NOVA IMS Information Management School

# MGI

Master Degree Program in
**Information Management**

## Bicycle Sharing Systems as a reliable option in Lisbon

Analysing the performance of GIRA

Frederico de Sousa Falcão Esteves Ferreira

Dissertation

presented as partial requirement for obtaining the Master Degree Program in Information Management

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

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

# BICYCLE SHARING SYSTEMS AS A RELIABLE OPTION IN LISBON

By

Frederico de Sousa Falcão Esteves Ferreira

Master Thesis / Project Work presented as partial requirement for obtaining the Master's degree in Information Management, with a specialization in Knowledge Management and Business Intelligence

**Supervisor:** Prof. Doutor Miguel de Castro Neto

**Co-Supervisors/Co-Orientadores***:* Nuno Tiago Falcão Alpalhão

November 2023

# STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledge the Rules of Conduct and Code of Honor from the NOVA Information Management School.

*Frederico de Sousa Falcão Esteves Fereira*

*Lisbon, 30th November 2023*

# ABSTRACT

The increasing attention to provide public transports that reduce pollutant emissions from private vehicles and offer sustainable solutions for traveling around a city is a major challenge to consider when developing urban areas. Therefore, the use of a soft mobility which guides citizens into their daily routines while delivering an eco-friendly option that tackles issues related to traffic congestion and the first-and-last-mile solution is crucial to be addressed. Bicycle-sharing systems emerge as a tech-driven solution that enables operators to access essential information for system optimization in the upcoming years, thereby supporting decision-making, since it has proven to be adaptable and effective, reshaping the way people commute and interact with their urban environments. In this sense, this study intends to clarify the overall performance of Lisbon's bicycle-sharing system (GIRA) and its overall acceptance by its citizens regarding this way of transportation by using data available from open sources. The creation of dashboards visuals allows a better understanding of the information gathered regarding the use of bicycles in Lisbon while aiding in resource allocation and providing a bigger picture of this sustainable transport which is becoming part of our lives every day. The overall analysis conducted in this study shows that GIRA is performing its job as an alternative form of transport in Lisbon.

# KEYWORDS

Bicycle; Bicycle Sharing Systems; Dashboards; GIRA; Sustainable Mobility

## Sustainable Development Goals (SGD):

# INDEX

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF ABBREVIATIONS AND ACRONYMS

| | |
|---|---|
| **API** | Application Programming Interface |
| **BI** | Business Intelligence |
| **BSS** | Bicycle Sharing System |
| **CRISP-DM** | Cross-Industry Standard Process for Data Mining |
| **ETL** | Extract Transform Load |
| **GHG** | Greenhouse Gas |
| **KPI** | Key Performance Indicator |
| **PBSS** | Public Sharing System |
| **TDB** | Trips per Day per Bicycle |
| **URL** | Uniform Resource Locator |

# 1. INTRODUCTION

## 1.1. Context and Background

Urban areas across the world have gone through significant population density shifts. Understanding and responding to this evolution while improving the quality of life for residents, aids in making cities more liveable while meeting the various requirements of their citizens.

As the world's population becomes increasingly urban, and as cities themselves expand, density becomes a defining feature of modern urban life. Transportation plays a big role in this section since it connects people, enables economic activity, and shapes urban development. That being the case, bicycle-sharing systems complement the existing transportation modes by reducing congestion and providing a convenient last-mile solution (S. A. Shaheen et al., 2010). Consequently, it supports inclusive urban planning, conforms to the idea of smart cities, and helps create liveable, environmentally friendly urban environments.

Bike-sharing systems have become a popular alternative for cities looking to promote sustainable transportation and reduce traffic congestion in urban areas (Olmos et al., 2020). These systems allow citizens and visitors to rent bicycles, in different places from urban areas, without the expense or inconvenience of ownership (S. Shaheen et al., 2012).

In past years, bicycle-sharing programs have quickly evolved from small-scale pilot initiatives to widely-used integrated transportation networks in numerous cities all over the world. Technology advancements, rising public desire for eco-friendly transportation choices, and greater awareness of the various advantages of bike-sharing programs have all contributed to this evolution. The great attention to this alternative and eco-friendly public transport is being considered as a viable option, since the use of bicycles as a way of commuting, has beneficial effects on public health, environmental sustainability, and the economy (Shi et al., 2018). Hence, developing more efficient, sustainable, and high-quality transport that addresses the opportunity of leveraging the quality of life that citizens experience in cities is a major goal to have in mind. BSS is seen as an opportunity to reduce the GHG emissions coming from cars, while presenting a convenient solution for transportation, addressing the "last mile" connection issue, and enhancing the current public transit systems.

## 1.2. RESEARCH FOCUS

According to (Machado et al., 2020), Lisbon is focusing on mobility as one of its main goals until 2030, due to its objective of reducing the dependence on private vehicles and, therefore, improving public transportation across the city. With the opportunity to develop a reliable bike-sharing system, that provides a convenient and flexible option for transportation, the GIRA system must meet the needs. GIRA is a bicycle-sharing system in Lisbon, Portugal, which allows residents and tourists to travel around the city while using a public service. This being said, there is no better time to address this issue than now since the GIRA BSS is still in an evolutionary phase, which leaves room to analyse and improve the impact on the city itself. It is important to comprehend this trend since bike sharing schemes have the potential to address a lot of the issues that modern cities confront regarding transportation.

A purpose must be established for a BSS to meet the criteria for success. Nevertheless, BSS success tends to be seen according to the number of trips completed and the emissions reduction made upon its use (Médard de Chardon et al., 2017) even though certain alleged benefits are difficult to quantify. In this way, the present study seeks to provide a thorough examination of Lisbon's BSS with a focus on its impact on urban mobility. It aims to provide actionable insights for the optimization of the city's BSS but also a clear and user-friendly perspective of displaying the information in dashboards that analyse the overall behaviour of bicycles. At the same time, promoting urban public transport that is eco-sustainable, but also inclusive and accessible to all, improving the lives of all residents.

## 1.3. OBJECTIVES

The purpose of this document is to provide a comprehensive overview of bike sharing systems, in a specific approach to Lisbon, focusing on their development over the years and implementation in order to have a reliable option in BSS for transportation within the city, for the future years.

With the data available from public shared sources, an analysis will be conducted to achieve a better quality of this service and encourage individuals to opt for bicycles as a way of transportation across the city of Lisbon. By using interactive visuals, to give a proper view of trends and actions made, this dissertation provides the capability to build a resilient analysis and demonstrate the behaviour of BSS, as a way of promoting sustainable development and improving the quality of life of residents and visitors that need to move around the city.

The following intermediate goals were established in order to accomplish these goals:

- Provide a user-friendly visualization of the data regarding the use of bicycles and the system, making it accessible and engaging for viewers.

- Conduct a comprehensive analysis of the performance over the past year and a half, after the pandemic crises, to assess its overall effectiveness.

- Enhance the user experience of using bicycles in Lisbon, by leveraging insights derived from the available data.

- Offer an in-depth overview of the development and progress of the GIRA BSS program had in Lisbon mobility.

- Discuss and present the results.

As GIRA is a service aimed at providing the necessary conditions for the use of bicycles by its users, one of the best ways to achieve this goal is by continually analysing the performance of the system. Furthermore, it studies how user patterns and behaviour can be applied to optimize every element of the system. This research aims to bridge this gap by examining these specific questions:

- What are the patterns for using bicycles in different seasons and times of the day?

- What is the overall trend with bicycle usage in Lisbon and is it unique or similar to other systems implemented?

- Are there any significant changes over time?

This study aims to investigate the behavioural patterns of BSS in a well-developed capital city and analyse its overall performance, while also improving the quality of life in Lisbon through the use of bicycles as a form of transportation among Lisbon citizens, while enhancing a sustainable way to commute and getting around the city.

## 1.4. RESULTS AND RELEVANCE

In the context of modern urban mobility and sustainable transportation, this study on the Lisbon BSS is extremely relevant. Bicycle-sharing programs show promise as a way to address the problems of traffic congestion, pollution, and the need for environmentally friendly transportation means in cities across the world while enhancing peoples' quality of life. At the same time, analysing the system's operability contributes to a higher understanding of the upcoming forecasts as well as for resource allocation.

The culmination of this study provides practical insights into the workings of GIRA PBSS, as a result of the analysis carried out on the overall performance of the system in the period from July 2021 to March 2023. This comprehensive exploration is crucial for improving Lisbon's current system as for guiding the overall acceptance from citizens towards this sustainable public transport implemented in 2017.

This study serves to provide a better understanding of the factors associated with the use of the bicycle system, as well as the general results it produces when implemented. In this specific case, the use of GIRA BSS as a way of traveling around the city has become a reliable transport that is part of the daily life of its residents. The overall acceptance by citizens shows a continuous increase in people using the system. Yet, the success of a BSS can be seen by different indicators, depending on the perspective of whoever analyses it. While for urban planners, the daily usage of BSS rentals and the investment made towards the use of bicycles as a way of traveling are important measures, for users it can be translated into the accessibility and density of stations regarding their locations, as for transport companies it can be seen as the number of car trips replaced or even changes in BSS rentals. Therefore, this study provides valuable insights into these different perspectives to have a future reliable public service of bicycle-sharing system in Lisbon.

## 1.5. Thesis Structure

This study is divided into 5 following chapters:

In the upcoming chapter, the literature review, it will be discussed and addressed work related to the topic of this dissertation. Mobility, sustainability, and the overall perception around the world about the topic of bicycle-sharing systems are the keys to better understanding the context and the need to address it now. Previous research on bike sharing systems is summarized and well explained, including information of the history and development, benefits and challenges, and its importance in urban mobility and their growth until now, so that a specific approach to Lisbon is easier to understand before moving on to the next chapters.

Next, the third chapter serves as the cornerstone of this thesis, outlining the systematic approach undertaken to investigate the performance of the bicycle-sharing system in Lisbon. Employing a CRISP-DM framework ensures that analytics or data mining projects are aligned with business goals. This chapter delineates the data collection methods, a description of the data that is going to be analysed, and the further steps applied to construct a robust dataset. At the same time, it also highlights the framework designed to yield a comprehensive understanding of the bike-sharing system model to be applied, paving the way for an analysis and evaluation of the results.

