Universitat de Lleida

TREBALL FINAL DE GRAU

ESCOLA POLITÈCNICA SUPERIOR
UNIVERSITAT DE LLEIDA
INSPIRING THE FUTURE

Estudiant: Pau Torné Enrich

Titulació: Doble Grau en Enginyeria Informàtica i ADE

Títol de Treball Final de Grau: Integration and Analysis of Company Reporting Data: A Business Intelligence Approach to Enhance Economic Activity Insights

Director/a: **Marta Oliva Solé i Sergi Balagueró Oliva**

Presentació

Mes: Maig

Any: 2024

Content

List of tables and figures

Abstract and Key Words

For a company, it is essential to have different reports where commercial and financial data can be analysed. These indicators are called KPIs, and it is crucial that the information they contain is reliable and accurate. The problem in our case consists of several factors, led by the inconsistency of data between reports, which is caused by the different data sources and constantly changing criteria set by the various departments of the company. The proposed solution involves building a new dataset with a standardized criterion that unifies all reporting information and can represent all data of this type, so that all company reports dealing with this data can rely solely on this set.

This solution requires knowledge of Computer Engineering such as software engineering and databases. However, it is also essential to understand the company's business model because if we do not understand the commercial and financial structure of a company, we cannot discern what we want and do not want to include. This is why knowledge of Business Administration and Management plays an important role, emphasizing management control and budgeting.

The methodology chosen for the project development has been SCRUM, belonging to the AGILE type methodologies, since given the characteristics of the project, it was considered the one that fits best.

Thanks to the successful development of this project, the company has been provided with a solution to the previously described problem, providing homogeneity and consistency in the data and including various visualizations in Power BI, according to the analysis to be carried out, also improving efficiency in the company's reporting structure, since a single dataset now feeds numerous reports.

Key words: Business Intelligence, Management Control, SQL Server, Datawarehouse, Budgeting, UML, Frontend and Backend.

1. Introduction

1.1. Context and justification

In the contemporary business environment characterized by rapid change and intense competition, companies must continually evolve to remain viable and competitive. A significant shift has occurred in the traditional decision-making paradigm, with organizations increasingly turning to advanced data analytics technologies to harness emerging opportunities and maintain their edge in the market.

One such transformative technology is Business Intelligence (BI). The integration of BI has heralded profound enhancements in data analysis and decision-making for numerous companies, including our own. As we embark on the development of a BI project, it becomes evident that a comprehensive understanding of the company's business model, coupled with proficiency in showcasing it to users, is paramount. Additionally, a deep grasp of databases and optimization processes is indispensable for the successful execution of such initiatives.

As students enrolled in the Double Degree program in Computer Engineering and Business Administration and Management, we find ourselves uniquely positioned to undertake projects of this nature. Our interdisciplinary background equips us with the requisite expertise to seamlessly integrate insights from both disciplines, enabling us to effectively apply knowledge acquired across diverse domains.

Central to the success of BI initiatives is the assurance of data consistency and reliability. This project underscores our commitment to constructing a unified dataset that can support a myriad of reporting requirements. By establishing a robust foundation for data management, we aim to enhance organizational efficiency and facilitate informed decision-making processes.

On a personal level, this project represents a compelling opportunity for professional and educational growth. Delving into uncharted territory has sparked a profound interest, while the prospect of gaining insight into the inner workings of corporate decision-making processes serves as a powerful motivator. Through this endeavour, we anticipate not only expanding our technical proficiencies but also gaining invaluable insights into the complexity of organizational dynamics.

In summary, the convergence of our academic pursuits in Computer Engineering and Business Administration and Management aligns seamlessly with the objectives of this BI project. As we embark on this journey, we remain steadfast in our commitment to delivering a solution that not only meets the needs of our organization but also fosters personal and professional development.

1.2. Project goals

The main objectives related to the Computer Engineering Degree would be:

- To implement a data model that suits the best in our case.
- To build this relational data model in SQL, optimizing its calculation time.
- To design an intuitive and attractive user interface for all type of users.

Related to the Business Administration and Management degree:

- To understand the business model, providing the best effective tools for facilitating decision-making and problem solving.
- To have the ability to audit economic data to ensure its reliability.
- To understand and apply the knowledge acquired from budgeting and its different ways of consecution measurement.

In terms of personal, professional, and educational development:

- To see and understand the functioning of a large company and decision-making process.
- To acquire better knowledge of Business Intelligence and databases concepts.

1.3. Methodology

After considering different methodologies for developing the project, such as waterfall model or AGILE methodologies, and analysing the pros and cons of applying them it was decided to apply a mix and adaptation of the well-known Agile methodology Scrum and Iterative Design Process.

Agile methodologies are a set of principles and practices for software development that prioritize flexibility, adaptability, collaboration, and continuous improvement. The core

values of agile methodologies, as outlined in the Agile Manifesto, include valuing individuals and interactions over processes and tools, working software over comprehensive documentation, customer collaboration over contract negotiation, and responding to change over following a plan. Agile methodologies emphasize iterative development, where small, cross-functional teams work in short cycles (iterations or sprints) to deliver working software frequently, incorporating feedback from stakeholders along the way. One of the most popular Agile methodologies is Scrum. It is a lightweight, iterative approach to project management that enables teams to deliver value quickly and adapt to change efficiently. Here are some key elements of Scrum:

- Roles: Scrum defines three primary roles: the Product Owner, who represents the stakeholders and defines the product backlog; the Scrum Master, who facilitates the Scrum process and removes impediments; and the Development Team, who are responsible for delivering increments of potentially shippable product at the end of each sprint.
- Artifacts: Scrum defines several artifacts, including the Product Backlog, which is a prioritized list of all desired work on the project; the Sprint Backlog, which is a subset of the Product Backlog items selected for the sprint; and the Burn-up or Burn-down chart which measures the amount of Product Backlog items assigned to the current Sprint that are pending completion.
- Events: Scrum events provide the framework for the team to collaborate and inspect and adapt their work. These events include Sprint Planning, where the team plans the work for the upcoming sprint; Daily Standup (or Daily Scrum), where team members synchronize their work and plan for the day; Sprint Review, where the team demonstrates the completed work to stakeholders; and Sprint Retrospective, where the team reflects on their process and identifies areas for improvement.

When applying this methodology to our project there were some Roles, Artifacts or Events that were not considered, due to the characteristics of the project. For example, in our case the Development Team was formed exclusively by me, or the Daily Meetings were not organised every day. Our way of working was the following one:

While developing the project some meetings were held by my company tutor and me. These meetings were arranged whether there was a significant progress (which would be the equivalent to Sprint Reviews), or also if a problem arose or there was a slight change

in the purposes of the project (which would be similar to Daily Meetings). It is important to mention that if there was any change in the criteria defined for the client classification, type of product…, a reunion was also held between them and us for being up to date in the aspects that our project deals with.

In the development of the technical part the followed procedure consisted of an iterative and incremental development. This is one of the characteristics of the iterative methodology, and consists in developing the project in increments, with each iteration adding new functionality or refining existing features. On the one hand, in the case of building the set of data in SQL, each iteration consisted of including a new table such as Sales, Budget, Stock… On the other hand, for the BI project each iteration meant building a new page or including more data and visual object in existing pages.

So, as mentioned before, mixing and adapting the iterative development and the Scrum methodology we could develop the project with a high adaptability, continuous improvement, early and continuous feedback and risk reduction, dealing also with small and almost insignificant cons such as the complexity of these methodologies, the dependency on feedback and the impossibility of applying them completely, having to adapt and mix them.

2. Business Intelligence

2.1. Basic concepts of Business Intelligence

Business Intelligence (BI) refers to the technologies, processes, and strategies used by organizations to analyse and interpret their data to make informed business decisions. It encompasses a variety of tools, applications, and methodologies that enable companies to gather, store, and analyse data from various sources such as internal systems, databases, and external sources.

The following are some basic concepts related to BI, which will be described in more detail and referenced all along the report.

- *Data Warehousing*: BI often relies on data warehouses, which are central repositories of integrated data from one or more disparate sources. These data warehouses are designed to support decision-making processes by providing a unified view of data for analysis and reporting.
- *OLAP (Online Analytical Processing)*: OLAP enables users to analyse multidimensional data interactively from different perspectives. It allows for complex analytical queries to be performed quickly and efficiently, supporting tasks such as reporting, budgeting, forecasting, and trend analysis.
- *ETL (Extract, Transform, Load)*: ETL is the process of extracting data from various sources, transforming it into a consistent format, and loading it into a data warehouse or other storage systems for analysis. ETL tools automate these processes, ensuring data quality and consistency.
- *Data Visualization*: Data visualization is the graphical representation of data to facilitate understanding and interpretation. Visualization techniques such as charts, graphs, and tables help users identify trends, outliers, and patterns in the data more effectively than raw numbers or text.

2.2. Power BI and its features

There are numerous tools available for developing BI solutions, each offering different functionalities, features and capabilities. Some of them are Tableau, QlikView/Qlik Sense and MicroStrategy. Although all those are powerful tools, the one that best fits in our case,

due to the characteristics of our project and the situation of the company is Power BI. The main features of this tool developed by Microsoft that made us decide to use it, are the following:

- *Data Connectivity*: Power BI supports a wide range of data sources, including Excel spreadsheets, databases, cloud services and many others. It enables users to import data for analysis easily.
- *Data Modelling*: Provides robust data modelling capabilities, allowing users to create relationships between different data tables, define calculated columns and measures, and build hierarchies.
- *Interactive Visualizations*: Has an extensive library of interactive visualizations, including charts, graphs, maps...
- *Scheduled Updates:* Allows you to schedule updates at the desired periodicity for each report independently.
- *Intuitive and clear:* Power BI has a very intuitive interface that allows users to get used to it easily.
- *Being part of Microsoft Office 365:* Being part of the Microsoft environment, which is the one used by the company, allows us to connect it with other apps such as Excel or Power Point. This unification on the Microsoft Office makes it also more affordable for the company.

3. Company overview

3.1. Company description

3.1.1. Sector and business scope

Rodi Motor Services is dedicated to the repair and maintenance of all types of vehicles (cars, trucks, agricultural vehicles, motorcycles, etc.), focusing primarily on the early stages of its history on tire replacement. Nowadays, however, the company offers comprehensive mechanical services, thereby increasing the number of clients and their loyalty to the company. Additionally, it also provides roadside assistance services.

As for the types of garages, there are two different ones available: mixed auto repair shops and those specialized in cars. In the first ones, services for all types of vehicles are offered, including trucks, agricultural vehicles, cars, and motorcycles, as well as vans and pick-up trucks. In car workshops, only light vehicles are repaired, including cars, motorcycles, vans, and light trucks. Regarding clients, all types are attended. Therefore, services are provided to individuals, companies, fleets…

Nowadays RODI owns more than 160 auto repair shops, with more than 1.700 employees and is the leader in its sector in Catalonia, Aragon and the Canary Islands, with presence in Portugal, Andorra and other provinces of Spain.

The company is divided into five businesses, which are Rms (North-West of Spain), Rodi Galicia (Galicia), El Paso (Canary Islands), Covipneus (Portugal) and Franchises.

Its business model, based on being the leader wherever the company is expanded, has led to numerous successes, including a turnover of over 260 million euros in 2023.

3.1.2. Internal organization

The internal organization of Rodi is the following one. At the highest level is the general management, which reports to four sections headed by a director. These are:

- Truck, Industrial and Agricultural: Takes care of all the business related to heavy transports, such as trucks, tractors…

- Light and Motorcycle: Takes care of all the business related to light transports, such as cars, motos…
- Corporate Management: Coordinates those departments of central offices that do not take part of the previous sections. Some of them are HR, Management Control, Expenses and investments…
- Operations Management: Coordinates the different auto repair shops. Those are divided in groups of 12-16, and each group has an operation director that supervises and manages the mechanical workshops.

