



CATOLICA  
CATÓLICA PORTO  
BUSINESS SCHOOL

PORTO

Business Intelligence  
Group Assignment

Vasco Moutinho 355422066  
Simão Magalhães 355422067  
Sofia Fernandes 355422038  
Tiago Silva 355422039



**Index**

Introduction .....	3
Database Design (Power BI) .....	4
SQL Language (MySQL workbench).....	6
Dashboards created in Power BI (Power Bi Report and Dashboards) .....	8
Operational Dashboard .....	9
Tactic Dashboard .....	10
Strategic Dashboard.....	11
Dashboard Q&A .....	13
Conclusion .....	15

## **Introduction**

Nowadays, especially in the business environment, it is increasingly crucial to extract maximum value from the data in the surrounding environment. This not only enables companies to maintain and establish a prominent position in the market, but also empowers them to gain a competitive advantage over their competitors.

Thus, within the scope of the Business Intelligence curricular unit, it was proposed to prepare a detailed analysis and interpretation of a database of an organization chosen by the students themselves or made available by the teachers, using SQL language and Power BI. Thus allowing the combination of theoretical and practical aspects taught in the classroom.

Therefore, the group chose to create a database for a fictitious chain of gyms called "myGYM", whose focus is on the satisfaction of its customers through the existing group classes, the individual training provided by the personal trainers (employees), the equipment available in each gym and the quality of its employees. This chain of gyms consists of 20 gyms, located in 5 different cities, namely Porto, Lisbon, Coimbra, Salamanca and Madrid.

To ensure a complete analysis, this report is organized in five parts, starting with a brief introduction, followed by the design and relational schema of the database, the execution of the SQL code for its creation, the development of the corresponding dashboards in Power BI, and finally, the most relevant and essential conclusions to assist the decision-making process.

### Database Design (Power BI)

To elaborate the present work, a business was originated related to a chain of 20 gyms, located in 5 cities, in Spain and Portugal. Therefore, we proceeded to the conceptual design of the database considering 14 entities (City, Country, Class, Class\_Equipement, Class\_Client, Employee, Gym, Client, Equipment, Gym\_Equipment, Pub, Payments, Equipment\_Training, Training), These are related to the city, country, classes, equipment used in each class, clients attending each class, employees, gyms, clients, equipment, equipment in each gym, advertisements, payments made by clients, equipment used in each training and training, respectively. It is also important to note the difference between the entities "Class" and "Training", the first referring to group classes taught by employees, and the second referring to the elaboration of an individual training plan for each client carried out by a particular employee. These entities are composed of multiple attributes and linked through various relationships, allowing the crossing of information and its use. It should also be noted that this database contains data for the last 5 years, that is, for the years 2018, 2019, 2020, 2021 and 2022.