The fourth chapter is related to the evaluation and discussion of the results previously gathered and pre-processed. This chapter bridges the gap between raw data and actionable insights, offering a nuanced understanding of how the bike-sharing system operates in the context of Lisbon's dynamic urban landscape.

Lastly, the fifth and sixth chapters denote the conclusions of the work carried out. It provides a comprehensive picture of the dynamics of Lisbon BSS by combining the major conclusions, trends, and implications discovered throughout the previous chapter, while presenting the limitations that influenced this work, as well as the future work to be applied in order to enhance a more conclusive analysis to the Lisbon BSS, taking into account in the context of this dissertation.

# 2. LITERATURE REVIEW

## 2.1. Sustainable Mobility

Urban areas have been increasing at a faster pace than what was expected. Cities have changed throughout time from being sparsely inhabited areas to tightly crowded metropolitan hubs due to different factors such as migration from rural to urban areas, economic opportunities, and changing lifestyles. In order to meet the needs of future generations and pursue a less polluting transport system, there’s an urge to develop different ways of commuting and travelling, although it can be very challenging to reduce the dependence on the private vehicle while giving a proper solution even though there are compelling social and environmental reasons to reduce the usage of cars (Holden et al., 2019). Transportation plays a big role in urban planning and is crucial to the well-functioning of any city. The ability to move from one place to another is critical, making it more or less appealing to live in.

The latest policies involving sustainable transport systems presented in the *2030 Agenda for Sustainable Development* contribute directly to the achievement of the *Paris Agreement*, by reducing greenhouse gas emissions given the fact that almost a quarter of those emissions come from transports (Nations, 2017). It is noticeable that sustainable mobility is often associated as mobility that minimizes environmental impacts since most of the interventions, made for the coming years, have the reduction of greenhouse gas and air pollution as their primary goal (Gallo & Marinelli, 2020). Therefore, providing access to accessible, affordable, and sustainable transport systems becomes a major goal to have in mind, considering the need of achieving sustainable mobility for the upcoming years. In addition to promoting accessibility and economic growth, sustainable mobility can also increase economic integration while preserving the environment.

Plans that promote public transportation and car restrictions without compromising future generations are the main goal of this concept of “sustainability mobility” and, therefore, are seen as a crucial subject to address as the world is changing at a rapid pace, driven by the evolution of technology (Gallo & Marinelli, 2020).

Addressing environmental inefficiencies and enhancing the performance of mobility with a focus on the environment, is one of the main goals of the future shared transport policy (Høyer, 1999). Sustainable mobility prioritizes eco-friendly transportation, enhances air quality, and reduces carbon emissions as well as traffic congestion made by vehicles.

As a result, there has been a growing focus on designing the built environment to facilitate outdoor soft mobility and physical activity. Specifically, there is a keen interest in encouraging built environments that support active modes of transportation, such as walking and cycling, which are the soft modes of mobility (D. Chapman & Larsson, 2019).

### 2.1.1. Soft Mobility

Soft mobility is considered any non-motorized transport means, also intended as zero impact movability which includes mainly cycling and walking. Referred as active or human-powered mobility, the tendency to use these modes of transportation is associated with the urban environment and is considered to be environmentally beneficial when substituting motorized travel (D. Chapman & Larsson, 2019). This type of mobility focuses on reducing the pollutant impact that daily motorized transport has on urban areas, which eventually contributes to improved road safety, reduced car accidents thanks to a lower use of private vehicles, and, therefore, a reduction in traffic congestion. It has also been proven to be partly responsive for urban regeneration of degraded areas despite that its initial objective was to reduce the dependence on private vehicles for short-distance trips as a mean to minimize vehicle emissions (Corticelli et al., 2022).

Cities that implemented urban mobility strategies to encourage public transportation and limit the use of private vehicles provide a bicycle-friendly atmosphere for the use of BSS. In addition, most cities with bicycle-sharing programs have extensive urban mobility plans that encourage cycling and reduce the need of having a car for daily purposes (Midgley, 2011).

By adopting bicycle-sharing systems as a new option for public transportation, consumption of natural sources and air pollution can be reduced. Motor vehicles are one of the major contributors to air pollution and environmental degradation, while contributing to climate change in a negative aspect. Furthermore, BSS contributes to a healthier lifestyle due to its increased physical activity besides the environmental benefits, while also contributing to a reduction of some travel times, and parking space, therefore being able to provide a larger creation of a greener environment (Eren & Uz, 2020).

Expanding and integrating cycling into transportation networks is the ultimate objective of bicycle sharing, as it makes it easier to become a regular mean of transportation (S. A. Shaheen et al., 2010). The majority of individuals are unwilling to walk medium-long distances, even though journey distance by mode differs by city. The typical cycling distance is between one and five kilometres. As a result, bike sharing might address that problem perfectly (Midgley, 2011).

## 2.2. General Overview of Bike Sharing Systems

### 2.2.1. History and Development of Bike Sharing Systems

Bike sharing systems consist on renting available bicycles in different docking stations, distributed around the city, for short or long distances trips, so that inhabitants of a city can travel between A-to-B either by electric or regular bicycles (Otero et al., 2018; Vallez, C.M., 2021). These systems typically involve a network of bike stations, where users unlock bicycles using a smartphone application or a membership card. BSS is seen as a way of encouraging cycling around a city and making a more accessible type of transportation (Buck & Buehler, 2012; S. Shaheen et al., 2012).

It all started back in 1965 in Amsterdam, when a plan called “White Bicycle” was presented as the solution to improve public transportation and fight the rise of pollution and cars (S. A. Shaheen et al., 2011). The main idea was to leave white bicycles around the city, to be freely used for one trip and then left for someone else in need of transport, by the inhabitants of the city, but it wasn’t a great success due to continuous thefts (DeMaio, 2009; O’Sullivan, 2022; Zee, 2016). Later, around the mid-90s, in Denmark, a second generation of BSS was developed which worked in specific locations, from the docking station, with a coin deposit that you could get back, if the bicycle was returned to one of the docking stations, although it would still experience theft. (S. Shaheen et al., 2012). Following, a third generation was developed, which brought information tracking through incorporated technologies (Ricci, 2015), that differs from the previous generations by its distinguishable bicycles, specific docking stations, smart technology for bicycles to check-in and check-out, user interface, and specific programs to identify users due to theft deterrents (S. Shaheen et al., 2012; S. A. Shaheen et al., 2010). Finally, a fourth generation, more technologically developed, arrived. This last generation consists in a dockless IT-based system, meaning that bicycles are located at arbitrary places and charges are based on the usage time of the mobile app/seasonal subscription. The third or fourth generation technology used by bicycle-sharing services allows users to better plan their routes and locate stations where they may pick up or drop off their bicycles (Patel & Patel, 2020).

A more summarised explanation of the overall evolution of bike sharing systems, that has been driven by technological advancements and the desire to make bike sharing more convenient and accessible for users is highlighted in (Beroud & Anaya, 2012; DeMaio, 2009). It wasn't long before it became clear that using information technology was essential to raising service standards. The systems eventually improved as technology was adopted, allowing, among other things, the tracking and storing of data about trips and their users. Nowadays, bike sharing programs all around the world contain several technological features that allow operators to track movements (Fishman et al., 2013). Nevertheless, with the use of GPS and automatic bicycle counters distributed around cities, it’s also possible to analyse track movements.

Since the beginning of Vélib’, back in 2007, a lot of different BSS’s have been developed across Europe, even though none reached the same proportions as the French BSS (DeMaio, 2009). Nowadays, Vélib’ is considered a landmark in Paris due to its impact and popularity, which makes it the largest bike sharing system in Europe and the third on a global scale, staying behind China’s two bike programs. The development of the PBSS in Paris and Rennes and its overall acceptance, created an interest in BSS as a means of transportation on a global scale (D’Almeida et al., 2021). Additionally, using BSS can help normalize the idea of riding a bicycle in a city or serve as a catapult for more frequent personal bicycle use (Hosford et al., 2018).

Bike sharing is frequently seen as a means of reducing the harmful social and environmental effects of global motorization. The bicycle offers a nearly emission-free alternative to personal car use as a way of mobility (S. A. Shaheen et al., 2011). Overall, as understanding has increased, so has the public's perception of cycling as a mode of transportation. In addition to the obvious health advantages, BSS makes moving around the city much more pleasurable than using the bus or subway.

BSS normally employs one of two types of stations: docking stations or dockless stations. The choice of station type depends on the particular demands of the community and the program's objectives

(Fishman, 2016). Docking stations are physical locations across the city where bicycles may be loaned and returned, being referred to as classic bike-sharing systems. The same station or another station in the system's service region must receive the bike when the trip is finished. Past studies suggest that the location of the stations is strategically defined to improve the quality of the surrounding pedestrian environment (Parkes et al., 2013; S. A. Shaheen et al., 2011; Zhang et al., 2015). On the other hand, dockless stations, commonly referred to as free-floating bike-sharing systems, don't require any physical infrastructure and bicycles may be picked up and dropped off anywhere. The money saved on building docking stations is the primary benefit of free-floating BSS versus docking BSS stations (D'Almeida et al., 2021).

At the same time, BSS also has two different types of bicycles: standard bicycles and electric bicycles. E-bikes are generally more physically appealing than normal bicycles due to the ability to ride farther without worrying about muscle fatigue while providing a versatile and effortless experience.

### 2.2.2. Challenges and Benefits of Bike-Sharing Systems

Bike-sharing systems have been widely implemented in cities around the world as a way to promote sustainable transportation and improve mobility. However, these systems also present several challenges that must be addressed for them to be successful.

On the one hand, there are a lot of benefits associated with the use of bicycles as an active transportation starting by the main ones: citizens' health, economic side, and environmental sustainability, but also reducing congestion, while improving transport flexibility and air quality (Shi et al., 2018). Many different cities are taking into care the associated climate changes connected with vehicle emissions and road congestion since road transport is responsible for up to 90% of urban air pollution (Programme, 2010).