Finally, each department is headed by a department responsible, and each mechanical workshop is lead by a responsible, too.

3.1.3. Decision making

The reporting section is crucial for the company when taking important business decisions, such as diversification on the products or market, changes of prices, opening of new garages, etc,

This section provides to the other departments all kind of data required so that it can be analysed and decisions can be taken based on this data.

Furthermore, every month there are three types of reunions: KPI (Key Performance Indicators) reunions, Operational reunions and Bimonthly reunions. The first ones deal with all types of KPIs, mostly divided by business. These are held monthly. The second ones include KPIs and other business strategies, changes… They are divided mainly by operation director and held monthly, too. Finally, the Bimonthly reunions include KPIs and information of all departments, divided by groups of mechanic workshops. Those groups are built considering the auto repair shop type and the geographical zone.

Most of the department managers, operation directors and senior managers assist to these reunions, which are also prepared by the reporting team, so it is very important to be sure that all the provided data is reliable.

3.2. Previous situation of the company's reporting section

For some years now, the company has had a report that we will call "Master", which contains the billing objectives of the auto repair shops alongside with their achievement status. For the preparation of this report, a sales classification criterion was created. The data in this report is considered valid when analysing billing and is the most important and reliable indicator for measuring the company's economic activity.

Currently, there is a large number of reports and data, and many of them have their own set of data independent of others, following the criteria considered appropriate at the time. This causes that different reports which include data that references on the same garage, the same product, etc., which, a priori, should have the same value in the data, do not. This discrepancy in the value of the data means that 100% reliability with the reports cannot be achieved. Complaints are often received from, among others, garage managers because one report shows that they have sold a certain number of, for example, truck tires, while another report shows a different value. This generates a problem because auto repair shops receive bonuses based on the percentage of sales, they achieve against the targets assigned to them, and it is important for managers to know for certain that the data they are analysing is correct.

4. Implementation of the BI Solution

4.1. Previous formation

As said in previous sections, this project demands a high level of company knowledge and understanding, as well as in databases and concepts related to Computer Engineering. That's why, at first, it was necessary to carry out a phase of learning and training.

Albeit knowledge from databases was acquired during the degree, previously to the development of the project it was necessary to receive some training on databases and Business Intelligence. This training also aimed on understanding the company's business model, the development of the activity and the way of working they carry out every day.

This training was received from the employee in charge of the BI Projects of the company during all the internship in the company, that took place all along the summer.

This phase of training consisted mostly of *learning by doing*, so by developing small projects I could understand all the aspects said above, acquiring all the knowledge necessary to cope with the project.

It was also necessary to revise all kind documentation from other projects developed in the company, such as the one that consisted of creating the report described in section "*Previous situation of the company's reporting section*" as "Master". This helped to understand the background of this project, which is fundamental for the development of our own.

4.1.1. Relation with other projects developed in the company

During the formation stage, revising other projects done in the company, we realized that we could include another purpose to the project. This purpose involves implementing a new presentation system for bi-monthly, KPI and operations meetings. The goal is for this system to be implemented in Power BI and to be much more efficient than the current one. We aim to create a report that has the structure of a presentation through which meetings can be conducted.

Last year, a classmate from the double degree, one year ahead, precisely carried out this task, focusing more on optimizing it using Excel, exporting data with ODBC queries to

the database, and automating the data update process. However, in terms of meeting preparation, the task is still quite laborious because one must go through each file, update the data, audit it, and copy the tables of interest to the PowerPoint presentation.

With the approach we propose now, there would be no need for any preparation tasks because the Power BI report would update itself and contain the correct data (since we use the previously created dataset).

4.2. Needs identification

In order to develop a successful and useful project it is essential to identify the needs of the company which must be satisfied.

As described previously in the document, in our case the main needs of the company are:

- Solving the discrepancy of the different KPIs between multiple reports.
- Having a set of data that can compare budget versus income, as well as stock, purchases and other activities of the company.
- Showing this data in an accurate BI report.
- Implementing a more efficient and automated way for developing the reunions.

4.3. Tasks' planification

A clear planification of the tasks to be undertaken is crucial for a successful elaboration of the project. The following are the tasks to be done:

1. Formation phase.
2. Needs identification.
3. Requirements analysis.
4. Analyse the criteria followed by the Master for classifying the sales. Analyse also the other tables that will be necessary for the project, the way they are created and how can we integrate all these tables.
5. Following this criterion building the dataset with SQL, creating new tables and adding all the columns or features that are necessary.
6. Auditing the data so that we can be sure that everything is correct. If it's not okay, we should return to step 2 (or maybe step 1).

7. Once we know everything is correct, importing the tables to PowerBI and connecting them, if necessary, with intermediate tables.
8. Ask to the people who analyse this data which indicators they like to see and make some research using the meetings available in the company's SharePoint.
9. Start building the report with Power BI. This process includes obtaining the KPIs using measures and building the visualization of this data.
10. Finding the way of automatizing the meetings and applying it.
11. Write the thesis.
12. During all the process we must be aware of new necessities that can appear and new functionalities to add. We may add some new tasks to this list or remake and adjust some other ones.

Task	July	August	September	October	November	December	January	February	March
1	X								
2		X							
3		X							
4		X							
5		X	X						
6			X						
7			X						
8			X	X	X	X			
9				X	X	X	X		
10							X	X	
11								X	X
12	X	X	X	X	X	X	X	X	X

Table 1: Detailed tasks planification. Own Source

4.4. Requirements Analysis

Before the implementation of any type of project it is crucial to perform a requirements analysis. This consists of the application of a series of techniques to detect the functional and non-functional system requirements.

4.4.1. Elicitation techniques

Elicitation techniques are methods used to gather requirements, information, or feedback from stakeholders. The identification of requirements is strictly linked to the application of these techniques. The following are the ones applied on our project:

- *Interview*: Interviews with the different stakeholders have been conducted in order to know which features the project may include. These interviews have been held with final users, which are the different responsibles of the sales and pricing departments, of both truck and car sections.
- *Prototyping*: Prototyping, which consists of building rough models or mock-ups of a product or system, have given us the opportunity to give to the user an initial version of the project to gather feedback and refine requirements.

4.4.2. Defined requirements

After the application of the techniques described above, the requirements defined for our project have been the following ones, divided into functional and non-functional requirements.

Functional requirements

- Get the different KPIs shown in each reunion.
- Get the business operations result.
- Get all the data (including budget) segmented by client type, type of business, business, director, centre, type of product, brand…
- All the data must be reliable and must match the "Master".
- Get the data for different periods of time, containing data from the last two years.
- Build as many different interfaces in Power BI as demanded.

Non-functional requirements

- The BI project must not weight more than 500 Mega bytes.
- The response time of the different visualization must be under 3 seconds.
- Intuitive and clear interface.
- The data set must be updated in less than 30 minutes.
- There report must be visible in a laptop but also in a mobile phone/tablet.

- There must be different roles to ensure that not every type of user has access to all the data displayed.

4.5. Data Model Construction

4.5.1. Company's Data Architecture

The company employs a Datawarehouse framework crafted to facilitate data analysis tasks. This framework effectively serves its purpose by housing the company's activity data via the respective ETL processes. It encompasses a stable database intended for integrating and standardizing data in a historical format. All the data is extracted from the ERP used by the company, Business Central.

The Datawarehouse structure is the represented in figure 1:

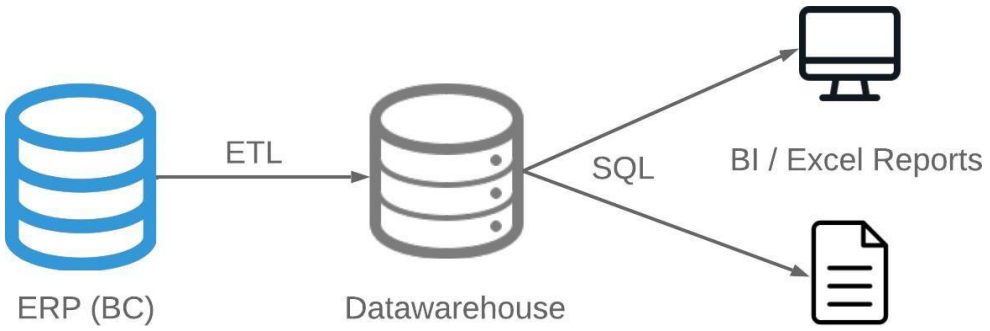


Figure 1:Datawarehouse structure of the company. Own Source

The data from the ERP is Extracted, Transformed and Loaded (ETL) to the Datawarehouse. These actions are developed by the company's IT team. Once we have the data in the Datawarehouse, SQL queries are applied by our team, allowing us to load (or add new tables) into the reporting tools such as Power BI, Excel, DBMaster… Once this data is displayed in these tools, final users can analyse it successfully.

This data architecture enables us to have a lot of control over the data, being able to create new tables, transform others and establishing all kind of relationships between tables. Another advantage of it is that processes can be created, so that some calculations can be

held during inactive hours, avoiding conflicts between these calculations and data loading.

The following is an example of the output of a Sales query launched to the Datawarehouse. Note that data values have been changed due to the company's privacy terms.