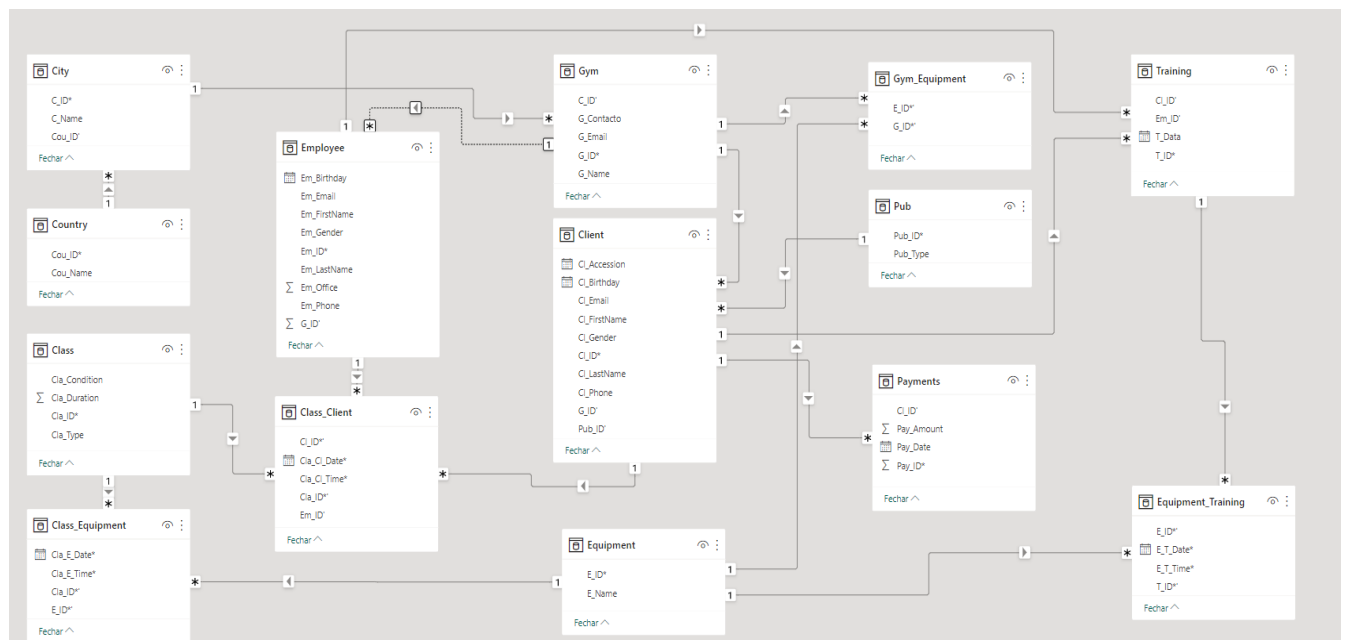


Image 1. Schema in UML notation of "myGYM".

Subsequently, we proceeded to the definition of primary and foreign keys, as can be seen through the relational schema in UML notation, Figure 1. In it, we can observe the schema of relationships between entities, as well as the primary keys, represented with a " \* ", and the foreign keys, depicted with a " ' ". Starting with the primary keys:

- **C\_ID** (primary key of the entity **City**)

- Cou\_ID (primary key of the entity Country)
- Cla\_ID (primary key of the entity Class)
- Em\_ID (primary key of the entity Employee)
- G\_ID (primary key of the entity Gym)
- Cl\_ID (primary key of the entity Client)
- E\_ID (primary key of the entity Equipment)
- Pub\_ID (primary key of the entity Pub)
- Pay\_ID (primary key of the entity Payments)
- T\_ID (primary key of the entity Training)

Regarding the table "Class\_Equipment", a primary key was defined, composed of the foreign keys Cla\_ID and E\_ID, together with the attributes Cla\_E\_Date and Cla\_E\_Time, referring to the day and time of use of a given equipment for a given class. The same line of thought was carried out in the creation of the composite primary key of the entity "Equipment\_Training", being the result of the foreign keys E\_ID and T\_ID, together with the attributes E\_T\_Date and E\_T\_Time, concerning the date and time of a piece of equipment, in a given training. Regarding the entity "Class\_Client", the primary key is formed by the junction of the foreign keys Cla\_ID and Cl\_ID, as well as the attributes Cla\_Cl\_Date and Cla\_Cl\_Time. Finally, the "Gym\_Equipment" table has as its primary key the foreign keys E\_ID and G\_ID.

With regard to the foreign keys present in the entities, and taking into account what was explained in the previous paragraph, the foreign keys that remain to be defined are:

- Cou\_ID is a foreign key in the entity City.
- Em\_ID is a foreign key in the entity Class\_Client.
- G\_ID and Pub\_ID are foreign keys in the entity Client.
- C\_ID is a foreign key in the entity Gym, referente à localização (cidade) do ginásio.
- Cl\_ID is a foreign key in the entity Payments e Employee\_ID.
- Em\_ID and Cl\_ID are foreign keys in the entity Training.

As mentioned above, through the relational schema in UML notation, Image 1, one can visualize the relationships between the tables, all of them being one-to-many, where relationships are associations between entities. As an example, a country has several cities, but a city only belongs to a country. Furthermore, another relationship present in the project is that a city can have several gyms, but a gym is only located in one city. In this way, relationships in a relational schema in UML notation are essential to represent associations, the structure and

the behaviour of the system. They provide a clear view of the interactions between entities, allowing a more complete and comprehensive representation of the system.

In order to avoid anomalies of insertion, removal or modification of data, special attention was paid to the structuring of data according to the Third Normal Form (3NF), in order to ensure the absence of repeated groups, as well as multiple information in the same line and the absence of partial and transactional dependencies, thus improving data integrity and reducing its redundancy.