Cycling has beneficial effects on public health such as the stimulation of physical activities which contributes to a much more healthy lifestyle and reduces the risk of health diseases but also the reduction of motorized vehicles on the road, and, therefore, minimization of the air pollution (Otero et al., 2018; Oja et al., 2011; Cavill & Watkins, 2007). This ultimately results in a decrease in noise pollution and in the emission of greenhouse gasses (Clockston & Rojas-Rueda, 2021).

In 2021, a brief resume about the emissions that a BSS emits along its manufacturing, maintenance, and daily use are described in (D'Almeida et al.). Specifically, it shows different perspectives of cities like Shanghai and, Barcelona, and the countries of China and the United States. Yet, Edinburgh is the specific case study of the article and its showed that BSS reduces a small amount of C02 emissions, only 0.5%, although it's contributing to a shift in the right direction. This reduction of GHG has a small impact relative to other transport emissions as a result of small number of trips, short distance trips, and also because most of the journeys don't transfer from private vehicles. In such a manner, most people switch from using public transportation or even walking (D'Almeida et al., 2021; Ricci, 2015).

The reduction in carbon emissions is essential for mitigating climate change. With the use of shared bicycles, what used to be a major problem can be twisted into phased manners of reducing greenhouse gas emissions, as the authors confirm through their articles (Kou et al., 2020; Li et al., 2022). However, it is relevant to mention that GHG emissions reduction has a linear relationship to the number of bicycles, docks, and trips made. The opportunity for reducing emissions as a result of some gradual changes in the transportation system can be significant.

BSS can be used as a one-way trip, round-trip, or even as a multimodal transport. These programs often cover the different costs associated with bicycle acquisition, upkeep, storage, and parking, which helps individuals who do not own bicycles, or even tourists, to take advantage of the benefits of cycling without thinking about the different problems associated with having a bicycle in an urban area (S. Shaheen et al., 2012).

These benefits differ from one person to another, while being influenced by the different factors associated with a bicycle like trip distance, traffic safety, infrastructure, air quality, and regularity. Physical activity is the perfect way to minimize sedentary lifestyles and improve public health, since physical inactivity is a major cause of death (Woodcock et al., 2014).

Bike commuting can be considered as the optimal way of traveling between work and home, since it is the cheapest and healthier due to its daily connection to the natural environment as well as it is an excellent way to meet the minimum needs of physical activity. According to Celis-Morales et al. (2017), people who use bicycles as a way of transport to work or who only use it as part of the route, have a lower risk for cardiovascular diseases and lower risk for several cancers. This study, conducted on 263 450 participants, emphasizes that cycling commuters and mixed-mode cycling commuters achieve the current physical activity guidelines needed to improve their health. In addition, a study was conducted to older volunteers, who don't regularly cycle, highlighting substantial benefits for mental health when riding bicycles weekly. Nevertheless, the participants also showed an increase in their mood and satisfaction from cycling regularly (Leyland et al., 2019).

Another study found that adult bicycle commuters had lower stress and higher satisfaction levels than those who used other modes of transportation (Avila-Palencia et al., 2017). For people to commute by bicycle more frequently and experience better mental health, a well-built and natural environment is crucial. Furthermore, bike sharing systems can be a catalyst for urban development and can revitalize neighbourhoods.

When talking about urban planning, BSS saves more space in the streets, while replacing parking spots for a whole dock station which can reach many more people (Buck & Buehler, 2012). Lastly, BSS can also be seen as an advantage, when compared to other transportation projects, through the low implementation costs and ease of implementation over time, contributing directly in a short amount of time.

Benefits are sometimes overstated, which might also lead to the demise of bicycle-sharing as a viable method of decreasing reliance on and usage of automobiles. Although there are several benefits to grabbing a bike and taking a ride around a city, cycling is also associated with risks, due to air pollution exposure and road traffic injuries (Clockston & Rojas-Rueda, 2021; Otero et al., 2018). Even though the physical benefits are greater than the health risks associated with bike sharing systems.

On the other hand, making sure that the necessary infrastructures are in place to enable bike sharing systems is one of the biggest problems. The decision of using bikes from the docking stations is also influenced by their surroundings. Docking stations that are closer to work locations, residential neighbours, colleges/schools/universities, or public transports stations tend to be more accessible to users (Caggiani et al., 2021; Fishman et al., 2015; Hosford et al., 2018). According to S. A. Shaheen et al. (2011), a study from China's Hangzhou PBB showed that users tend to use bike sharing systems more often if stations are closest to either workplace (40%) or near their house (40%). Likewise, there's an increase in usage for those who inhabit close to a docking station (Hosford et al., 2018).

Bike paths that are isolated or separated from vehicles have a positive relation with the demand, as it increases driving safety and, of course, comfort (Eren & Uz, 2020). At the same time, the least favoured bicycle lanes among cyclists are those without infrastructure or a mixed one, since the choice of the driving path is influenced by the feeling of safety (Hunt & Abraham, 2007). The number of motorized cars tends to decline on the road as more people choose to travel by bicycle, also reducing traffic congestion (Woodcock et al., 2014). A safer cycling environment might be promoted by a higher bicycle presence overall since the presence of more bicycles, on the road, improves driving behaviour as a result of increased awareness towards bikes (Jacobsen, 2015). Additionally, using PBSP can help normalize the idea of riding a bicycle in a city or serve as a catapult for more frequent personal bicycle use (Hosford et al., 2018).

Additionally, building needed infrastructures, operating and maintaining bike sharing programs may be expensive, and there may be problems with funding and finance, as many systems depend on government funding or corporate sponsorship to function, which can be challenging to get. It can also be difficult for BSS to generate sufficient money to pay operational expenses.

Theft and vandalism are one of the most common challenges applied to the BSS, despite the user identification technology already developed (Midgley, 2011). Theft and vandalism are one of the main reasons why most of the past BSS generations didn't work, and to prove that theory, just two years after the launch of Vélib's bicycles, a report was made, which showed that almost 20,000 bikes disappeared or were vandalized. Some systems have implemented measures to prevent theft, although this is a recurrent problem, according to S. A. Shaheen et al. (2010).

Another important aspect to consider is bicycle redistribution, as it is needed to redistribute some bicycles into key locations after use, in order to satisfy the constant demand. With this, the system must be accessible to all members of the city and ensure that citizens are able to find or park a bicycle at each station (Chemla et al., 2013). Rebalancing bicycles is often done by a fleet of specialized, capacity-limited trucks. This fleet is headquartered in a central depot, and the vehicles distribute bicycles to stations in need after collecting them from stations with greater occupancy levels than expected. Another way of redistributing bicycles is through a points system, in which people are invited to look for stations that have low levels of bicycle occupancy, so that there is constant repositioning and rewarding of users who do so.

Weather conditions also have a great impact on bicycle usage, if not already familiarized or having the proper equipment. Low temperatures, precipitation, and snow tend to have a negative correlation when compared to the bike sharing demand, while increased temperatures until 30ºC tend to have a positive correlation (Eren & Uz, 2020).

At the same time, BSS has a negative effect on the pedestrian city culture, since it makes previous journeys made by foot now by bicycles, which is more pleasing due to its saving time, comfort, and accessibility. Nevertheless, by increasing the proportion of journeys made by bicycle, most of the short routes that were done before by foot, can now be replaced by bikes. According to Jensen et al. (2010), the average trip distance in Lyon's BSS was 2.49 km, which shows that several trips were short route paths.

### 2.2.3. Bike Sharing Systems in Lisbon - GIRA

With the aim of creating sustainable habits in cities and increasing the use of bicycles, which is the most sustainable mean of public transport, GIRA, Lisbon's bicycle-sharing service, was born. This service began in 2017 and currently belongs to the Empresa de Mobilidade e Estacionamento de Lisboa, better known as EMEL. One of the purposes of EMEL is the construction and operation of infrastructures supporting pedestrian mobility, electric mobility systems, and shared mobility products (EMEL, 2019).

This service is only located in Lisbon, even though there are more systems scattered around the country. Although GIRA is the first system in Lisbon, it is not the first bike sharing system in Portugal, as the first bicycle programme developed in Portugal was in Aveiro 2010 (do Nascimento Liberto, 2018). After that, several other programs were developed until the creation of GIRA in 2017. Even though the introduction of this program was later than expected, the city's orography seemed to limit riding only to athletes due to Lisbon being referred to as the city of seven hills, but actually, Lisbon has 73% of flats streets, with slopes lower than 5% according to Machado et al. (2020). Regardless, this makes it a possible candidate for a cycling city since the orography is not such a big problem after all. At the same time, with the arrival of electric bicycles, this challenge was eliminated because they increased the system's accessibility and range, making it much simpler for users to travel farther or across challenging landscapes. In fact, the hills of the city are much more accessible at this scale (Machado et al., 2020).

The Lisbon City Hall has plans to invest in a strategic vision for mobility until 2030 in order to reduce the use of private cars from 46% to 34%, and reduce by 15% the rate of sedentary lifestyles (SAPO, 2022; Machado et al., 2020). Of course, this is only viable if efficient alternatives are provided, therefore, not only bike sharing is taken into account in this goal for 2030. However, BSS's are seen as one of the potential sustainable mobility keys for cities due to its popularity, it's user-friendliness, affordable prices, short distance trips, and stress saving due to traffic congestion (Bullock et al., 2017; Médard de Chardon et al., 2017). This would relieve the city of additional traffic and pollution by lowering the number of automobiles that are in circulation every day in Lisbon, where 68% of trips conducted inside the city are for less than five kilometres and 31,8% of journeys up to 5 km are performed by private cars (André, 2021).

*MOVE2030* is a strategic reference document, which defines the municipal actions in the area of mobility for upcoming years. According to it, most workers and students use the car for their daily commutes, therefore the presence of a private vehicle is very notable. Back in 1981, Lisbon public transports was the main mode of transportation with a share of 67% when compared to the 14% of cars. Nevertheless, there has been a paradigm shift because in 2017 only 22% of total journeys were made by public transports while private cars added a total of 46%. Bikes also saw a powerful increase of 200%, when compared to 2011, representing 0,6% of the total journeys. All this information can be seen in Machado et al. (2020).