Empresa	tipo	Centro	Tienda	Documento	linea	Corresponde_Albaran	Nif	Cliente	fecha
Empresa_4786591	Factura	Centro_3336052	Tienda_3738016	Documento_3968744	247.500	Corresponde_Albaran_1246377	Nif_4667574	Cliente_3364490	2024-01-08 00:00:00.000
Empresa_3727417	Abono	Centro_2202237	Tienda_3129384	Documento_68907	80.000	Corresponde_Albaran_2321331	Nif_2370477	Cliente_594193	2024-02-22 00:00:00.000
Empresa_4791909	Factura	Centro_4788862	Tienda_253152	Documento_328036	1.960.000	Corresponde_Albaran_4174771	Nif_2396240	Cliente_3321652	2024-02-14 00:00:00.000
Empresa_4791817	Factura	Centro_4786389	Tienda_1322429	Documento_328074	530.000	Corresponde_Albaran_4174730	Nif_2393229	Cliente_3321348	2024-02-14 00:00:00.000
Empresa_4360978	Factura	Centro_1704695	Tienda_2780627	Documento_1361458	140.000	Corresponde_Albaran_3617934	Nif_2607430	Cliente_3058738	2024-04-08 00:00:00.000
Empresa_3718955	Factura	Centro_1952509	Tienda_2981068	Documento_3496384	150.000	Corresponde_Albaran_3767241	Nif_2609947	Cliente_3103720	2024-01-26 00:00:00.000
Empresa_1078065	Factura	Centro_2437058	Tienda_3270984	Documento_2289851	10.000	Corresponde_Albaran_3892327	Nif_2666154	Cliente_3758841	2024-01-24 00:00:00.000
Empresa_4525144	Factura	Centro_2104360	Tienda_3075276	Documento_554089	100.000	Corresponde_Albaran_2628677	Nif_451323	Cliente_346851	2024-02-12 00:00:00.000
Empresa_3675356	Factura	Centro_1952339	Tienda_2985256	Documento_1471288	15.000	Corresponde_Albaran_3768516	Nif_3579259	Cliente_2945832	2024-02-14 00:00:00.000
Empresa_3674779	Factura	Centro_1952723	Tienda_2985280	Documento_1470718	20.000	Corresponde_Albaran_3767458	Nif_3582420	Cliente_2957770	2024-02-03 00:00:00.000
Empresa_978707	Factura	Centro_2321197	Tienda_3206237	Documento_2234625	72.499	Corresponde_Albaran_3826799	Nif_915287	Cliente_1253902	2024-01-23 00:00:00.000
Empresa_2144037	Abono	Centro_2520731	Tienda_3333408	Documento_84891	50.000	Corresponde_Albaran_2322330	Nif_2405715	Cliente_1671176	2024-04-17 00:00:00.000
Empresa_3674784	Factura	Centro_1950425	Tienda_2983877	Documento_1470705	70.000	Corresponde_Albaran_3767464	Nif_3573803	Cliente_2944610	2024-02-03 00:00:00.000
Empresa_4793281	Factura	Centro_4788124	Tienda_490401	Documento_371749	550.000	Corresponde_Albaran_1524955	Nif_2393882	Cliente_3322095	2024-03-26 00:00:00.000
Empresa_3524468	Factura	Centro_3004458	Tienda_3609014	Documento_704759	60.000	Corresponde_Albaran_2818980	Nif_2298287	Cliente_2774204	2024-01-30 00:00:00.000
Empresa_4838311	Abono	Centro_642299	Tienda_2161503	Documento_4482	90.000	Corresponde_Albaran_2228095	Nif_4454982	Cliente_753121	2024-03-07 00:00:00.000
Empresa_4615602	Factura	Centro_2726848	Tienda_3487385	Documento_406111	10.000	Corresponde_Albaran_2425135	Nif_1873114	Cliente_148827	2024-04-06 00:00:00.000
Empresa_4137193	Factura	Centro_1215561	Tienda_2467536	Documento_1156213	10.000	Corresponde_Albaran_3376907	Nif_1285348	Cliente_516727	2024-04-05 00:00:00.000
Empresa_3815542	Factura	Centro_202045	Tienda_1990858	Documento_848489	50.000	Corresponde_Albaran_2992613	Nif_3563570	Cliente_3237129	2024-03-01 00:00:00.000
Empresa_3604597	Factura	Centro_1455520	Tienda_2613608	Documento_478528	10.000	Corresponde_Albaran_2527070	Nif_3606379	Cliente_3237009	2024-04-04 00:00:00.000
Empresa_4462377	Factura	Centro_1891122	Tienda_2932292	Documento_384802	10.000	Corresponde_Albaran_2392793	Nif_3597093	Cliente_3237084	2024-02-13 00:00:00.000
Empresa_4209000	Factura	Centro_1349682	Tienda_2557097	Documento_1212623	10.000	Corresponde_Albaran_3439714	Nif_3582690	Cliente_3237137	2024-03-06 00:00:00.000
Empresa_4216839	Factura	Centro_1374780	Tienda_2567098	Documento_1218526	10.000	Corresponde_Albaran_3446255	Nif_3601734	Cliente_3237093	2024-01-18 00:00:00.000
Empresa_4301658	Factura	Centro_1598933	Tienda_2694310	Documento_1296092	10.000	Corresponde_Albaran_3539192	Nif_3571694	Cliente_3237080	2024-01-12 00:00:00.000
Empresa_4441622	Factura	Centro_1804122	Tienda_2877564	Documento_1430389	10.000	Corresponde_Albaran_3702945	Nif_3586495	Cliente_3237058	2024-02-22 00:00:00.000
Empresa_3815543	Factura	Centro_203640	Tienda_1968640	Documento_848491	60.000	Corresponde_Albaran_2992612	Nif_3579748	Cliente_3237117	2024-03-01 00:00:00.000
Empresa_3602569	Factura	Centro_1461585	Tienda_2616739	Documento_476500	10.000	Corresponde_Albaran_2524755	Nif_3574086	Cliente_3237092	2024-03-11 00:00:00.000

Figure 2: Output to Sales table query. Own Source

4.5.2. Followed strategy

4.5.2.1. Previous versions (regarding Sales)

After analysing all processes in charge of calculating all the data of the "Master" and observing the way the information is displayed in the document, the first option for building the dataset was the following one (now we are only considering the Sales table).

The idea was building a table containing all the Sales following the criteria defined by the "Master". This was done inserting all this data in a new table named LabelSales. The name of this table comes from the term 'Label'. 'Labels' are a string of text used in "Master" to classify sales.

Once this table was created, an incremental process for updating it had to be created. This process required the creation of two auxiliar tables. One contained the criteria followed for each label, while the other contained information such as the last update date of a label, which was necessary for beginning to insert sales from this date on.

This process had given a successful result in terms of the reliability of data, albeit its update time took so long (almost one hour). This fact made us take the decision of finding a new way for obtaining the data desired.

4.5.2.2. Definitive version (regarding all types of data)

→ **Sales**

Taking a deep view on the tables used for the "Master" calculation we realized that one of them contained valuable information for us. This table contained all the Sales classified by Label. This table, though, lacked valuable and indispensable data such as the income or cost of each sale line. This problem was solved by joining this table with Sales on different columns (document, line, company, date and client) and selecting all this data from the Sales table. There are mainly two advantages of this version:

- The updating time is, at most, 30 minutes (half of the time it took in the previous version).
- The "Master" updating processes are reused, as we are using a table updated by these processes. We don't have to create and execute new processes.

Finally, we also aimed to classify the sales by client type, type of product and type of business, as the rest of classifications such as family product, brand, etc, are included by default in the Sales table. This classification was made via the structure CASE … WHEN in the SELECT clause. Depending on the label associated with the sale line, a concrete type of client, product, etc, was selected.

Multiple conditions had to be added in order to ensure the correctness of the data and avoiding duplicated data. Furthermore

→ **Budget**

Once all the Sales data was successfully obtained, the next goal was to classify the budget by the same criteria as Sales. The strategy followed for successfully doing it was the following one:

An additional table was created in the Datawarehouse, cloning all the data from the table that contains the budget for all the centres, grouped by item, year and month. In this table, several columns were added. A different column is created for each existing classification. Therefore, there is an extra column for client type, type of business and type of product

(this one is divided into 2 columns, one for discerning whether the sale is labour or product, and another for distinguishing tyres from replacement).

Regarding SQL this classification was also implemented using the CASE ... WHEN clause when updating the table. This clause evaluates the column type of the original budget tables, which gives the description for each amount stablished to a centre, month and year.

→ **Products**

In order to have the opportunity to classify stock and purchases items by the same criteria as the previous tables, a Product classification had to be created. The main idea was to join this tables by product (which is the primary key of the products table) so that the classification could be easily obtained.

The steps followed for the classification of the products table was the same as the one implemented for the budget. A new Product table was created adding new columns for the classification. The only difference is that in this case, the criteria was taken from the one defined for the "Master" instead of from an existing column in the original table.

→ **Other tables**

Regarding other tables included in the data set, such as Centres, Clients, Suppliers… No special changes have been made. The queries used include several conditions to guarantee the correctness of the data and, in some cases, to include additional information that we don't have at first sight, but these cases are not as remarkable as the other ones described above.

All this data selection and tables construction has been carried out using Heidi SQL, a tool provided by the company that, establishing a connection to the Datawarehouse allows us to extract the desired data from this source using the Transact-SQL programming language.

4.6. BI Report Construction

As explained in previous sections, the chosen BI tool has been Power BI, developed by Microsoft, due to the numerous advantages that it has compared to the others.

Building a BI report requires various steps, the following are the ones that have been followed.

4.6.1. Data Importing

Data importing in Power BI is such a simple task. There are many ways available, such as importing data from excel, connecting to cloud-based data sources (Azure, SharePoint, etc.), or direct query to a data source, which has been the one chosen.

A connection must be established between this tool and the Datawarehouse, indicating its server's IP address. Afterwards, for each table that needs to be included in the data model, a SQL query has to be introduced. This involves a task of identifying the data necessary for the data model, as extra information must not be imported for guaranteeing that the report does not occupy more memory than desired. It is important to import all the columns necessary for posterior data modelling, as primary and foreign keys are fundamental.

The picture below shows the importing method in Power BI.

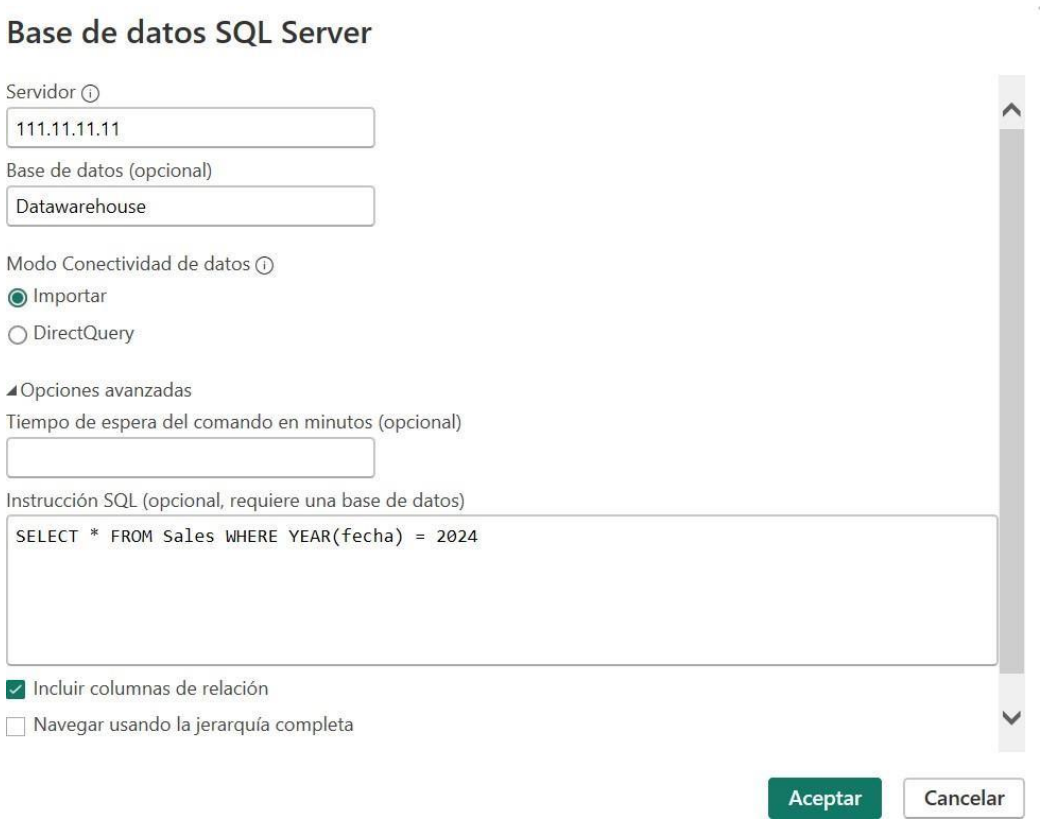


Figure 3: Importing data to Power BI. Own source

For posterior data modelling decisions, some tables had been included manually, as Power BI also has an option for creating your own tables on your own.

4.6.2. Data Modelling

Data modelling plays an important role in our project. The structure of our dataset and the relations between tables are fundamental to show valuable and concise data when this one includes information present in more than one table.

This is not an easy task, as our dataset consists of 35 tables, having each one multiple columns and thousands of rows (even millions). As an example, the Sales table includes around 20 columns and, at the end of the year, consists of over 7 million rows.

Some been connections have been added without difficulty, as the one between Centres table and others like Sales, Budget, … These relations are established by the primary and foreign key, without need of adding any additional column or table.