In addition, during the design of the database, care was taken to ensure that both entities and attributes were unique, i.e. all having a different name. It was also ensured that there were no duplicate records, and that each relation had a unique value, thus avoiding duplicate and redundant information and ensuring data accuracy and integrity.

The structure presented above, Image 1, can help the chain of gyms in the management of their daily operations, providing information regarding a given client, that is, both in terms of personal information, as well as in terms of management of class scheduling, training, and cash situation, that is, if the client in question has his situation regularized before the gym. In addition, it allows to conclude which advertising method is the most important and then define what the company should bet more. This database can also benefit the organization regarding the management of its employees, that is, it allows the visualization of their personal data, which gym they work in and which classes and training they teach, and which employees are most requested by customers. Finally, the database structure will allow a better management of the equipment, being possible to obtain the information of which equipment is most used in each gym and which are the most requested times to use them.

### **SQL Language (MySQL workbench)**

In today's world, businesses and industries deal with a huge amount of data and it is critical that this data is stored, understood, and analysed so that informed strategic decisions can be made, resulting in greater profitability and solving real problems. This can be done using the SQL language, which plays a crucial role in this process.

Understanding and analysing databases involves modifying index and table structures, as well as adding, updating, and removing rows of data. This also enables the retrieval of subsets of information in relational database management systems. This capability is useful for transaction processing and other applications that need to communicate with each other via a relational database.

Regarding the work itself, all the tables of the database were created, and the first three rows of each table were inserted in MySQL Workbench, as suggested by the teachers. However, due to MySQL Workbench restrictions, especially regarding Foreign Keys, it was necessary to adapt some values so that the code could be executed without errors. These adaptations were made to ensure compliance with the rules of the application and to allow the successful execution of the code.

```

20 • create table class (class_id int(9),
21                       class_type varchar(25),
22                       class_duration int(9),
23                       class_condition varchar(25),
24                       primary key (class_id));
25 • insert into class values (1,"Abs",15,"Medium");
26 • insert into class values (2,"HIIT",20,"Advanced");
27 • insert into class values (3,"Full Body Training",25,"Advanced");

```

Image 2. Example of creation of table “Class”.

After completing the initial step, some essential commands were added for data analysis. A total of twelve queries were created in order to analyse parameters that were considered pertinent and relevant to the study. As such, we proceeded to evaluate how many and which customers joined the gym from 2022 (inclusive). Next, the most popular gym was verified considering the number of classes and presence of customers.

```

169 #See all clients that joined our gyms since 2022
170 • select *
171 from client
172 where client_accession > '2022-01-01';
173
174 #Finding the most popular gym
175 • select class_type, count(*) as total_clients
176 from class
177 join class_client on class.class_id = class_client.class_id
178 group by class_type
179 order by total_clients desc
180 limit 1;
181
182 #Table of employees and their gyms
183 • select employee_firstname, employee_lastname, gym_name
184 from employee
185 join gym on employee.gym_id = gym.gym_id;
186
187 #Calculating the total revenue by gym
188 • select gym_name, sum(payment_amount) as total_revenue
189 from payment
190 join client on payment.client_id = client.client_id
191 join gym on client.gym_id = gym.gym_id
192 group by gym_name;

```

Image 3. Example of some Queries utilized.

Subsequently, the "Employees" table was verified, which allowed the collection of information regarding the identification (name and surname) and the total number of employees that existed in each of the gyms in addition to the number of customers distributed by them.

By analysing the customer table again, it was possible to ascertain the total number of customers per city, the customer with the highest number of payments made and the average age of each customer.

Finally, exploring the table of "Gym", "Payments" and "Class", it was possible to conclude the total revenue generated per gym, the total equipment divided by locations and the total participation in classes, per client.

All this analysis proved to be very important, since it represents another way of understanding the results obtained for the chain of gyms, allowing the understanding of several aspects that were relevant to the success and progress of the business and the possible strategic decision making, both in terms of revenue, as well as in terms of employees and classes / machines to be kept in operation.

### **Dashboards created in Power BI (Power Bi Report and Dashboards)**

As discussed in the previous point, companies have an abundance of data at their disposal, coming from various sources and systems. However, this data alone does not provide significant value unless it is transformed into useful, actionable information. This is where dashboards come in.