Initially, tests were made in the Parque das Nações area, with a total of 200 bicycles and 20 stations, later evolving to the Greater Lisbon area, where most of the docking stations are now located. After only one year of existence, GIRA's bicycles had already cumulated more than one million journeys and a daily bicycle trips usage of more than seven (Félix et al., 2020). Due to a great adhesion of the citizens to the alternative presented by EMEL to circulate in the city of Lisbon through the use of the bicycle as transportation or as a complement to the public transportation already available, GIRA evolved in a phased manner, always presenting a gradual growth over the years since implementation.

In 2021 there were already 102 operational stations and around 1600 bicycles in total (Vasconcelos, 2021). SIC Notícias (2022) states that only 1000 tend to be in circulation on Lisbon streets, due to maintenance problems. The rest of the fleet is kept at the workshop due to constant repair effort that is needed to meet the demands of a system and, therefore, ensure the reliability and convenience of the system. Nevertheless, in the last quarter of 2022, GIRA grew the number of stations in operation to 139, which translates into around 2,700 docks (André, 2022b).

GIRA is a docked bicycle-sharing system that offers traditional pedal bicycles and electric pedal assist bikes, which tend to be the preferred ones, according to the preferences demonstrated by users (André, 2022b). The system fleet has 60% of electrical bicycles and 40% of standard bicycles, with future investments plans of electrifying a portion of its conventional fleet and acquiring more than 500 electrical bicycles due to the fact that on average 80% of the total daily trips are made by electrical ones (Vasconcelos, 2021). The fact that most of the Gira's bicycles are electric, contributes to a lower physical demand by its users, making it much more appealing to use.

Giving some more detail about the usage of GIRA, an observation in the first years after the implementation of the system was developed by Félix et al. (2020). The timeline that presented the more flows between cyclists was between 8:00-10:00 am and in the afternoon between 5:00-7:0pm. At the same, there was a significant increase in the cycling volume from 2017 to 2018, in locations near bike-sharing system, proving that users started to embrace this method of commuting more often.

According to EMEL estimations, since September 2017 over 10 million kilometres have been cycled by GIRA bicycles which is equivalent to almost 5.6 million trips. Also in February 2021, the system broke the record of 10 000 daily trips, in a week where the average was always above nine thousand uses (André, 2022a). At the same time, a study conducted by *Instituto Superior Técnico* to EMEL relating GIRA BSS showed that more than 300 tonnes of CO2 equivalent in GHG are thought to have been prevented since it was implemented on more than 3.6 million trips (André, 2021).

As know, BSS is a way of promoting combined travel in a city, making it easier for citizens to opt for different sustainable alternative solutions. Considering this, if the rise in public transport generated by the usage of public bikes is taken into consideration, the consequential reductions in carbon emissions produced by the use of bicycles would be far larger.

### 2.2.4. Overall Perspectives from Other BSS

Leonardo Caggiani demonstrates that public acceptance of the use of BSS can be seen as a pertinent factor that influences the success both in Swedish and global contexts (Caggiani et al., 2021). At the same time, also emphasizes the contrast of usage of bicycles before and after the COVID-19 pandemic. Although the pandemic of COVID-19 stroke recently, and could be seen as a step back, actually it worked all the way around given the fact that the number of trips and distance covered, when compared to the non-pandemic time increased, which ultimately shows its prospective mobility towards the future.

The evolution of BSS is considered global since 2010, when there were nearly 100 BSS programs operating in an estimated 125 cities (S. A. Shaheen et al., 2010). In 2015, they significantly grew in prevalence and popularity to include over 800 cities across the world and a global fleet exceeding 900,000 bicycles (Ricci, 2015). In 2021 there were more than 2,500 open systems, according to Urban Solutions (2021).

As stated before, S. A. Shaheen et al. (2011) gives an overall view of the usage impact on mobility in Huangzhou and other successful BSS programs. In that sense, and according to a survey made, the surprising finding is that 78% of the participants who had cars said they had utilized bike sharing for trips they had previously made in their cars. Yet, more than 80% of participants who don't have cars also stated that they now use BSS for the same journeys made previously by bus. Following, 50% of households with cars also used bike sharing as a bus substitute, which can be translated to a switch from bus transit to bike sharing as a replacement for these kinds of transportation. These findings clearly imply that bike sharing is encouraging more individuals to ride bicycles.

On the other hand, China's experience insinuates that a great portion of the bicycle-sharing system users tend to choose this way of transportation to complete their journey and not as a way of only commuting between home and work (Zhang et al., 2015). The shift from private vehicles to bicycles seems to only occur for a small percentage of users. Still, amongst European examples, Barcelona shows the higher percentage of transition with 9.6%, followed by Lyon with 7% and London with 2%. On the other hand, examples outside European countries include Australia's two BSSs with 21% and 19%, Washington D.C. with 7% and Montreal with 2% as the author shows in (Ricci, 2015).

Nevertheless, either in Lyon or Washington D.C., reports show that trips that would otherwise have been made by private vehicles, were replaced by the use of bicycles (S. A. Shaheen et al., 2010). A growth in the usage of this program highlights the general acceptance towards this mode of transportation making it easier, faster, or more convenient than other options. There has been a rise in the number of bicycle trips in cities with effective bike sharing systems, while other cities have shown a boost in cycling after the launch of programs. Lyon increased its bike usage by 44% in the first year of Vélo'v BSS program (S. A. Shaheen et al., 2011). Paris also became quickly a system to look up for experience by its notable system and performance, later becoming one of the largest programs in Europe.

Ricci, (2015) highlights that the implementation of the scheme also contributes to a start in cycling as a way of commuting, making it a very engaging mode of transportation, which can be seen due to its examples of London, Valencia, and four North American cities where a survey was conducted previously. Due to its popularity, the systems are increasing their number of bicycles gradually. Although Europe and North America both reveal utilisation growth patterns, Europe had a bigger number of cities adopting initial large systems, while North America seemed to develop a system in a more gradual way (Parkes et al., 2013).

The number of daily rentals per bike is a useful measure when analysing the performance of the system (Jiménez & Nogal, 2021). The same report shows that in the first years after Dublin's BSS implementation, there was an average daily rental per bike of 6.2. According to reports, daily bicycle trips usage ranges from 3 to 8 in a good system, and they have been seen to rise sharply when public transportation is connected (Ricci, 2015). This being said, there are some disparities of values when comparing different BSS across the world since either London or Washington D.C. bikes produce on average 3 trips per day, while Australia BSS have a less successful attracting program, by only producing 0.3–0.4 trips per day per bike.

A perfect comparison between TDB across the world as well as providing information from the number of stations and bicycles from each system. In this study, Brisbane, Brussels, and Minneapolis, which have a system much larger than Dublin, Ljubljana, or Lithuania either by number of stations or bicycles (estimation of having 4 to 12 times the system bigger), present a TDB much lower than the supposed, between 0.3 and 1.1, while the small systems by the European countries presents some of the highest TDB scores reaching values between 6.0 and 8.2 is demonstrated by (Médard de Chardon et al., 2017). Besides that, Helsinki's BSS which only operates in months that are expected to be frost-free, is also seen as a system capable of providing users the usage needed thanks to its 6.0 and 8.7 TDB, making it a close competitor for the world's most successful BSS (Willberg et al., 2021). The claim that a BSS must be large to be considered successful is easily refuted by the existence of systems like the ones in Ljubljana, Dublin, and Vilnius, which have high TDB values.

As said before, Dublin is considered a system with lots of usage from its users. Either way, its success cannot be determined by a single ratio (TDB) because it doesn't give enough information about the system. (Jiménez & Nogal, 2021) provides an overall overview of BSS in Ireland where he concludes that the determination of whether a BSS will be successful in surviving over time comes from a successful integration into a public transportation system. Additionally, the main common result it showed is that BSS are considered aesthetically pleasing and flexible urban mobility systems that grow and develop quickly.

A case study conducted in Poland in 2020 gives an overall view of the usage, acceptance, and satisfaction of the system among its users. The results of the analysis showed that more than half of the participants were fully satisfied with the functioning of the system while only a very small percentage were not satisfied at all (about 3%). This contrast also highlighted that there is a substantial association between the use of the bike-sharing system and the degree of pleasure associated with its use. Nevertheless, the paper also indicates that there's a yearly growth of BSS users in Poland with peak hours between 3:00 pm and 9:00 pm, mostly on working days and in months where the temperature is higher (April to September) (Macioszek et al., 2020).

Trips made by bicycle tend to be under 10 km. At the same time, an Italian survey shows that bike trips made up to 2 km have 8,7% of use when compared with 27 % of use by cars, 58% by foot, and the rest by motorbikes and public transports, according to (Gallo & Marinelli, 2020). On the other side, trips between 2 km and 10 km had a decrease in both soft mobility means, 4,9% by bicycle and 4,6% by foot, instead cars (72,8%) and public transports (13,9%) raised their %.

When analysing the behaviour of age groups, younger people tend to use the system often when compared to middle and older age groups (Böcker & Anderson, 2020. This affirmation comes from a study conducted to a group of regions in Norway where it's also highlighted that women tend to be more interested in using the system than men, opposing to Fishman's vision which states exactly the opposite (Fishman et al., 2015).

A study in Helsinki shows that most of the trips made in the system usage tend to be generated by young adults until the age of 30 and male users. The proportion of those living in areas near the system is also a contribution to its use. Results evidence that it is usual for users that grab bicycles as a way of transportation regularly, to make most of trips starting from the same particular station at a given time and that users tend to pick up the bicycles in stations where the previous trip ended. At the same time, the last station in a day where the user left the bicycle is more frequently the same station where he grabbed the first daily bicycle. These findings suggest that users may have more habits regarding the use of bicycles, which helps in the prediction of future investments but also, by showing that BSS is considered an important mean of transportation across citizens as a way of commuting (Willberg et al., 2021).

# 3. METHODOLOGY

This chapter gives us a proper aspect of the methodology used in the making of this thesis and all the steps needed to conduct a study about the need to improve Lisbon's bike sharing system – GIRA. It will give a glance of the work done, and how the data was gathered, pre-processed, and handled, to later be analysed and evaluated so that a proper overview of the results is presented. Each of these procedures is crucial to assure the quality and dependability of the data analysis and to provide results that are useful to be applied to Power BI.

