page makes analysis easir or more difficult?
    - It facilitates analysis significantly
    - It is balanced
    - It slightly hinders analysis
    - It hinders analysis a lot

- Are the presented visualizations appropriate to convey the indeed information?
    - Yes, completely
    - In most cases
    - In some cases, they are not
    - Many visualizations could be improved

- Since the implementation of the dashboard, do you believe the time needed for analysis and decision-making has decreased?
    - Yes, significantly
    - Yes, moderately
    - No difference
    - The process became slower

- Does the dashboard effectively support your price positioning decisions
    - Yes, completely
    - Yes, in most situations
    - Partially
    - No