Dashboards are data visualization tools that provide a clear, organized and concise view of an organization's performance, metrics and key indicators. They bring together data from different sources and present it in a visually appealing and understandable way, allowing managers and decision-makers to gain relevant insights in real time to aid decision-making.

There are different types of dashboards designed to assist different levels of decision making. Taking into account the proposal and with the aim of providing the best information about the gym chain, 3 different dashboards were used: one operational, one tactical and one strategic.

The link to the Dashboards Cloud is below:

[https://app.powerbi.com/links/sVcFGCD7Rr?ctid=b3e38b18-ba39-42b1-ac1a-3725b548a1ba&pbi\\_source=linkShare&bookmarkGuid=a88d6de7-b00f-48a9-bac8-c433c5c8fb46](https://app.powerbi.com/links/sVcFGCD7Rr?ctid=b3e38b18-ba39-42b1-ac1a-3725b548a1ba&pbi_source=linkShare&bookmarkGuid=a88d6de7-b00f-48a9-bac8-c433c5c8fb46)

And the Cloud link to the tables is as follows:



[https://ucppt-my.sharepoint.com/:f/g/personal/s-sobrfernandes\\_ucp\\_pt/EpWDXxs8PtRLgqkzpYQyWJsBIJk08UnEfnwir7VmFTmORg?e=HIzvog](https://ucppt-my.sharepoint.com/:f/g/personal/s-sobrfernandes_ucp_pt/EpWDXxs8PtRLgqkzpYQyWJsBIJk08UnEfnwir7VmFTmORg?e=HIzvog)

## Operational Dashboard

The operational dashboard provides a real-time view of daily activities and key operational performance indicators. It has the function of projecting the day-to-day operations of the organization so that any problems can be identified and corrected quickly.

Taking this into consideration, the operational dashboard focuses only on the year 2022 and the month of December, since it is the last one in the database considered.

Going deeper into the operational dashboard, one can find several graphs and tables with relevant information about the company's activity. In the upper left corner, it is possible to filter the information according to the desired. The reader/manager can select from the desired country to the gym. In the lower right corner there are two ring graphs with data on the equipment of the gyms: one shows which equipment is being used most in group classes, the other reports which equipment is most used in personalized training. To create these two graphs, two new columns, "Cla\_E\_ShortDate" and "E\_T\_Short Date", were created using the month and year functions of Power Bi, which contain the number of the month and year of equipment use.

On top of the information about the equipment of the gyms, two tables about the staff members were added. These tables give a quick overview of how many trainings and classes each staff member has led and in which gym they were given. From this information, it is possible to analyse if there is a good distribution of classes and trainings among employees, and if it makes sense to change the approach practiced or not.

On the left side of the dashboard, three information cards were placed: one with the number of classes of the month, another with the number of personalized trainings and finally one with the value of the total revenue of the gyms, to make known the performance of the company.

Finally, a horizontal bar chart was added with the number of group classes of the month and sorted in descending order, in order to understand which classes have the highest adherence, and two funnel charts: one with the data of payments per client, in order to understand which are the largest sources of income of the gyms in the month under analysis, and another with the top 3 of the gyms by revenue.