This methodology chapter outlines the structured approach for studying bike sharing systems in Lisbon using the CRISP-DM framework.

## 3.1. ARCHITECTURE OF THE WORK

The architecture of this work is divided into four different steps, with Power BI being the primary tool used for work. Power BI is a service provided by Microsoft that helps to build a *sophisticated reporting system*. Its ability to load data from many different sources and build the model in a much simpler and powerful way provides an effective way of dealing with all the procedures for BI analysts (Ferrari & Russo, 2016).

However, the first step is collecting data from sources, in this specific case open data sources. The second step is to apply all the necessary changes to the data downloaded by cleaning, transforming, and extracting to have only the information that is going to be analysed. This process is called pre-processing. After that, the third step is creating the data model that links all the information from the different tables uploaded to create interactive visualizations which is the fourth and last step of this architecture process. The creation of dashboards with interactive tools to analyse the overall growth and performance of the system is only possible to achieve if the previous steps are handled carefully. This last step reveals all the data into meaningful metrics.

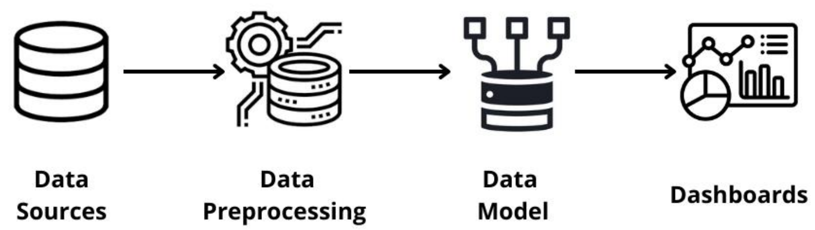


Figure 1 - Architecture of the Work

## 3.2. CROSS-INDUSTRY STANDARD PROCESS FOR DATA MINING (CRISP-DM)

To create and assess a bicycle-sharing system that satisfies user demands and addresses the problems highlighted, the Cross-Industry Standard Process for Data Mining (CRISP-DM) approach will be employed. CRISP-DM methodology is widely used for data analytics and mining methodologies as it serves as a base for a data science process, from an initial business understanding to final deployment, providing an organized approach for interpreting data (Wirth & Hipp, 2010) in this specific case of GIRA bicycle sharing system, which enables to identify patterns, insights, and trends that improve the efficiency of this sustainable transport alternative. It enables a comprehensive analysis of the factors influencing BSS usage and contributes to valuable insights for urban transportation planning in Lisbon.

It's known for having 6 different steps which will be described, one by one: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment as it is possible to see in Fig.2 below:

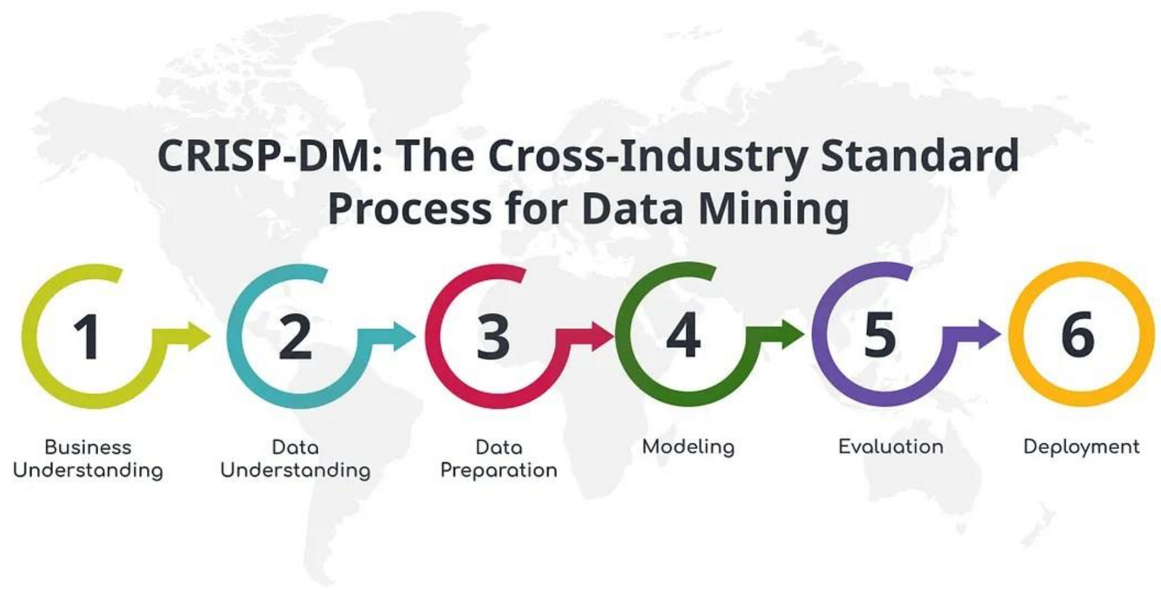


Figure 2 - CRISP-DM Method (source: https://medium.com/@yennhi95zz/0-an-overview-of-crisp-dm-the-cross-industry-standard-process-for-data-mining-1078ba8cd6da)

### 3.2.1. Business Understanding

As stated before, Lisbon is becoming each day more congestioned by traffic all around the city and there's an urge to use sustainable options (Machado et al., 2020). Bike sharing systems are emerging as a widely used and environmentally friendly mode of transportation around the world, especially in urban areas. These systems allow an easier access of bicycles for short-distance trips, which eventually contributes to preventing traffic congestion, the well-being of its users due to physical exercise encouragement, and also creating a greener environment for the upcoming years (Médard de Chardon, 2019). In this particular context of analysis of Lisbon, the bicycle-sharing system has experienced a significant growth and evolution over the past few years, since its implementation.

It's still unsure whether this substitute is the right one to be applied, even though there are several examples of European cities where the results have been positive and, therefore, there has been a gradual growth of the system to include even more stations and bicycles due to its use (S. Shaheen et al., 2012). Public bicycle systems can be characterized as emerging initiatives that facilitate the use of bikes within urban environments, utilised as a one-way means of daily mobility or as a component of the public transportation network existent since it differs from traditional bikes thanks to its easy access and simple automated renting procedure made with a smart card or smartphone.

GIRA program, which is Lisbon BSS, has grown in scope by adding more stations and bicycles to its user base, since its introduction in 2017 (Vasconcelos, 2021). This evolution can be characterised due to the increasing recognition of cycling as a way of commuting as well as the increasing desire for new modes of transportation since bicycles don't release harmful emissions that pollute the atmosphere, making it a perfect and sustainable form of transportation (Hunt & Abraham, 2007).

Although Lisbon's bike-sharing system has made significant progress over the years there are still opportunities to improve the program's effectiveness. For that reason, understanding the behaviour patterns regarding BSS utilization and enhance the system's performance based on data insights are two major challenges.

To address this issue, it is essential to use key performance indicators (KPIs) and different visuals. KPIs evaluate and provide a thorough picture of the effectiveness and performance of the program into measurable metrics. Additionally, it enables decision-making based on the data gathered to optimize the distribution of resources and improve user experience in general. Visualizations, such as charts, graphs, and maps offer a visual representation that aids in interpreting data trends and patterns.

The main objective of this artifact is to give a proper view of the data available through Power BI dashboards using specific KPIs to improve the efficiency and user experience of Lisbon's bike sharing system, as GIRA is evolving daily into becoming a reliable transport substitute due to the goals defined in MOVE2030 (Machado et al., 2020). By setting specific and achievable objectives ensuring the promotion of sustainable urban mobility, optimization of the bike sharing system in Lisbon, and its impact on the city's transportation ecosystem, this analysis intends to analyse the behaviour of GIRA BSS users in recent years, as well as the total number of bikes for improving the performance of the BSS scheme in Lisbon.

Simultaneously, there is an absence of studies focused on Gira's program and its characteristics, requiring the conduct of research to fill this void and produce insightful data. It is crucial to identify and address these challenges to enhance the efficiency, effectiveness, and reliability of Lisbon's bike-sharing system – GIRA. An alternative way of mobility which also contributes to a healthier lifestyle and a decrease in traffic and, therefore, the dependence on private vehicles, which tend to be the major cause of air pollution in big cities.

When deciding which modes of transportation to use, the price comes as one of the critical success characteristics. For lower-income groups and young people, the price to pay by using public transports is certainly a reason to take into account (Fishman, 2016).

The accessibility and cost of the Lisbon bike sharing program are other success factors. With a modest membership fee and reasonable daily fees, the system is intended to be usable by everyone. Due to this, it has been a well-liked choice for locals and guests, which has increased its high rate of utilization. Yet, the city council implemented free transportation to students until 23 years and citizens over 65 years old, with a residential address in the city, so that, mobility by bicycle starts being considered a real alternative for locals and the future generations (Moreira, 2022).

The distribution of bicycles across the city is an issue that most BSS programs come across in the first years of implementation (S. A. Shaheen et al., 2011). GIRA, when talking about this particular subject, has a way of transporting bikes to different docking locations by knowing where to place them, due to the associated technology prediction on demand. On top of that, GIRA also has a points system for rebalancing bicycles, which rewards users who return the bicycles, after the end of the journey, in stations with less than 30% occupancy of parking places. In turn, to balance very full stations, the user is also rewarded with points if picks a bicycle in a station with more than 70% of the docks occupied. These points are then used to discount the fares of the 2nd and following periods of use and can be accumulated for 2 years (EMEL, 2021).

To keep up with the growth of the system, providing favourable conditions for all bicycle users in terms of cycling safety, Lisbon's existing bike paths have been constantly reinforced through investments made by the city council.

Since 2017, when the Gira mobility system was implemented, the cycling path has almost doubled from 76km to 156km. The goal is to reach 200km and, therefore, be able to move around Lisbon with a bicycle in a proper environment, that is designed for its use while providing safety and comfort for everyone who chooses the riding mode in response to the needs of its users (André, 2022b). It's important to complement the usage of bicycles with bike lanes or cyclable streets in order to meet the rising demand for riding bikes.

Below, Figure 3 shows the growth of bicycle lanes in Lisbon, over the years, since GIRA's system implementation.