The complexity of this task of data modelling lies in the filters that we want to apply for the type of product, type of client, etc. Multiple intermediate tables have been added, one for each filter. This tables only consist of a single column with a unique value of the type of client, product, etc. Then, they are connected to each table that include a column with the specific filter (with a relation one to many). This way we can have a unique segmentation of each type that filters all values for all tables. It is necessary to do it that way due to the fact that filters are only applied when the relation between the table that includes the filter and the one filtered is a one to one or one to many relation. Otherwise, a filter for each table should have been added, which generates so much confusion and visual harm to the user.

There is a representation of the whole data model:

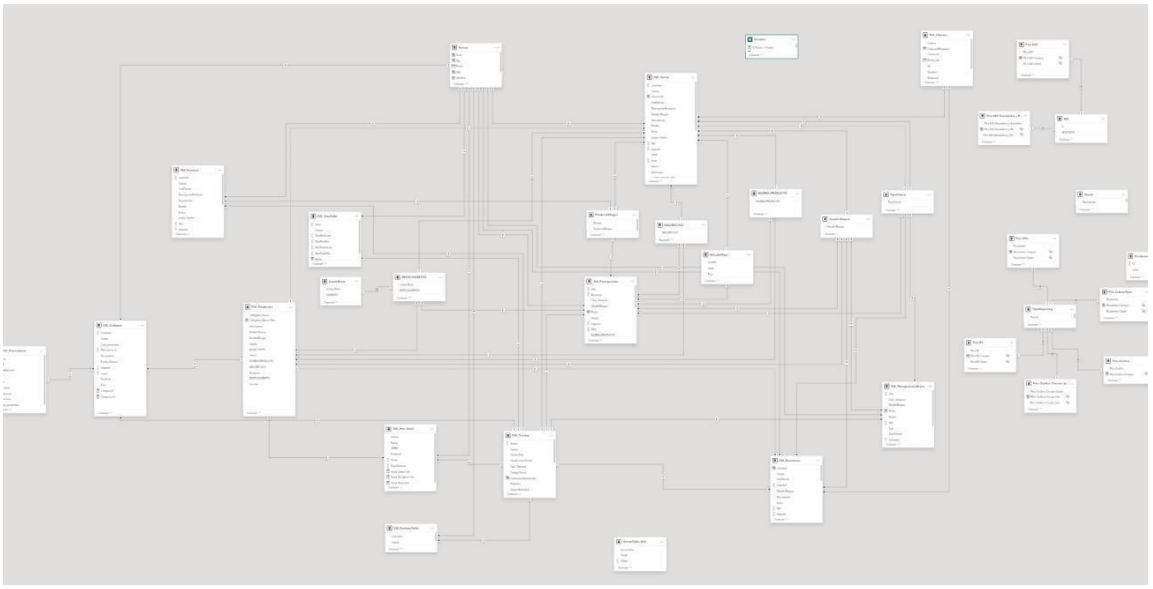


Figure 4: Entire data model. Own Source

As it is seen, this data model is very complex, and tables and relations are not seen clearly. Therefore, in the next image we can see an extraction of the most important tables and how they are related.

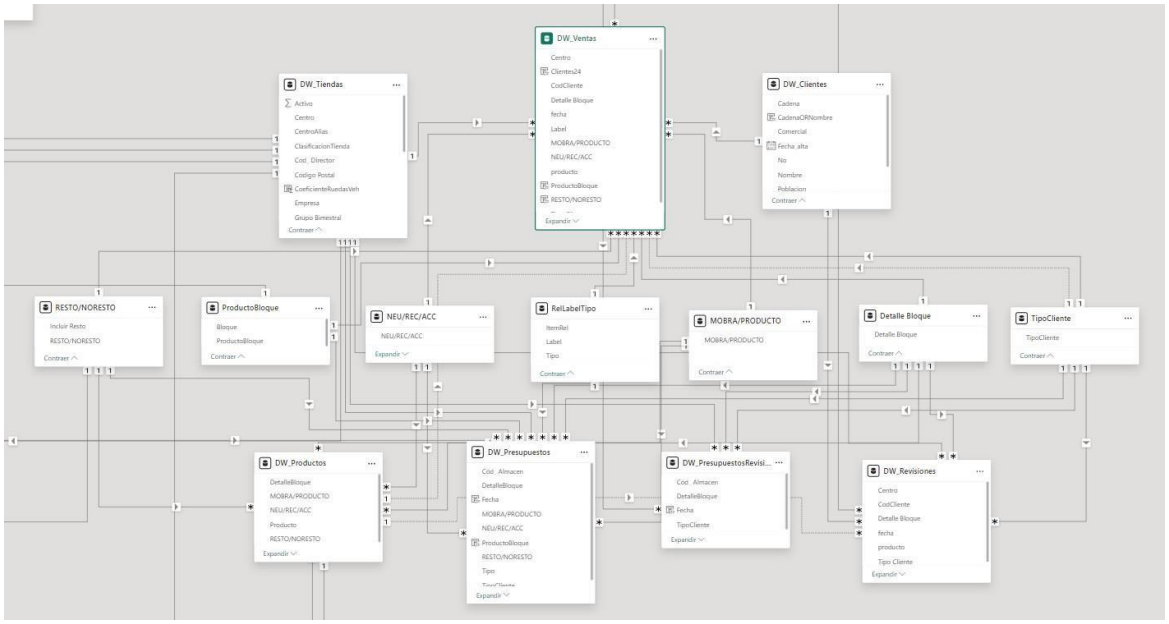


Figure 5: Relations between most used tables. Own source

As explained before, the main tables such as Sales, Budget, … are related using intermediate tables (the ones in the centre). These tables have the finality of filtering the data by the different possible parameters. The structure of these tables is very simple, the following is an example of the labour/product one:

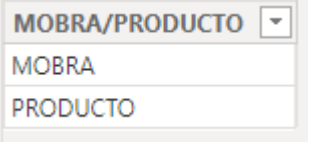


Figure 6: Table used for segmentations. Own source

The values of this table are related to the columns of the main tables with the same name. This is how we can have a one-to-many relation and effectively apply the segmentations.

In addition to these kinds of tables there are others used for format issues.

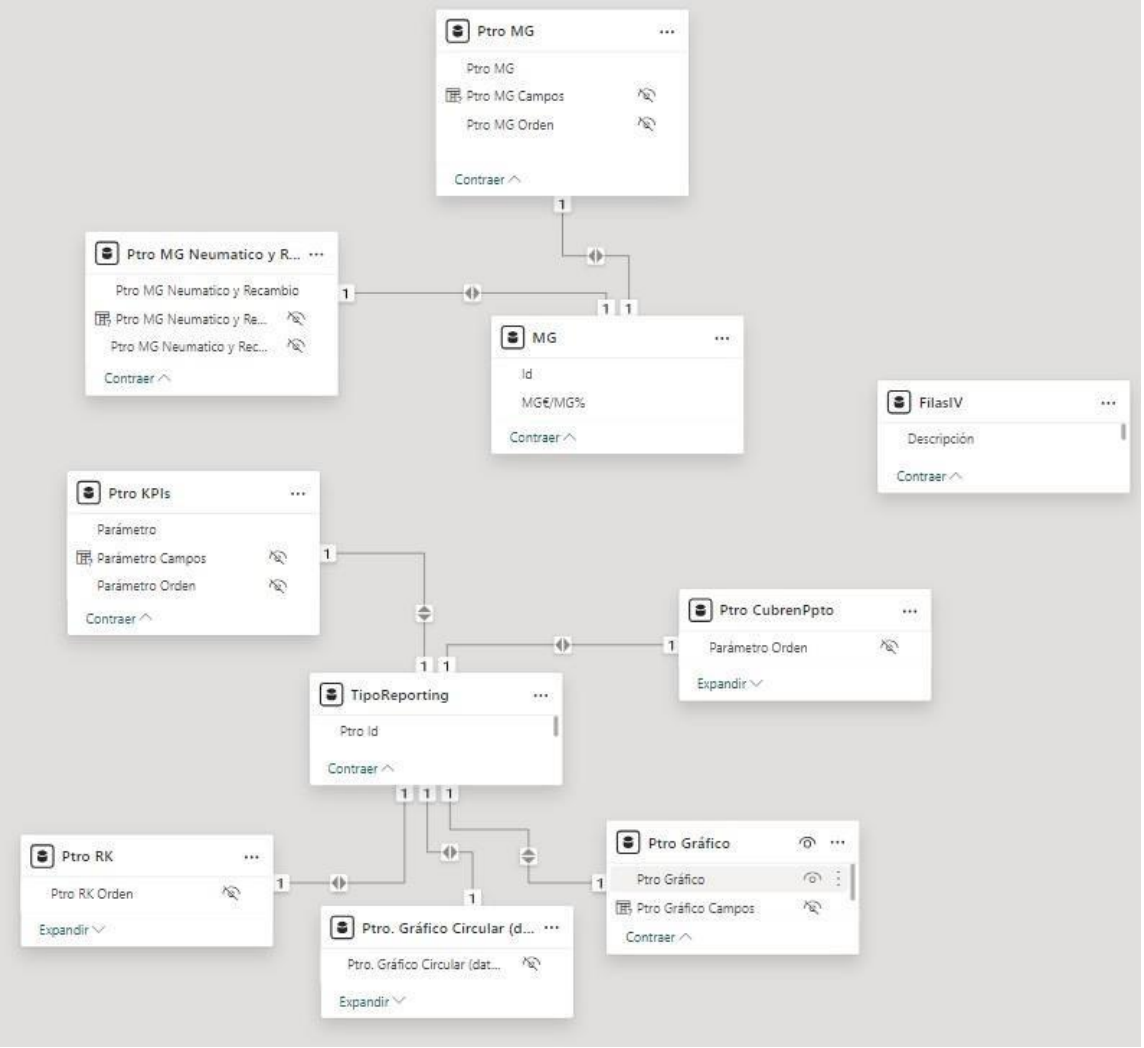


Figure 7: Tables used for the report's format. Own source

These ones are the so-called parameters, in addition to manually created tables. Parameters are used for showing a specific data or another depending on the kind of information the viewer wants to see. The structure of the parameters is the one shown in the picture below:

```
Ptro KPIs = {
    ("2024", NAMEOF('Medidas'[Ptro Neum Cant aCt]), 0),
    ("Presupuesto", NAMEOF('Medidas'[Ptro Neum Ppto aCt]), 1),
    ("24vsPpto", NAMEOF('Medidas'[Ptro Neum Consec aCt]), 2),
    ("2023", NAMEOF('Medidas'[Ptro Neum Cant aNt]), 3),
    ("ActvsAnt", NAMEOF('Medidas'[Ptro Neum ActvsAnt]), 4),

    ("2024", NAMEOF('Medidas'[Importe Acum aCt]), 10),
    ("Presupuesto", NAMEOF('Medidas'[Presupuesto aCt Acum €]), 11),
    ("24vsPpto", NAMEOF('Medidas'[Consec Acum € aCt]), 12),
    ("2023", NAMEOF('Medidas'[Importe Acum aNt]), 13),
    ("ActvsAnt", NAMEOF('Medidas'[ActvsAnt Acum €]), 14),

    ("2024", NAMEOF('Medidas'[Ptro Rec Imp aCt]), 20),
    ("Presupuesto", NAMEOF('Medidas'[Ptro Rec Ppto aCt]), 21),
    ("24vsPpto", NAMEOF('Medidas'[Ptro Rec Consec aCt]), 22),
    ("2023", NAMEOF('Medidas'[Ptro Rec Imp aNt]), 23),
    ("ActvsAnt", NAMEOF('Medidas'[Ptro Rec ActvsAnt]), 24),

    ("2024", NAMEOF('Medidas'[Revisiones Ud Acum aCt]), 30),
    ("Presupuesto", NAMEOF('Medidas'[Presupuesto aCt Acum Revisiones]), 31),
    ("24vsPpto", NAMEOF('Medidas'[Consec Acum Rev aCt]), 32),
    ("2023", NAMEOF('Medidas'[Revisiones Ud Acum aNt]), 33),
    ("ActvsAnt", NAMEOF('Medidas'[ActvsAnt Rev]), 34)
}
```

Figure 8: Parameter's structure. Own source

As we can see, each parameter has a name, a measure and an identifier. So as to group the different calculations by the type of reporting they refer to, we have built an annex table and related it with the parameter using the identifier. This table is used as a segmentation and part of its content is the following one:

TipoReporting	Ptro Id
NEUMÁTICO	0
NEUMÁTICO	1
NEUMÁTICO	2
NEUMÁTICO	3
NEUMÁTICO	4
NEUMÁTICO	5
NEUMÁTICO	6
NEUMÁTICO	7
NEUMÁTICO	8
NEUMÁTICO	9
NEU+REC	10
NEU+REC	11
NEU+REC	12
NEU+REC	13
NEU+REC	14
NEU+REC	15
NEU+REC	16
NEU+REC	17
NEU+REC	18
NEU+REC	19
RECAMBIO	20
RECAMBIO	21
RECAMBIO	22

Figure 9: Parameter segmentation table. Own source

4.6.3. Data Visual Representation

4.6.3.1. Information to be shown

As it has been remarked all along the document, the development of a project like this one requires a complete understanding of the company's business scope and functioning, being able to know which information must be shown to be analysed posteriorly.