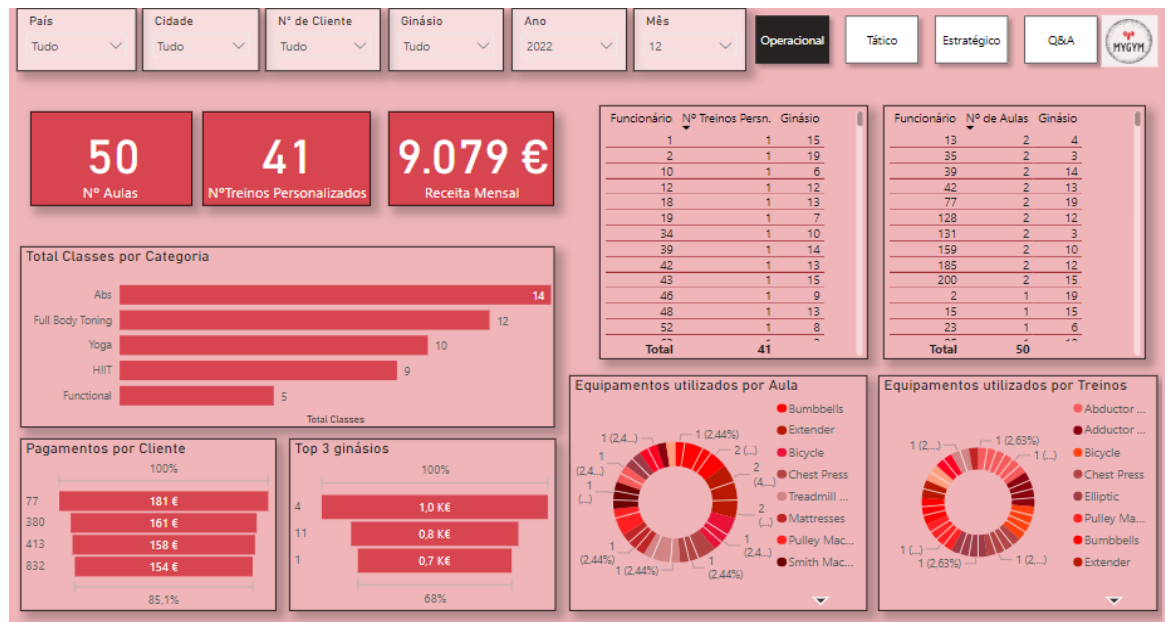


Image 4. Operational Dashboard of “myGym”.

## Tactic Dashboard

The tactical dashboard aims to provide in-depth insights into the performance and effectiveness of strategies and initiatives at departmental or functional level. It enables a more comprehensive analysis of data to identify trends, patterns and areas for improvement that can drive efficiency and success of operations in the medium term. Thus, the focus of this Dashboard is on 2022, which is the most recent year. However, with the help of the filters, you can always check previous years and make a comparison between them, to understand if the company is on the right or wrong track.

Analysing the tactical dashboard in detail, it is possible to draw some conclusions regarding the performance of the gyms.

Through the bar chart "Total Revenues per Month", you can check and compare the level of money generated by the gyms or you can select, in the "Gym" filter, the gym you want to check, and analyse the evolution of these same revenues over the months of the year. In this graph, a new measure called "Pay\_Total" has been added, which is the sum of payments from gym customers, this was originated using the DAX Sum expression. Com o propósito de perceber o ponto de situação do valor total das receitas dos ginásios em relação à meta definida, foi adicionado um medidor à dashboard, “Valor da Receita Total”, onde se percebe rapidamente onde se encontra e qual ainda o caminho a percorrer até atingir esse objetivo. Ao lado do medidor, pode-se encontrar uma tabela com as mensalidades por clientes que justificam o valor total das receitas e onde dá para identificar os principais clientes dos ginásios.

Another element of this tactical dashboard is a WordCloud chart with the title "Advertising". The purpose of this graph is to let the manager know which advertising medium(s) are most effective in acquiring new customers, in this case friends. It is relevant information, since it allows the company to understand where it should invest more in its marketing.

Another relevant graph in this dashboard is the "Customers by Gender", because through it you can identify customer patterns and thus guide and develop some activities in gyms based on this information to captivate and attract more people.

Lastly, two graphs more focused on employees were created. By the graph "No. of Trainings per Employee" it is possible to identify the number of personalized trainings that each employee has carried out, and to observe the distribution of trainings by employees, that is, to understand if there are employees with many trainings in their charge and if others do not have any training. In the graph on the right, the object of analysis is to understand which officials teach the most classes. For its elaboration, the data "Em\_ID" and "Cla\_Cl2022" were considered, the latter column being generated through the Dax IF syntax, thus allowing to filter the number of classes of 2022. With these two graphs, conclusions can be drawn about the employees, and decide what to do with each of them, in order to be more efficient and profitable for the company.