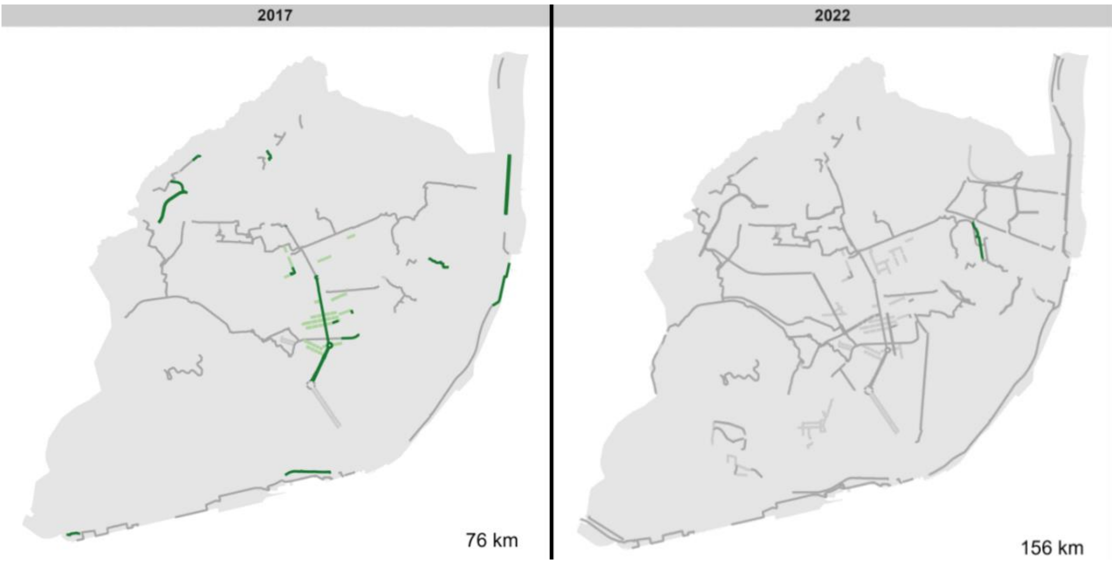


Figure 3 - Bicycle Lanes in 2017 and 2022 (source: https://shiny.desv.io/ciclovias/)

### 3.2.2. Data Understanding:

#### 3.2.2.1. Data Collection

This data is composed from the GIRA bicycle network in Lisbon city, therefore, is related to the location of stations, information relating to the station's capacity, bicycle trips and information more detailed about trips, automatic bicycle counters, and their location, and a map showing the bicycle path network in Lisbon.

When talking about improving the performance of an actual BSS, already developed, the data gathering was easier due to its current functioning and the existence of an open data source with data regarding Lisbon's PBSS. The possibility of acquiring most of the necessary data in a combined open source with all the information from GIRA BSS makes it easier for the later procedures by having the same data sources for all the dimensions.

The thesis's initial phase consisted of searching for relevant and usable data. In this particular case, this step was not a concern due to EMEL's open data source, which everyone can access, without restrictions. In that way, data was collected from EMEL open data source, and later on conducted to some elimination process and transformation as well, based on the information needed for this study, in order to have the data in a presentable form and easier to access. Simultaneously, historical bicycle data was collected from Lisbon City Council's open data source.

This was a simple process, thanks to the accessibility that the system provides for downloading the information required. At the same time, there was a specific use of python for downloading and combining the data from the bicycle's counters spread around Lisbon, where the parameters of an

API were continuously updated to provide daily information since the system was too slow to combine more than one day in a worksheet. Python was also used to transform the data and group all the observations into hours to provide a clear observation. The download made from the API is related to the counters data which was the only different method of gathering all the data needed. Is important to highlight that all the counters placed throughout Lisbon count not only bicycles but also vehicles that run on two wheels considered as micro-mobility devices.

However, bicycles exhibit significantly greater utilization compared to other forms of two-wheeled transportation, therefore, this analysis will suppose that all its counts are regarding bicycle activity from the GIRA bicycle sharing system, which ultimately provides specific information about which routes do users most use and the total number of cyclists crossing bike lanes in real-time.

Power BI is the tool where it all happens because it allows users to connect data, and create dashboards and reports, making it much easier to understand and communicate large amounts of data. It gives the opportunity to create an accurate analysis of the problems and solutions to be presented while providing real-time or near real-time information, bringing together data from various sources and making it appellative to viewers.

The data is analysed between July 2021 and March 2023, in order to make a comprehensive analysis about the performance of the system, so as to understand the growth that GIRA has had in the last years, but also to compare the several seasons' riding patterns as user behaviour patterns. The time set for analysing the data is restricted by data scarcity. At the same time, the data is as recent as possible because it contributes to a real and detailed analysis of how the system is currently working. As there was a data limit lower than one year for each invocation made on the open data source, various downloads were made and combined, so that all the data was in place for a good pre-processing. Most of the files were downloaded as a JSON file, within the date format from 08/07/2021 to 31/03/2023, combined and loaded into Power BI so that a proper analysis with interactive visuals in dashboards was made. On the other hand, the data from the python transformations were introduced into Power BI by a CSV file. Is important to highlight that the data from the bicycle’s counters are only available from 08/07/2021, which is the reason why it starts on that specific day.

Table 1 - Data Collection Table

| Data Source | Data Information |
| --- | --- |
| API | Bicycle Counters Data |
| Open data EMEL | Bicycle Counters Location |
| Open data EMEL | Gira Statistics |
| Open data EMEL | Gira Stations |

#### 3.2.2.2. Data Description

This dataset covers a substantial period of one year and a half, providing sufficient data for analysing the BSS over an extended period of time. This timeline allows for a more comprehensive understanding of seasonal patterns and long-term trends to enhance the operational efficiency and the overall performance of the bike sharing system. The reason behind the choice of this date range can be explained in two different ways:

First of all, data regarding the bicycle counters is only available since 08/07/2021, and at the same time data regarding the usage of bikes is only available until the first quarter of 2023. This being said, there's no possibility of adding data to the analysis as it would not provide enough information to be conclusive.

Secondly, before 2021 COVID-19 was still present in our daily routines and if those values were added to the analysis it may lead to less accurate conclusions due to the unprecedented impact the pandemic had on various aspects of society, making the data potentially outdated and unrepresentative of current conditions.

Still, this makes the information as fresh and updated as viewers would like to see it. The fact that the data used in this study is the most recent one, provides a more accurate understanding of the user behaviour lately while ensuring relevance and accuracy. Analysing the most up-to-date GIRA data contributes to staying current for future adaptations. Yet, the dataset regarding information from GIRA stations contains a much smaller number of stations than those that are active today. This analysis is therefore compromised from the very beginning by the scarcity of data.

As this dataset covers the daily usage of bicycles around Lisbon, almost no missing values were found in the dataset regarding the GIRA usage. Therefore, not many adjustments were made in the process of analysing the data to have a better understanding of the transformations needed to apply later on. Is noticeable that the data collected has already suffered some transformations and alignments before being available for anyone to access, use, or share.

One of the biggest limitations that this dataset has is the lack of information from users due to privacy laws and restrictions. The scope of this research is constrained by the absence of user-specific data, particularly when analysing user preferences or behaviours. For example, it's impossible to identify patterns in bicycle usage on aspects like age groups or genders because that information is not shared, due to data privacy. Therefore, it becomes challenging to make any presumptions about the characteristics or preferences of different user groups, which would ultimately provide a better overview of the system. It's important to highlight that due to this limitation, the overall analysis won't be as sharp as if user data were available to analyse. Personalization heavily relies on user data for gaining insights into patterns, trends, and correlations. It allows a deeper understanding and enables more accurate and targeted information, which potentially hampers the ability to draw meaningful conclusions.

Furthermore, the assumption that all the data regarding the counters is only from bicycle usage also constrains the accuracy of the final work. As said before, although most of the occurrences are made by bicycle users thanks to the overall perception towards this way of transportation, it is more difficult to provide accurate statistics as the precise percentage associated with the use of bicycles is unknown.

Regardless of these limitations, the dataset still provides important information on trends in bicycle usage, station popularity, peak times, popular routes, and system performance as a whole, which are going to be the aspects analysed through this artefact. With the use of interactive visuals, data is easily understandable and enables easy access to valuable information that might otherwise be difficult to see.

In the tables below is possible to have a clear view of the data as they provide information about each row of the dimensions downloaded. Some minor transformations of the columns were already applied thanks to the pointless information it had, as said before:

Table 2 - Trips dataset example

| Date | Avg. Trip Seconds | Avg. Trip Seconds Rush |
|---|---|---|
| 19/04/2022 | 818,824621 | 822,383449 |

Table 3 - Counters Location dataset example

| ID Counter | Postal Code | Latitude | Longitude | Location |
|---|---|---|---|---|
| 5 | 1700-063 | 38,75657505 | -9,146406395 | Avenida do Brasil - Campo Grande |

Table 4 - Usage dataset example

| Date | Bike Use per Day in Seconds | User Quantity | Avg. Trip by User in Seconds |
|---|---|---|---|
| 01/06/2021 | 4424481 | 2535 | 1745,357396 |

Table 5 - Stations Location dataset example

| ID Station | Postal Code | Latitude | Longitude | Location | Opening Date |
|---|---|---|---|---|---|
| 220 | 1100-150 | 38,70878 | -9,137125 | Rua do Comércio | 31/01/2022 |

Table 6 - Counters Detections dataset example

| Date | Time | ID Counter | Total Nº | Direction |
|---|---|---|---|---|
| 23/06/2022 | 03:00:00 | 30 | 4 | 1 |

Table 7 - GIRA dataset example

| Date & Time | ID Station | Location | Nº Bicycles | Nº Docks |
|---|---|---|---|---|
| 01/01/2022 13:38:03 | 102 | Portas do Mar | 2 | 14 |

### 3.2.3. Data Preparation:

Data Preparation addresses the format that will later be used for the modelling so that a proper dataset is built. This phase also consists in selecting, cleaning, constructing, integrating, and formatting data multiple times until it reaches the criteria needed to provide a dataset capable of progressing to the next stage (P. Chapman, 2000).