To do this task properly we did the following actions:

- We engaged different meetings with the heads of departments with the finality to know the KPIs analysed by them and the new ones they wanted to incorporate.
- We assisted to one meeting of each type. With this action we could look at what they put more emphasis and it also served as a learning tool for knowing how decisions are taken in such a big company.
- We revised other reports and previous reunions to ensure that there were no other aspects left out by us.

After this task of detecting all the valuable information, this had to be calculated. DAX allows us to do this task.

4.6.3.2. DAX (Data Analysis Expressions)

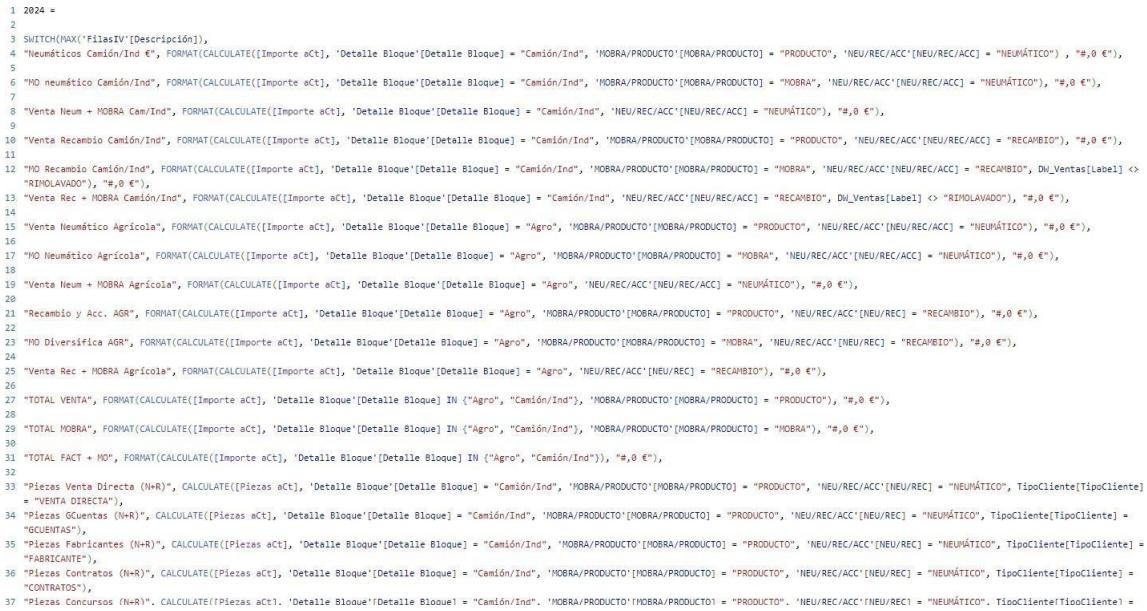
Data Analysis Expressions (DAX) is a formula language used in Power BI, Excel Power Pivot, and SQL Server Analysis Services (SSAS) Tabular mode for data analysis and calculations. DAX is a powerful tool for performing data analysis and creating advanced calculations, enabling users to derive insights and make data-driven decisions. Some key points about DAX are:

- *Syntax:* It has a similar syntax to Excel formulas but is more powerful and designed for complex analysis tasks.
- *Functions:* DAX provides a wide range of functions for different purposes such as aggregation, filtering, date and time calculations, text manipulation, and more. The following is a list of the most important/used functions in DAX.
    - CALCULATE(<expression> [, <filter1> [, <filter2> [, …]]]): Evaluates an expression in a modified filter context. It is certainly, if not the most, one of the most used functions in DAX, as it allows to do any calculation applying all the necessary filters.
    - ALLCROSSFILTERED(<table>): Clear all filters which are applied to a table.
    - MAX|MIN(<column>): Returns the largest/lowest value in a column, or between two scalar expressions.
    - SUM(<column>): Adds all the numbers in a column.
    - CALENDAR(<start_date>, <end_date>): Returns a table with a single column named "Date" that contains a contiguous set of dates. The range of dates is from the specified start date to the specified end date, inclusive of those two dates. This function is very useful for creating the Dates table, which allows us to classify all the data by dates (this includes month, year…).
    - TODAY(): Returns the current date.
    - YEAR|MONTH(<date>): Returns the year or month of a date as an integer.
    - FILTER(<table>,<filter>): Returns a table that represents a subset of another table or expression.

        - ALLCROSSFILTERED(<table>): Clear all filters which are applied to a table.
        - IF(<logical_test>, <value_if_true>[, <value_if_false>]): Checks a condition, and returns one value when it's TRUE, otherwise it returns a second value.
- *Calculated Columns:* These columns are derived from existing data in the data model and can perform calculations based on that data. They are useful for getting information that cannot be obtained directly via SQL queries, depends on the value of other columns and is static.
- *Measures:* Measures are calculations performed dynamically on the data displayed in visualizations, often based on filters applied to the data. Measures are used to create Key Performance Indicators (KPIs) and other aggregated values. They are particularly useful for creating dynamic and interactive reports. In our project there has been a multitude of measures created, for getting different KPIs. The structure followed has consisted in calculating in a single measure the general value for a KPI and then concreting more in other measures. For instance, for income there is a measure that calculates all the income of the Sales table. Then, there are others that, using the previously mentioned CALCULATE function, call the value of this measure adding a filter for only getting the value of the current year sales. This way, if any changes have to be applied, they only are applied in the general measure, which saves so much time for corrections.

  Note that measures are dynamic, so their value changes based on the filters applied. This makes the report so much interactive.
- *Relationships:* DAX leverages relationships established between tables in the data model. These relationships define how tables are connected and allow DAX calculations to aggregate data across related tables.
- *Error handling:* DAX includes functions for error handling, such as IFERROR and ISERROR, to manage potential errors in calculations.

An example of a complex used measure is the following one:



```
2024 =

SWITCH(MAX('FilasIV'[Descripción]),
"Neumáticos Camión/Ind €", FORMAT(CALCULATE([Importe aCt], 'Detalle Bloque'[Detalle Bloque] = "Camión/Ind", 'MOBRA/PRODUCTO'[MOBRA/PRODUCTO] = "PRODUCTO", 'NEU/REC/ACC'[NEU/REC/ACC] = "NEUMÁTICO") , "#,0 €"),

"MO neumático Camión/Ind", FORMAT(CALCULATE([Importe aCt], 'Detalle Bloque'[Detalle Bloque] = "Camión/Ind", 'MOBRA/PRODUCTO'[MOBRA/PRODUCTO] = "MOBRA", 'NEU/REC/ACC'[NEU/REC/ACC] = "NEUMÁTICO"), "#,0 €"),

"Venta Neum + MOBRA Cam/Ind", FORMAT(CALCULATE([Importe aCt], 'Detalle Bloque'[Detalle Bloque] = "Camión/Ind", 'NEU/REC/ACC'[NEU/REC/ACC] = "NEUMÁTICO"), "#,0 €"),

"Venta Recambio Camión/Ind", FORMAT(CALCULATE([Importe aCt], 'Detalle Bloque'[Detalle Bloque] = "Camión/Ind", 'MOBRA/PRODUCTO'[MOBRA/PRODUCTO] = "PRODUCTO", 'NEU/REC/ACC'[NEU/REC/ACC] = "RECAMBIO"), "#,0 €"),

"MO Recambio Camión/Ind", FORMAT(CALCULATE([Importe aCt], 'Detalle Bloque'[Detalle Bloque] = "Camión/Ind", 'MOBRA/PRODUCTO'[MOBRA/PRODUCTO] = "MOBRA", 'NEU/REC/ACC'[NEU/REC/ACC] = "RECAMBIO", DW_Ventas[Label] <>
"RIMOLAVADO"), "#,0 €"),
"Venta Rec + MOBRA Camión/Ind", FORMAT(CALCULATE([Importe aCt], 'Detalle Bloque'[Detalle Bloque] = "Camión/Ind", 'NEU/REC/ACC'[NEU/REC/ACC] = "RECAMBIO", DW_Ventas[Label] <> "RIMOLAVADO"), "#,0 €"),

"Venta Neumático Agrícola", FORMAT(CALCULATE([Importe aCt], 'Detalle Bloque'[Detalle Bloque] = "Agro", 'MOBRA/PRODUCTO'[MOBRA/PRODUCTO] = "PRODUCTO", 'NEU/REC/ACC'[NEU/REC/ACC] = "NEUMÁTICO"), "#,0 €"),

"MO Neumático Agrícola", FORMAT(CALCULATE([Importe aCt], 'Detalle Bloque'[Detalle Bloque] = "Agro", 'MOBRA/PRODUCTO'[MOBRA/PRODUCTO] = "MOBRA", 'NEU/REC/ACC'[NEU/REC/ACC] = "NEUMÁTICO"), "#,0 €"),

"Venta Neum + MOBRA Agrícola", FORMAT(CALCULATE([Importe aCt], 'Detalle Bloque'[Detalle Bloque] = "Agro", 'NEU/REC/ACC'[NEU/REC/ACC] = "NEUMÁTICO"), "#,0 €"),

"Recambio y Acc. AGR", FORMAT(CALCULATE([Importe aCt], 'Detalle Bloque'[Detalle Bloque] = "Agro", 'MOBRA/PRODUCTO'[MOBRA/PRODUCTO] = "PRODUCTO", 'NEU/REC/ACC'[NEU/REC] = "RECAMBIO"), "#,0 €"),

"MO Diversifica AGR", FORMAT(CALCULATE([Importe aCt], 'Detalle Bloque'[Detalle Bloque] = "Agro", 'MOBRA/PRODUCTO'[MOBRA/PRODUCTO] = "MOBRA", 'NEU/REC/ACC'[NEU/REC] = "RECAMBIO"), "#,0 €"),

"Venta Rec + MOBRA Agrícola", FORMAT(CALCULATE([Importe aCt], 'Detalle Bloque'[Detalle Bloque] = "Agro", 'NEU/REC/ACC'[NEU/REC] = "RECAMBIO"), "#,0 €"),

"TOTAL VENTA", FORMAT(CALCULATE([Importe aCt], 'Detalle Bloque'[Detalle Bloque] IN {"Agro", "Camión/Ind"}, 'MOBRA/PRODUCTO'[MOBRA/PRODUCTO] = "PRODUCTO"), "#,0 €"),

"TOTAL MOBRA", FORMAT(CALCULATE([Importe aCt], 'Detalle Bloque'[Detalle Bloque] IN {"Agro", "Camión/Ind"}, 'MOBRA/PRODUCTO'[MOBRA/PRODUCTO] = "MOBRA"), "#,0 €"),

"TOTAL FACT + MO", FORMAT(CALCULATE([Importe aCt], 'Detalle Bloque'[Detalle Bloque] IN {"Agro", "Camión/Ind"}), "#,0 €"),

"Piezas Venta Directa (N+R)", CALCULATE([Piezas aCt], 'Detalle Bloque'[Detalle Bloque] = "Camión/Ind", 'MOBRA/PRODUCTO'[MOBRA/PRODUCTO] = "PRODUCTO", 'NEU/REC/ACC'[NEU/REC] = "NEUMÁTICO", TipoCliente[TipoCliente]
= "VENTA DIRECTA"),
"Piezas GCuentas (N+R)", CALCULATE([Piezas aCt], 'Detalle Bloque'[Detalle Bloque] = "Camión/Ind", 'MOBRA/PRODUCTO'[MOBRA/PRODUCTO] = "PRODUCTO", 'NEU/REC/ACC'[NEU/REC] = "NEUMÁTICO", TipoCliente[TipoCliente] =
"GCUENTAS"),
"Piezas Fabricantes (N+R)", CALCULATE([Piezas aCt], 'Detalle Bloque'[Detalle Bloque] = "Camión/Ind", 'MOBRA/PRODUCTO'[MOBRA/PRODUCTO] = "PRODUCTO", 'NEU/REC/ACC'[NEU/REC] = "NEUMÁTICO", TipoCliente[TipoCliente] =
"FABRICANTE"),
"Piezas Contratos (N+R)", CALCULATE([Piezas aCt], 'Detalle Bloque'[Detalle Bloque] = "Camión/Ind", 'MOBRA/PRODUCTO'[MOBRA/PRODUCTO] = "PRODUCTO", 'NEU/REC/ACC'[NEU/REC] = "NEUMÁTICO", TipoCliente[TipoCliente] =
"CONTRATOS"),
"Piezas Concursos (N+R)", CALCULATE([Piezas aCt], 'Detalle Bloque'[Detalle Bloque] = "Camión/Ind", 'MOBRA/PRODUCTO'[MOBRA/PRODUCTO] = "PRODUCTO", 'NEU/REC/ACC'[NEU/REC] = "NEUMÁTICO", TipoCliente[TipoCliente] =
```