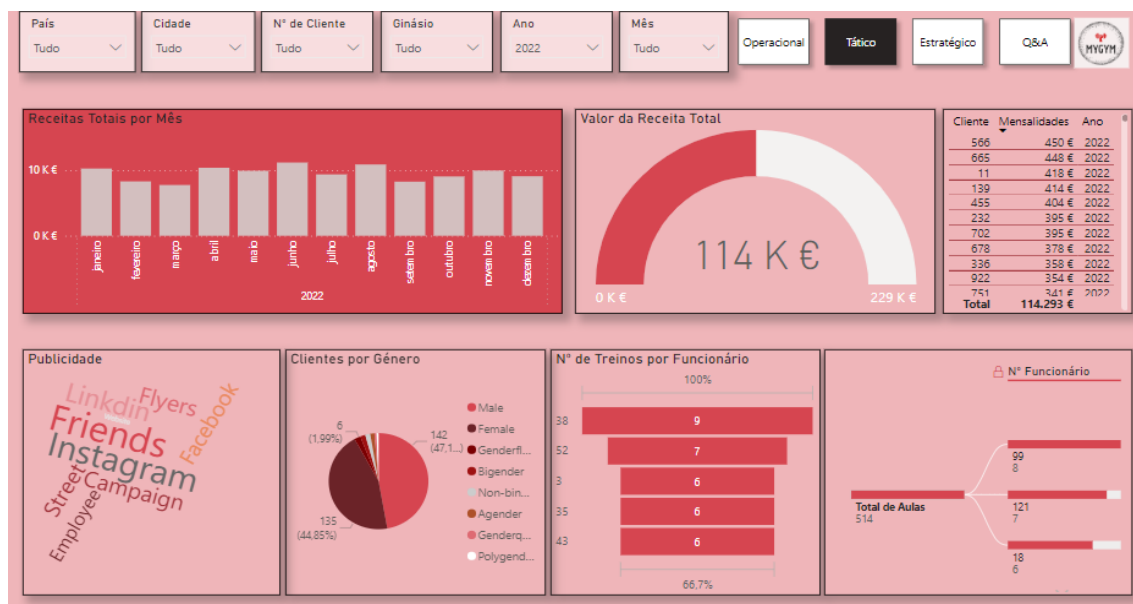


Image 5. Tactic Dashboard of "myGym".

## Strategic Dashboard

The strategic dashboard provides a comprehensive view of key performance indicators related to long-term strategic objectives. It is designed to assist leaders and executives in assessing progress against strategic goals, identifying opportunities and challenges, and making informed decisions to drive long-term organizational success.

In this dashboard, it is intended that the manager can have an overview of the business, so it was decided to place four cards at the bottom of the page with information on the number of gyms, the total amount of revenue generated, the number of active customers in the 20 gyms and also where these gyms are located, in how many countries and in how many cities. The map in the bottom right corner illustrates this same information, i.e. it shows the location of each gym.

In order to analyse in more detail the revenues of the gym chain, four graphs have been added for this purpose, the first one corresponding to the "Payments per client". The purpose of this graph is to inform the top 5 of the main clients, and from there it is possible to calculate the amounts spent by these clients to understand whether the revenues are very concentrated in a restricted number of clients. From here, some strategies can be discussed to combat this concentration of clients, if any, as this can be a threat in a situation where these same clients drop out. With the line chart "Revenues per year", you can see and analyse the evolution of revenues over the years of activity, which is a good indicator to understand how the company is evolving. Subsequently, you can find a stacked line and column chart, "Revenue and number of customers per Gym", which gives readers very useful information. From the outset, it is possible to identify which gyms have earned the most revenue in recent years, or by selecting the desired filters in the upper left corner, a specific year or country can be analysed. In this graph it is possible to visualize the relationship between customers and revenues of each gym, thus analysing the relationship between the two variables.

Finally, following the analysis of revenues, a bar chart called "Revenues by City and Country" was added, where in a direct and clear way it makes known the value of the billing by city, organizing them in descending order in terms of value. In the upper right corner of the graph, the same analysis was made, but in this case by Country.

With the graphs "No. of classes per category" and "No. of personalized trainings per client" the intention is to give more technical information. The reason why these graphs are in this dashboard is so that top managers know which classes are most attended by category and the number of personalized trainings per client, and with this, if necessary, give the indication to middle management whether or not to continue with this type of strategy. At the origin of these graphs are two measures that were created through the DAX count expression: "Number T Cl" which calculates the number of trainings per clients and "Total Classes" which calculates the number of classes.