After understanding which data will be used for this project, it's time to address the form in which the data will be organized. Along these lines, different steps of formatting data were applied with the aim of having a clean dataset.

#### 3.2.3.1. Data Cleaning and Transformation

After all the data was gathered and connected to Power BI, a proper process of removing incorrectly formatted or incomplete information was carried out to enable analytics. It is noteworthy that most of the information was already cleaned when the downloads were made, either way, a second review was conducted in Power Query because it is imperative to comprehend both the content and structure of the data. Power Query is an instrument that empowers users to extract, transform, and load data from many different sources. It is considered as a powerful tool for data preparation and analysis tasks by handling data efficiently. Therefore, most of the processes of cleaning and transformation were carried out in this ETL tool, although Excel also contributed to this process.

Duplicated values can compromise the integrity of analytical outcomes, therefore action was taken to address the problem and remove duplicated records from the dataset to improve data quality. Dashboards created with raw data are less reliable and cannot be used to make informed decisions or provide valuable insights. Power Query has the ability of removing data by simply selecting which columns in a table that are going to be addressed.

Following, null values and some “0” values that appeared on the dataset were handled differently. While some values were deleted because didn’t interfere with the goal of this thesis, others were carefully replaced. The dataset had some empty values regarding counter detections. Some of the values were single days while the other scenario was a range of consecutive days, longer than a week which ended up completely altering the outcome of the data for that same month. The month in question is August 2022 between the dates of 20 and 28 August. Therefore, these values were fulfilled with the average value of the previous two weeks, regarding each weekday. Single values were not addressed due to not knowing if the system had an error or if it was a consequence of unfavourable conditions. These values were carefully changed in Excel through a simple function. This analysis conducted on the behaviour of bicycles around Lisbon generates better insights, although not being as accurate and precise as if the data were complete.

A time filter was applied to “Fact_GiraData” in order to have only the values regarding the timestamp defined by the system which allows users to unlock bicycles from stations. The filter chosen comprehends only values that are from 06:00 am until 12:00 pm so that a proper visualization of the number of bicycles available per station would be reliable. The misleading values would include data from hours in which the system doesn’t function for renting, only for dropping bicycles, and data regarding the repositioning of bicycles made by the entity itself.

Simultaneously, a focus on data transformation was dealt with more precaution, hence the importance of converting some features into more detailed and easily accessible to reach when needed and to provide a more detailed information sharing making it more accurate and precise to provide support decision making.

Following this scenario, there was a need to create a Date Dimension to organize the data to compare it across different periods, while enabling the use of time filters on the reports. Adding the preference of analysing by quarter, month, day of the month, and day of the week, tends to be more accessible as a way of assembling the data and comparing it in different periods. However, there’s a need to have a real-time understanding of the data in order to make real assumptions and provide valuable information. At the same time, a dimension was also created to assemble all the data regarding Lisbon parish to provide a more comprehensive analysis when it comes to bicycle usage around the city, this dimension is called “DIM_Parish”. On the other hand, in order to have all data related to trips and usage by bicycles together both tables were merged into one named “Fact_Trips&Usage”. The need to make data more accessible to its use and to easily create other columns regarding its information was the main reason for this assemble.

Nevertheless, custom columns were also created to improve the accessibility of data. Below are some examples of the work done either in Power BI or Excel:

- Duration Columns: Different columns from two different tables showed the data in a total number of seconds. Therefore, the creation of columns with the values presented in the proper format of HH:MM:SS makes it easier to understand the overall time. A DAX was created by programming language.

- ID_DATE: Creation of column that gives the date format into YYYYMMDD so that a SK is defined to later link to other tables.

- Nº Bicycles Column: In order to retrieve the total number of bicycles used per day, a simple DAX expression was created to divide the total of “Bike Use per Day in Seconds” by the “Avg. Trip Duration”, giving an approximate figure for the number of daily journeys made by bicycle. This measure contributes to a more accurate analysis of the number of bicycles used per day while comparing it with the number of daily users of the system.

- DIM_Date Columns: Creation of different date columns relating to the specific time defined previously, subdivided into year, quarter (results showed in “Q1”), month, day, and day of the week. Fulfils the purpose of analysing different aspects of date.

- Parish Column: Creation of a column that helps identify the movement of bicycles between parishes more concretely to provide a more detailed analysis related to location usage.

- Custom Columns: To organize data and split columns, various transformations were done in different columns so that values like “Location” and “ID_Station” were separated. At the same, in order to get the proper number of the Postal Code, two columns were merged. This process was necessary to connect tables.

- Variation Column: By dividing the difference of each update regarding the number of bicycles in a station by the number of docks each station has, when sorting by date and time, it’s possible to create a percentage that shows the variation of each occurrence. This column was created in Excel through the use of a simple “IF” function and serves to inform the variation of bicycles at each station.

- Hours Columns: Creation of a "hours" column to analyse trends, patterns, or correlations based on the time of day of the table regarding information from GIRA bicycles.

### 3.2.4. Modelling:

Modelling phase consists in selecting the technique to be applied to the project (P. Chapman, 2000). Since it is linked with Data Preparation, is important to apply the right technique because there are several ways of doing it for the same data mining problem type, due to specific data formats. This includes the selection and building of the model as assessing it.

There's an urge to provide a visualization of the data collected from bike-sharing systems while analysing and interpreting in a Power BI dashboard. This step provides a different and easier approach to the data while giving an overall view and more comprehensive understanding.

#### 3.2.4.1. Data Model

After all the data has been pre-processed, it's time to develop a conceptual model that links all the tables with the data regarding to the BSS usage in order to provide information for the creation of the dashboards. After taking into account the information regarding the development, evolution, and use of BSS to have a broader concept of this subject, Power BI is responsible for the next steps from now on. In the context of this thesis, data modelling is the basic procedure that forms the core of our analytical methodology.

The choice of the schema was decided according to the accessibility of the values. In this way, a snowflake model was developed as Fig 3 shows. This snowflake model is composed of four dimensions, and three different fact tables named "Fact_Trips&Usage", "Fact_CountersData" and "Fact_GiraData" that link all the data together to facilitate its use in the next steps. Furthermore, this schema makes it easier to retrieve data when creating reports, hence this analysis is based on the fact tables and its dimensions. Each fact table has a primary key that links the information to the dimensions associated and each dimension has a surrogate key to fulfil the same purpose.

All three fact tables are linked to the "DIM_Date" by the creation of the attribute "YYYYMMDD", which makes it viable to reach the information needed from each table. Therefore, this new attribute created to connect the tables, makes it possible to develop dashboards with informative visuals.

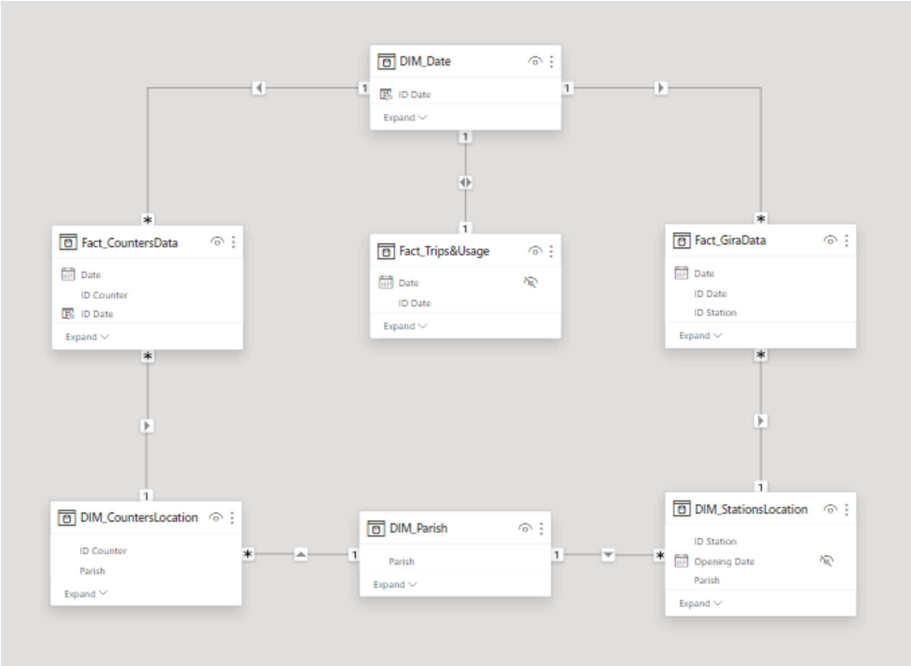


Figure 4 - Data Model

As Fig. 4 shows, each dimensional table needs to include a primary key that corresponds to a foreign key in the fact tables in order to link the information. In this sense, it is important to emphasise those connections to better understand the functioning of the model.

DIM_Date is connected to all the fact tables by the attribute ID_Date providing the ability to study behaviour and trends over a period of time. Following, DIM_Parish is linked with both location dimensions (DIM_LocationCounters and DIM_LocationStations) by the attribute "Parish" which enables a better understanding of the data by parish. At the same time, DIM_LocationCounters is linked with the fact table that has information regarding counters by the attribute "ID_Counter" as DIM_LocationStations is also linked with the fact table that has information relating to GIRA bicycles by the attribute "ID_Station", which gives the ability of analysing data by each counter/station. Therefore, "Fact_Trips&Usage" is the only fact table to have only one connection to other dimensions.

### 3.2.4.2. Definition of Dimensions and Measures:

A brief explanation of each dimension and fact tables will be explained in this step as the attributes that make them up. At the same time, a visualization of each dimension is displayed below the explanation of each table and an example of each table rows can be found in the appendices from Fig. 36 to Fig. 42, respectively.

**Location of Counters:**

“DIM_LocationCounters”, exactly what it's called, has information regarding the location of the counters in the city of Lisbon. Therefore, it contains the ID Counter and its identifying number, Street Name, Latitude and Longitude, Postal Code, and Parish.