Figure 10: Complex measure. Own source

In this measure, using the SWITCH operator, depending on the item that has to be calculated, a certain measure is executed with some filters applied. The use of the FORMAT operator is necessary for applying the "€" format when the value calculated refers to an income related KPI.

As said before, this measure is an example of a complex one. Most of them include easy calculations. The following calculate the number of products sold:

```
CantidadUds = CALCULATE(SUM(DW_Ventas[Cantidad]))
```

Figure 11: Total number of products sold (all years). Own source

```
Cantidad aCt = CALCULATE([CantidadUds], Fechas[N] = 0)
```

Figure 12: Total number of products sold (this year). Own source

To give an idea of how powerful are measures in Power BI, we include another type of measure used in the project. This one calculates different values and concatenates them representing them as a string of text. Variables are used for intermediate calculations.

```
Cubren Ppto Neumaticos =
Var cubren =
CALCULATE(
    DISTINCTCOUNT(DW_Tiendas[Centro]),
    FILTER(
        DW_Tiendas,
        [Consec Acum Ud aCt] >= 1
    )
) + 0

var total = DISTINCTCOUNT(DW_Tiendas[Centro]) + 0
var porcentaje = ROUND((cubren / total) * 100, 0) + 0

return
porcentaje & "% (" & cubren & "/" & total & ")"
```

Figure 13: Mixed integer and text format measure. Own source

4.6.3.3. Power BI Visual Objects

After creating all the necessary calculated columns and measures (there is a total of 20 calculated columns and there are around 450 measures, most of them contain calculations, but some other ones are created due to format issues), we had to design a way to represent all this KPIs in a visual way, allowing the users to interact with them and choosing the nest ones for carrying the analysis on the company's economic activity.

Power BI provides multiple visual objects to represent this information. These visuals are used to present data in a meaningful and interactive way, enabling users to gain insights and make data-driven decisions. Here are some important aspects of this Power BI feature:

- *Types of visuals:* Power BI offers a wide range of visualizations to suit different data analysis needs. Some common types of visuals include bar charts, line charts, pie charts, scatter plots, maps, tables, matrices, gauges, cards, and custom visuals. The last ones are created by third-party developers or users. These custom visuals extend the capabilities of Power BI by offering specialized visualizations or unique ways of presenting data.

  The most used types of visuals have been charts, tables, matrices and cards, as they represent the data in a clear and visual way, and users can draw conclusions more rapidly.
- *Interactivity:* Power BI visual objects are highly interactive, allowing users to interact with the data in various ways. Users can drill down into details, apply

filters, highlight data points, sort and filter data dynamically, and perform other actions to explore the data and gain deeper insights.

- *Formatting options:* Users can customize the appearance of visuals to match their preferences or corporate branding guidelines. They can change colours, fonts, labels, axes, backgrounds, borders, and other visual elements.
- *Responsive design:* Power BI visuals are designed to be responsive, meaning they adapt to different screen sizes and orientations. This ensures that reports and dashboards remain usable and visually appealing across various devices, including desktops, laptops, tablets, and smartphones.

4.6.4. Report Design and Structure

The report must have a solid and well-designed structure. UI design principles play an important role in this phase. Those principles that need to be accomplished are:

- *Simplicity:* Aims to minimize complexity and cognitive load for users. This can be reached by keeping interfaces clean, uncluttered and easy to understand.
- *Consistency:* This is one of the most important principles to be achieved. Elements and interactions across different screens and sections of a digital product are cohesive and predictable.
- *Visual hierarchy:* Refers to the arrangement of elements on a screen to prioritize their importance and guide users' attention. We must consider that users visualize the information from left to right and from the top of the screen to the bottom. Therefore, the most important information must be placed on the top-left of the screen and the less important information, on the right-bottom screen.
- *Clarity:* Focuses on presenting information and content in a clear and easily understandable manner.
- *Scalability and Adaptability:* Involves designing interfaces that can seamlessly adapt to different screen sizes, resolutions, and orientations. A mobile interface for the report must be created to ensure the accomplishment of this principle.

There exist other UI principles that are not applicable to our project, though they are as important as the mentioned ones.

After having a clear idea of how the report should be, the data it had to include and how this data should be represented, we developed a prototype with the first version of the report, organising the data by the principles defined above.

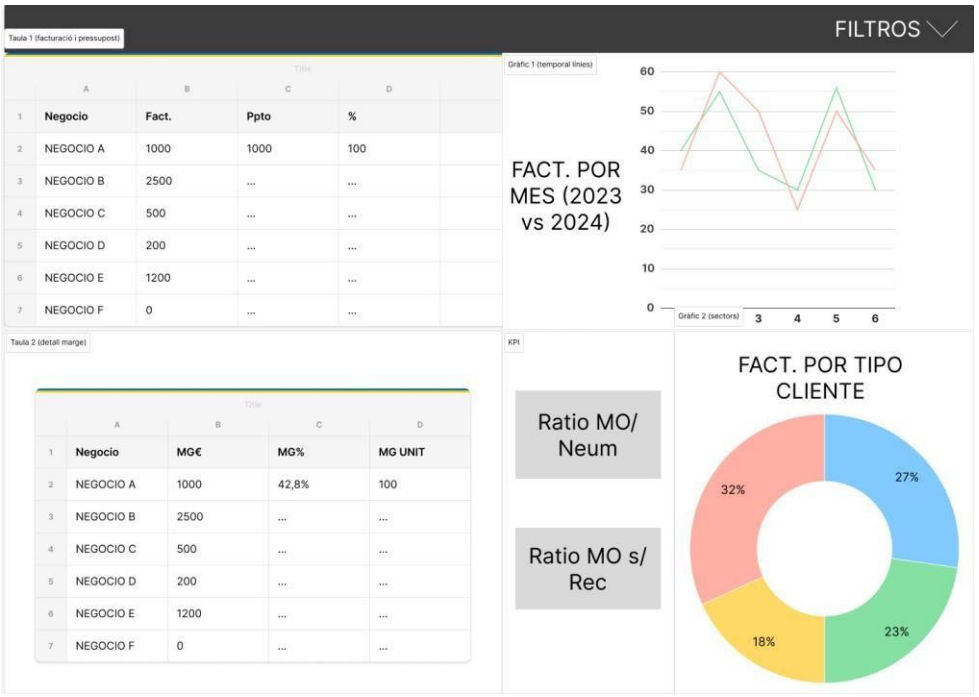


Figure 14: Prototype of the report. *Own source.*

Considering all the feedback received by the future users of the report, we designed the current active version of the report.

This version included some visual and interactive aspects which were not included in the prototype, but they were well-received by the users, which rapidly became familiar with this report and agreed that it was intuitive to use and easy to understand, as well as useful for the task of analysing the economic activity of the company.

Here are the different pages of the report. Note that data values have been modified due to company privacy terms and conditions.

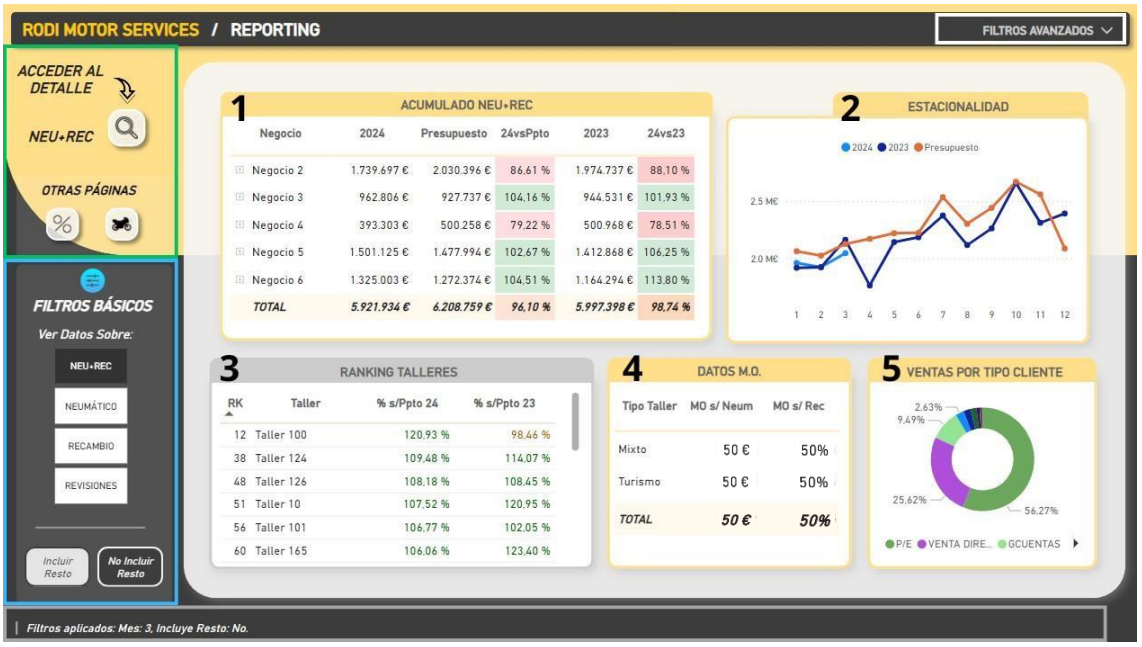


Figure 15: Report's main page adding four rectangles (white, blue, green and grey) remarking information further explained. Own source.