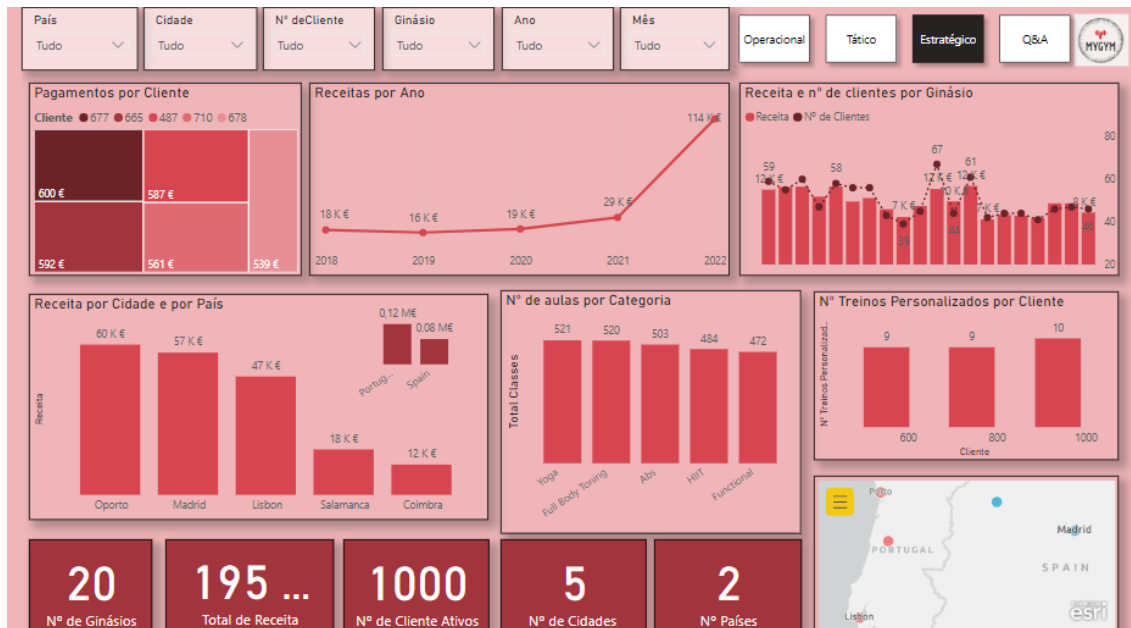


Image 6. Strategic Dashboard of "myGym".

## Dashboard Q&A

Finally, a Q&A dashboard was developed, where the questions and answers that can help in the elaboration of a more complete analysis can be found. It is important to note that all dashboards contain the logo of the fictitious company, "myGym". Through this tool, it is possible to ask questions that cover all the analyses made throughout the work, which help the reader / manager, making it possible to obtain concrete and exact answers to a specific topic.



Image 7. Q&A Dashboard of “myGym”.

## **Conclusion**

In summary, this work of analysing and interpreting a database of the "myGYM" gym chain (fictitious) using SQL and Power BI, provided a comprehensive view of the organization's operations and performance. Through the structuring of the data and the analysis performed, it was possible to gain valuable insights into customers, employees, gyms, equipment and revenues. The dashboards created in Power BI allowed for a clear and concise visualization of KPIs, facilitating operational, tactical and strategic decision-making. This data analysis contributed to the strengthening of the gym chain, providing important information to optimize operations and gain a competitive advantage in the marketplace.

First, it was essential to create a properly structured database, encompassing fourteen distinct and connected entities. Only after completing this step was it possible to proceed with the conversion to the Power BI and MySQL platform.

Analysing the result of the set of Queries elaborated in SQL language, it was possible to make a study of the business itself, answering certain questions regarding customers, revenues, employees and classes held in each gymnasium.

Moving on to Power BI, once the conversion and verification of the links between entities was done, three dashboards were executed, according to three analyses, these being operational, tactical and strategic. Through them, it was possible to answer both the questions previously elaborated in SQL and to make a more complete and detailed analysis of the different attributes of each entity. This tool proved to be the most effective and easy to analyse, since it allows an interactive and detailed view of the case worked, allowing a quick, complete and cohesive understanding.

Through the strategic dashboard, it is visible that the year 2022 presented a sharp increase in its revenues compared to the other previous years, being the gym with the ID equal to 13, located in Madrid, which has a higher revenue and is the second gym that has a higher number of customers. Portugal is the country that obtained the highest volume of revenue, with the city of Porto having a turnover equal to 60238€. It is also important to note, in general, that the classes with the highest adherence are those referring to "Yoga", however, in the last month of our analysis (December 2022) the class with the highest adherence was "abs".