The main page includes the main KPIs by type of service/product offered.

The table 1 includes the accumulated invoice of the current year and month by business and centre compared to the budget and the year before. To make the data more visual a conditional colour has been added to the columns that include the comparations depending on the level of consecution of the objective achieved.

The graphic 2 contains the invoice reached every month, compared to the one of the previous year and the budget. With this graphic the user can easily analyse in which months the budget has been achieved and if the number of sales has been superior to the one of the year before or not.

Table 3 is a ranking of the centres ordered by the percentage of consecution achieved the current year. This way users can see which centres are going better and will earn more bonuses at the end of the year.

KPI 4 is a summary of the labour per tyre and the percentage of replacement income that refers to labour. This data is given by the type of centre.

Graphic 5 includes the percentage of sales invoiced by client type. This way we can see rapidly which clients give us more profit.

This report has been created with the intention of being highly interactive, so that multitude of data can be obtained applying filters. By clicking to the zone remarked with the white rectangle a filter pop-up is opened:

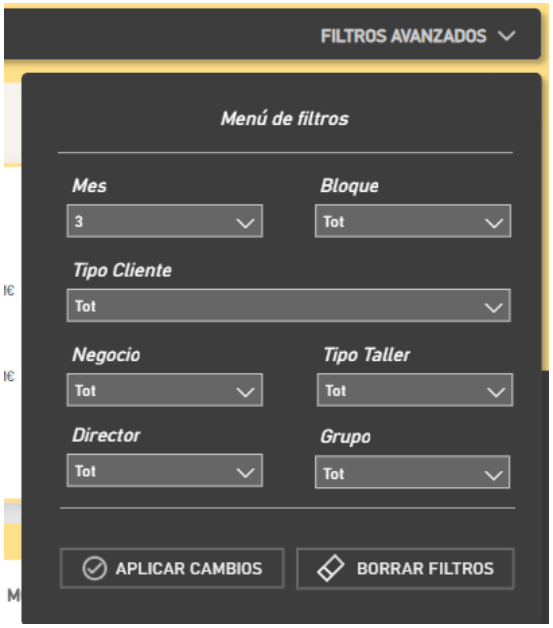


Figure 16: Report's filter window. Own source

The appliable filters of this sections are: month, block (this is light vehicles, motorcycles, truck or agricultural vehicles), type of client, business, type of centre, director of centre and group (centres are grouped depending on the zone they are placed and type of centre).

They grey rectangle remarks a zone used for showing to the user which filters have been applied. With this information the user knows any time what segmentations are active in the data they are seeing.

The blue rectangle signals the zone of basic filters. They are very accessible for the user because they are intended to be changed more frequently than the other ones. In the first segmentation the user can choose which type of product wants to see the data from. So, when one type of product is selected, all the data from all the tables and graphics previously described is shown over this type of product. This is done by using parameters, which have been explained in previous sections.

Finally, the green rectangle is the zone that gives access to other pages of the report that contain more detailed information of the type of product selected (icon with a magnifying glass). The other two icons give access to a page that contains different ratios and specifical information about motorcycle block.

The following is an example of the detail page of tyre plus replacement products:

Figure 17: Tyre plus replacement detail page. Own source

As we can see the structure is similar to the principal page. On the left-top of the main window we have the main KPIs. Then there is a graphic of the distribution of sales between tyre and replacement by business (note that there is one business which is not present, due to the low impact it has on the company). On the lower part there is all the margin information by business. It is possible to have this value in percentage or in €. Finally, we have the top 10 clients (major accounts or contests) who have spent more in our company, added to the percentage they represent among all the others.

There are other detail pages, but as they are like this one (one or two visual objects are changed), they have not been added to this thesis.

The ratios page has a finality of data extraction and analysis, and it is not as interactive as the other ones. This is the aspect of one of the 8 different ratios that this page contains:

MO s/ Piezas	2024	2023	Dif
Taller 10	50 €	50 €	0 €
Taller 100	50 €	50 €	0 €
Taller 101	50 €	50 €	0 €
Taller 102	50 €	50 €	0 €
TOTAL	***50 €***	***50 €***	***0 €***

Figure 18: Ratio example. Own Source

A remarkable characteristic of the report is that depending on the business selected in the filters, the background, and visual objects colour change, giving them the brand colour of the current business.

Although most of the users are familiar with the use of Power BI, we had to arrange a few meetings for showing how this report works and how they could take more benefit of its use. This small training phase did not take much time.

4.7. Translating business theory to project practise

Working on this project has required applying to practise some of the knowledge related to business acquired during all the five years of the degree. In addition, much knowledge related to this aspect has been obtained.

It is essential to comprehend the underlying business model to effectively showcase project-relevant information. This involves grasping the complexities of the operational framework and strategic objectives that drive the project forward.

4.7.1. Projections for closure and budget assignment

Regarding projections for closure and budget assignment both methods based on the percentage of monthly expenditure and methods based on moving averages have been used in the implementation of the project.

The first one tracks spending as a percentage of a monthly budget. Each cost centre gets a slice of the annual budget, divided by twelve for each month. By comparing actual spending to this monthly allocation, you can calculate a percentage spent and identify

areas that consistently go over or under budget. It's a simple way to monitor spending trends and find room for cost reduction.

Moving average methods track spending trends by averaging expenses over a set period.

- We use the previous year's actual spending and budget for the current year.
- We calculate the change in spending between years for the period we are studying.
- Based on the previous year's closing balance and this spending change, we forecast the closing balance for the current year.
- Finally, we compare the budgeted amount to the forecasted closing balance to see if there's a potential budget surplus or shortfall.

Each method has its pros and cons, taking it into account, both were added in the project, so that it could be evaluated in different ways.

4.7.2. Shown data according to different factors

Data can be shown in different ways and measurement units according to the product that is being analysed and the indicator used.

4.7.2.1. Income

Income can be displayed as units or as euros. We have seen that when seeing tyres data, the best way of showing income, budget and consecution is via units sold. This gives us more information than euros, as when analysing the same type of vehicles, the price of each tyre is similar to the other ones. It’s when analysing the profit per brand when we will show data like unit margin, with the purpose of enhance the sales of the ones with the higher margin per unit.

Replacement, though, cannot be analysed with units. The main reason is that it is not logical to analyse a product that has a very changeable price by units. Therefore, the income of kind of product is displayed mainly in euros. Another fact highly analysed on replacement is the percentage of labour that refers to the income, as it is the factor that provides the most benefit to the company, rather than the product itself.

4.7.2.2. Margin

The difference between income and costs, known as margin, can also be seen in different ways. The most common one is via euros. This one gives us a quick overview of which products produce more margin. Nevertheless, it is also interesting to analyse margins in

percentages. With this way it is easier to see the percentage of the price that has been a net profit for the company. With this variation we can compare the margin taken from cheap products with the one of more expensive products.

Finally, as said before, basically with tyres it is very common and useful to analyse the margin per unit, being able to compare between brands and different types of tyres.

4.7.2.3. Clients

Depending on the type of client analysed, the shown data is also relevant.

Therefore, when we see data based on particulars or companies, mainly global data is seen, without entering on concrete clients.

However, with major accounts and contests we want to see which ones have spent more in our company, and its weight in the total income. That's why we have built special rankings and KPIs in our project, so that customer acquisition and preservation can also be worked with it.

4.7.3. Working on achieving business maturity

Another thing learned with the development of this project is the process of establishing a new business in a new area.

We have seen that, with the centres recently acquired in zones where there is not presence of our company, a certain strategy has been followed to reach to new clients and gain market share.

As an example, it is evident in figure 17 that in consolidated businesses, replacement plays a more important role than in the new ones, in which tyre change leads by far the percentage of income obtained. The objective is to offer more services of replacement, which give the company more benefits, but this can not be implemented in the early stages of a business. A certain level of maturity has to be reached, and the process might take long, but once the objective is accomplished, the percentage of spare parts services offered does not decrease at all.

4.8. Performance tests

Once the report is created, it is necessary to check if its performance satisfies the previously defined requirements.

There is a tool provided by Power BI that allows us to know how much time each visual object needs to be loaded. This is the performance analyser. The results obtained have been the following ones:

Figure 19: Performance analyser execution. Own source

Observing the time of load of each visual object we can confirm that, as demanded, every single object loads in less than 3 seconds.

4.9. Report publication

When having confirmed that the report contains valid data and has good performance numbers, it can be published to the Power BI cloud service, provided by Microsoft, so it can be available for all users.

An automatic updating of the data set has been scheduled every day, so that each KPI is calculated up to date.

4.10. Encountered problems and difficulties

As it was expected, during the process of developing the project, many problems and difficulties had to be faced. Here is a list of the ones that were more laborious to solve:

1. Since the job is performed on work that is already done, this involves understanding everything that is done in order to advance further on the subject. There was a long learning curve.
2. Knowing which labels to take. At first, it was thought two tables were needed, one with sales and labels (where there would supposedly be duplicate records) and one with unique sales. We saw that it was not necessary since there are one kind of labels for income items and others for other parts (control panel), so if we only take the labels that interest us, there won't be any duplicates.
3. This also led to the false assumption that a join could not be made directly from sales table to the one created for the calculation of the “Master”. For this reason, the LabelSales table was created, with which we had sales with labels, for classification. A process was also made to update this table, with auxiliary tables of criteria, configuration, etc. *(see section 4.5.2.1.)*

   By joining sales and this “Master” table we could make good optimization because that way it is not necessary to execute the LabelSales update script. *(see section 4.5.2.2.)*
4. Storage problems. Columns and tables had to be deleted to make reports smaller. We went from 1GB to 200-300 MB, a reduction of about 75%.
5. In a more advanced phase of the project, we wanted to do a classification that has not yet been done. We had to manage to classify budgets, sales, purchases, stock... by block, product type, customer type, etc. What we ended up doing was creating new tables with additional columns. With Budget it was relatively easy because

there is the group and type columns with which we could more or less base our classification. With Products, we had to look at the criteria of the different labels and see which products are put in one label and which in another, in order to classify them. It was a more laborious process. *(see section 4.5.2.2.)*

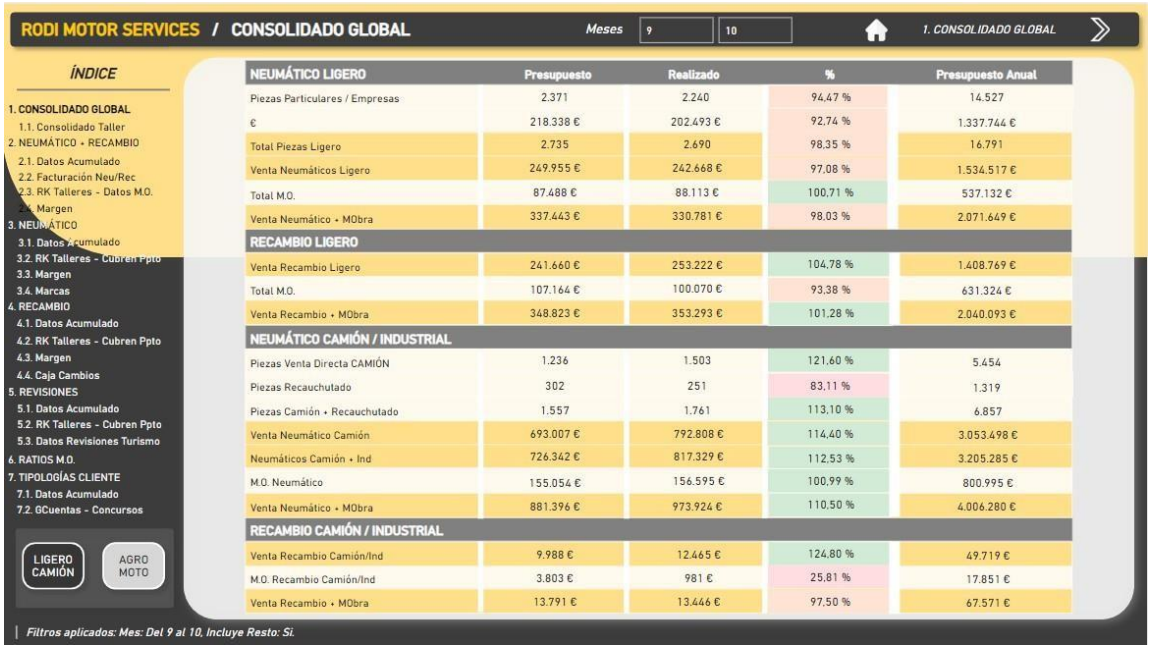
6. The stock value in the Stock table has a photo of the concrete day, so we cannot sum the amount, but take the value of the last possible day.
7. To have the data of revisions and ratios an additional table had to be added, as some data of this type is not present in the "Master". This meant implementing a new SQL query which is quite complex, as gets information from different tables and unifies it with the UNION ALL operator.

4.11. Other applications of the project

In Power BI, there is the option for multiple reports to share the same set of data. Thus, a single set of data can have more than one graphical interface. Why might this be useful? Consider that the data in the reports needs to be updated every day. By completing the project, it would be possible to have all the reports updated with just one update per day (rather than one update per report).

Taking advantage of this feature we have designed more than one visualization for the project. In parallel to the development of the main project we made a test to integrate this set of data to another report used by the company. We could successfully do this task, so we saw that the project was progressing really well.

Also related to this fact we implemented and designed another visualization of the project. This one had the mission to replace the actual format of the bimonthly meetings. A simulation of a Power Point presentation was created using Power BI, with all the information seen in this kind of meetings. There is an example of what would the equivalent of a Power Point slide:



NEUMÁTICO LIGERO	Presupuesto	Realizado	%	Presupuesto Anual
Piezas Particulares / Empresas	2.371	2.240	94,47 %	14.527
€	218.338 €	202.493 €	92,74 %	1.337.744 €
Total Piezas Ligero	2.735	2.690	98,35 %	16.791
Venta Neumáticos Ligero	249.955 €	242.668 €	97,08 %	1.534.517 €
Total M.O.	87.488 €	88.113 €	100,71 %	537.132 €
Venta Neumático + MObra	337.443 €	330.781 €	98,03 %	2.071.649 €
RECAMBIO LIGERO				
Venta Recambio Ligero	241.660 €	253.222 €	104,78 %	1.408.769 €
Total M.O.	107.164 €	100.070 €	93,38 %	631.324 €
Venta Recambio + MObra	348.823 €	353.293 €	101,28 %	2.040.093 €
NEUMÁTICO CAMIÓN / INDUSTRIAL				
Piezas Venta Directa CAMIÓN	1.236	1.503	121,60 %	5.454
Piezas Recauchutado	302	251	83,11 %	1.319
Piezas Camión + Recauchutado	1.557	1.761	113,10 %	6.857
Venta Neumático Camión	693.007 €	792.808 €	114,40 %	3.053.498 €
Neumáticos Camión + Ind	726.342 €	817.329 €	112,53 %	3.205.285 €
M.O. Neumático	155.054 €	156.595 €	100,99 %	800.995 €
Venta Neumático + MObra	881.396 €	973.924 €	110,50 %	4.006.280 €
RECAMBIO CAMIÓN / INDUSTRIAL				
Venta Recambio Camión/Ind	9.988 €	12.465 €	124,80 %	49.719 €
M.O. Recambio Camión/Ind	3.803 €	981 €	25,81 %	17.851 €
Venta Recambio + MObra	13.791 €	13.446 €	97,50 %	67.571 €

Figure 20: Slide of the report designed for reunions. Own source

5. Result evaluation

Upon completion of the project, it is interesting to do an evaluation of the job done, looking aspect by aspect whether the tasks planned have been completed successfully or not.

5.1. Data set evaluation

The building of a unique and consistent data set that contains all the information related to the company's economic activity has provided us a reliable source of data that can be used in multiple projects and does not lead to conflicts among different reports.

This task has been possible thanks to the previously identified company's data structure and the using of its Datawarehouse, as well as to Heidi SQL, the tool provided for extracting data.

Power BI has also given us the opportunity to add additional columns and tables to the dataset, as well as establishing relations between tables. This has been a key aspect to complement the previously extracted data.

5.2. BI report evaluation

The design and implementation of the BI report has been received as a powerful tool by the company. Many users have agreed that it allows them to get and analyse all the desired data related to the company's economic activity, giving them the opportunity to identify and delve into the detail of each sale/budgeting problem that must be addressed.

Power BI has also provided us all the needed tools to perform this task, in terms of backend but also in terms of frontend, since a well-designed and attractive report has been built.

The importance of studying and understanding the company's business model, taking part of the meetings and stablishing reunions with the different heads of the departments has also been capital to resolve this task. Knowledge of how a company operates is essential when performing a task to analyse its economic activity.

5.3. Tasks planification evaluation

The different tasks have been carried out in accordance with their time schedule. Following this planification has given us the comfort to work without an imminent deadline, allowing us to develop the tasks with critical thinking and always choosing the allegedly best solutions.

Although in some cases those have not been the best ones, this task planification has enabled us to find another solution without having to worry about deadlines.

5.4. Functional requirements evaluation

All the functional requirements have been satisfied. This means that the report effectively represents all the data desired, with the segmentations previously required and the consistency that was demanded. Different interfaces have also been created in order to satisfy multiple needs.

The satisfaction and feedback received from users has been used as a measuring device for determining the compliance of such requirements.

5.5. Non-functional requirements evaluation

Related to non-functional requirements, all of them have also been accomplished, with one little exception. Depending on the database server speed, the report is updated in more than 30 minutes. Usually this requirement is satisfied, but some days (mostly at the beginning of the month), it takes longer than usual. Maybe this a fact than could be improved in the future.

Regarding the rest of the requirements, we can guarantee that the report is accessible and adaptable, it weights between 200-300 Mbytes and the different visual objects are represented quite fast (less than 3 seconds).

6. Conclusions

6.1. Dataset building conclusions

A critical objective of this project was the construction of a unified dataset encompassing the entirety of the company's economic activity data. This accomplishment, achieved through the strategic application of database knowledge (including SQL and data modelling principles), has yielded a dataset characterized by the following attributes:

- Multi-report Utilization: The dataset serves as a centralized source of information, readily accessible and utilizable by various reports.
- Rapid Update Cycle: The data refresh interval adheres to a strict timeframe, ensuring updates occur within 30 minutes.
- Dynamic Adaptability: The dataset demonstrates a high degree of flexibility, readily accommodating modifications to user-defined criteria.
- Elimination of Conflicting Criteria: Prior inconsistencies in criteria definitions that plagued legacy reports have been successfully eradicated.

As an ongoing effort, the integration of inventory and purchase data into the data model is currently underway. This addition will further enhance the dataset's robustness and analytical capabilities.

To sum up, the creation of a consolidated and flexible dataset has laid the foundation for the project's success. This central repository of the company's economic data ensures information is readily accessible and up to date within 30 minutes. Furthermore, the data model's design allows for easy adaptation to evolving user-defined criteria, eliminating inconsistencies that previously hampered decision-making.

6.2. BI project conclusions

This Business Intelligence (BI) project has equipped the company with a powerful tool to gain deeper control over its economic activity. This translates to improved decision-making across strategic and operational levels.

- Product Performance: The company can quickly identify its most profitable and best-selling products, while also pinpointing areas needing improvement.

- Comparative Analysis: The BI tool allows for side-by-side comparisons between current and past years' data, added to budget data. Users can define the scope of the study by selecting specific months, centres, product categories and many more filters for analysis.
- Customer Activity: The tool tracks customer activity, revealing trends in purchase behaviour and product preferences. This allows the company to identify customer churn and acquisition patterns.
- Customer Lifetime Value: By understanding customer purchase history and frequency, the company can gain valuable insights into customer lifetime value.

Ultimately, this project equips the company with the ability to analyse economic activity with real-time precision. This empowers data-driven decision-making based on factors such as sales trends, product margins, customer buying habits or performance across different centres. Its highly interactive interface also allows the user to extract a lot more information, reaching a detail level which was not available before.

Overall, this project has demonstrably improved the company's data analysis capabilities. The success can be attributed to the well-organized data set, a user-friendly and informative BI report, adherence to the project plan, and the fulfilment of both functional and non-functional requirements.

6.3. Personal conclusions

This project has been an exigent yet enriching personal challenge that lived up to my expectations. Here's what I gained from this experience:

In terms of Computer Engineering:

- Structured Approach: I implemented the project following established principles and steps of software engineering.
- Technical Implementation: The backend design leveraged database concepts and languages. The frontend design involved selecting and applying all basic principles of human-computer interaction (HCI) to obtain a good interface and ensure a good user experience.

In terms of Business and Administration:

- Business Understanding: The project allowed me to apply my Business Administration and Management skills. This involved understanding the company's operations, which is indispensable for implementing the BI system.
- Data-Driven Decision Making: I learned to audit and interpret business activity data effectively by showing indicators like KPIs.

In personal terms:

- Practical BI Experience: the project has opened up the world of Business Intelligence to me. I learned about managing and developing such projects and the relevance they have nowadays in big and successful companies. I personally see it as a very positive fact, as I have discovered what might be the future of my professional career.

Overall, this project fuelled my personal satisfaction and significantly enhanced my experience. Now I am very motivated to keep implementing these kinds of projects, deepening my knowledge of BI and improving data analysis in companies as RODI, which has given me an unbeatable opportunity that I could not reject.

7. References

Microsoft (2020). Power BI Desktop Quick Start: Learn DAX Basics. URL: https://learn.microsoft.com/es-es/power-bi/transform-model/desktop-quickstart-learn-dax-basics

Microsoft Power BI Community (n.d.). Power BI Forums. URL: https://community.fabric.microsoft.com/t5/Power-BI-forums/ct-p/powerbi

NALS Engineering (n.d.). 10 principles in UI design: Enhancing user experience through practical examples. URL: https://medium.com/@NALSengineering/10-principles-in-ui-design-enhancing-user-experience-through-practical-examples-9d519e91b515

8. Acknowledgments

I would like to express my deepest gratitude to Sergi Balagueró Oliva for his invaluable contribution to this project. Sergi not only provided the initial inspiration for the direction of this research but also shared his extensive knowledge and expertise in the field, which has been instrumental in shaping the content and methodology of this work. Without his guidance and mentorship, this project would not have been possible.

I would also like to extend my sincere appreciation to Marta Oliva Solé for her unwavering support and guidance throughout the development of this project. Marta's dedication to this endeavour, her willingness to hold regular meetings, and her insightful feedback have been invaluable in refining the ideas and ensuring the successful completion of this work.

Finally, I would like to express my gratitude to Rodi Motor Services for providing the opportunity to develop this project within their organization. Their support and resources have been indispensable in the successful completion of this work.