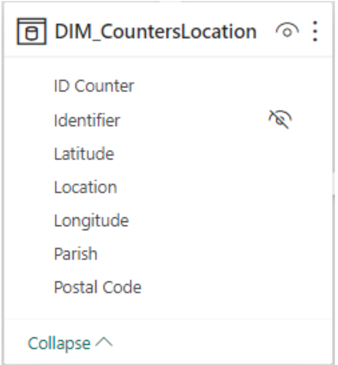


Figure 5 - Dimension Location Counters

**Stations:**

The station's dimension is shown as “DIM_StationsLocation”. The table has information regarding the location of each station such as their identification numbers. Thus, it contains ID_Station, Location (name of the street where is placed), Latitude and Longitude, Postal Code, and Parish. The goal of this dimension is to connect the data from the GIRA stations with its location. A column named “Parish” which provides information on the parish of each station, was created to better comprehend cycling behaviour between parishes.

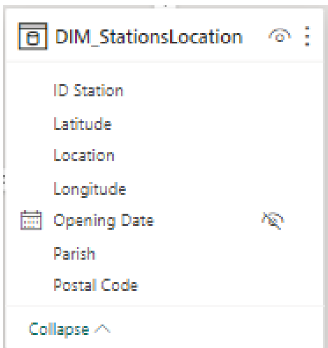


Figure 6 - Dimension Station Location

**Parishes:**

This dimension was developed with the purpose of linking both tables of data regarding counters and stations by a mutual connection. Therefore, an analysis through Lisbon parishes gives a more detailed perception about areas where bicycle usage has made an impact. It contains information regarding the name of the city, each parish's name, and their area size. It is connected to "DIM_LocationCounters" and "DIM_LocationStations" by the attribute parish.


Figure 7 - Dimension Parish

**Date:**

This dimension gathers a continuous set of dates covering the time frame needed for the analysis. The identification of "DIM_Date", includes various ways of seeing the date by filtering it by almost any date range. In this way, contains ID_Date, Day, Day of the Week, Month, Quarter, Year. Nevertheless, it also contains a Date hierarchy with Year, Quarter, Month, and day in this exact order.

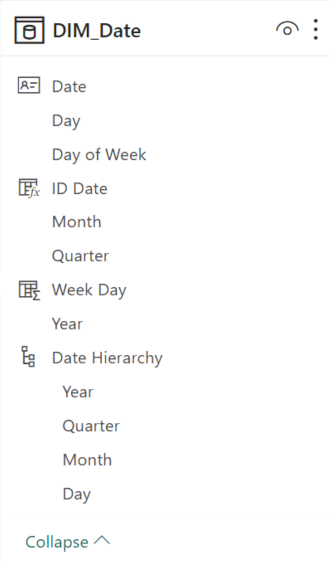


Figure 8 - Dimension Date

**Data regarding Counters:**

Also, “Fact_CountersData” resembles to its name. It holds Date, ID_Date, ID_Counter, the direction of the path (either 0 or 1), the total number of bicycles, and the hours regarding each counter. As said before, all detections within one hour range were clustered to have a clear perspective of which hours have the biggest peak. This table has almost 950,000 rows of information regarding cycling activity and is linked to “DIM_Date” by ID_Date and to “DIM_CountersLocation” by ID_Counter.

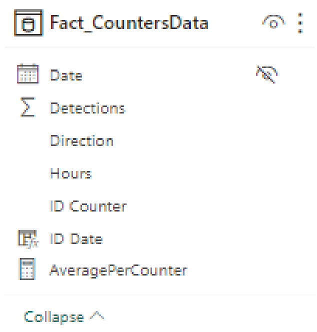


Figure 9 - Fact Table Counters Data

**Data regarding Gira Usage:**

This fact table has a huge amount of data, being composed by more than 2 million rows. It has information relating to ID_Date, Date, Time regarding each occurrence, ID_Station, Nº Docks each station has, the number of bicycles in a station regarding each occurrence, and a variation of the number of bicycles in each update. This table is linked with “DIM_StationsLocation” by ID_Station and to “DIM_Date” by ID_Date. The “Time” attribute was filtered to contain only values from 05:55 until 00:00 due to the timeframe defined by GIRA to unlock bicycles from the system.

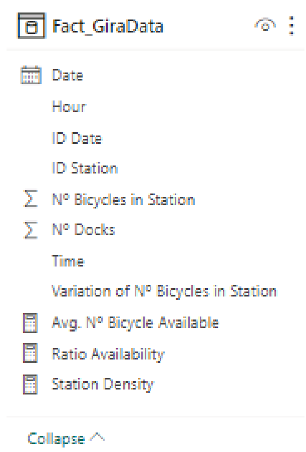


Figure 10 - Fact Table Gira Data

**Trips & Usage:**

This fact table contains information about the number of users, bicycle usage, and bicycle usage by user. The fact that it contains information regarding the average time taken per trip as the average time taken by a user with a bike when considering a daily perspective. This table is entitled by the name of “Fact_Trips&Usage” because it’s a merge of both tables downloaded previously (“Trips” and “Usage”). Therefore, it has information regarding ID Date, Date, User Quantity, Daily Bicycle Use, Average Duration of Trips by User, Average Trip Duration, and Nº of Daily Trips.

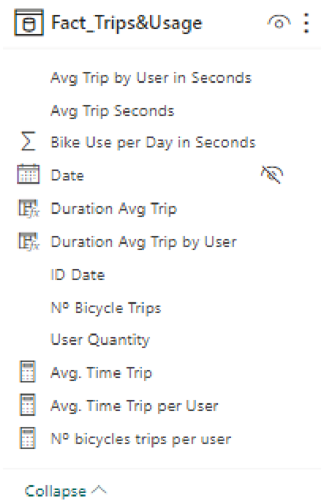


Figure 11 - Fact Table Usage & Trips

**Hierarchies**

Defining hierarchies in Power BI is essential as it enables users to explore data from a variety of viewpoints which ultimately allows a better understanding of the information regarding the ability of drilling up or down different levels. When analysing Date Hierarchy, we have a calendar time with a depth four. This helps comparing different journeys through time, getting a proper notion of time, while also being able to get a piece of more detailed information about each journey made. Fig. 12 shows the hierarchy built in this data model.


Figure 12 - Date Hierarchy

**Measures Creation**

Power BI enables users the flexibility of creating custom measures to perform calculations and aggregate data within a data model through DAX language. Thus, measures implement elaborate operations that would be challenging to carry out using the default Power BI visuals. In this sense, several measures were developed to increase the quality of the data without being stored in the dataset. Following, are some examples of measures created in the report part to later be applied in the report part:

- Average Nº Bicycle Available: Gives the average number of bicycles each station has.

- Average per Counter: Displays the average value of detections a counter makes.

- Average Time Trip: Provides the average duration a trip takes.

- Average Time Trip per User: Gives the average duration a user takes when riding bicycles in a day.

- Nº bicycle trips per user: Regarding the number of trips done, a simple division of the number of bicycle trips by the number of users.

- Ratio Availability: Gives the average number of bicycles available in a station by dividing the number of bicycles in a station by the number of docks each station has.

- Station Density: Provides a distribution and concentration of bike-sharing stations across a given geographical area, in order to understand the overall distribution of stations per parish.

### 3.2.4.3. Dashboard Creation

A dashboard is a *visual display of the most information needed to achieve one or more objectives which fits entirely on a single computer screen so it can be monitored at glance*, according to (Few, 2006). Dashboards are not significantly dissimilar from other methods of conveying information. When meticulously designed, a finely tuned single-screen display can deliver insights in a particularly impactful manner (Few, 2006).

The choice of using dashboards as a way of displaying data facilitates an extensive examination of the information and visually communicates performance metrics while using different visualisations according to its attributes. The implementation of reports through the combination of visuals improves the decision-making and the evaluation of BSS performance to determine if the system is operating at its best capacity or whether modifications are required to increase its effectiveness.

This strategy is a useful tool for measuring success, as it provides valuable insights in a clear and succinct manner that ensure the long-term viability of the system and identify actual or future deviations. The distinctive feature of dashboards lies in their capability to consolidate and show diverse data sets in a unified, easily comprehensible visual format. This enables users to swiftly comprehend complex information, facilitating well-informed decisions.

To optimize the presentation of information and enhance user comprehension it is essential to follow an organised framework for all the reports done. Therefore, all the reports displayed below in Chapter 4 follow the same transition, highlighting comparison visuals between different categories of data and correlations between distribution across the performance of both counter and data from stations over time to identify trends or changes. Nevertheless, the layout of the reports was created with the aim of providing data in an enlightening way, following a line of thought that allows for extensive analysis of each chosen graph. Each report contains slicers that allow you to change the date of the observation as well as slicers that allow you to get a more concrete understanding of each location or station to be dealt with. The way the visuals are laid out results from the inclusion of main KPIs at the top to prioritise small-scale visuals, larger visuals in the middle to demonstrate trends over time in a detailed way, and finally granular metrics and maps to understand the variability of the results.

KPIs enable the measurement of system performance by providing quantitative metrics. This allows for a comprehensive evaluation of the system's effectiveness over time. It also acts as an objective measure for evaluating the performance of BSS concerning the goals defined, which eventually makes it easier to determine if the system is operating at its best capacity or whether modifications are required to increase its effectiveness in a simple and short visual. It’s important to have different ways of analysing the data coming from the usage of GIRA system since it must ensure the long-term viability of the system and provide a comprehensive analysis that allows you to conclude pertinent information without getting confused.

### 3.2.5. Evaluation:

Evaluation embodies the revision of the model, the validation, and the certification that the model meets the business objectives defined earlier (P. Chapman, 2000). In other words, provide an overview of the process and emphasize key initiatives.

In this part, an overall evaluation of the artefact is done while addressing the objectives and solutions defined previously. The result of this phase helps understanding and better explaining the further steps needed, but also the behaviour of users. By analysing patterns that contribute to a better understanding of Lisbon’s BSS performance and its results, with the help of Power BI dashboards, information becomes visually appealing and interactive in many different ways.

This section must contribute to existing information regarding the use of BSS around well-developed cities while giving a realistic view of the behaviour patterns by its users.

### 3.2.6. Deployment:

Lastly, the Deployment is focused on giving reports about the data analysed previously in order to share the knowledge acquired with viewers (Wirth & Hipp, 2010). Works as a summary of all the previous steps and organizes the results in a way that makes use of the created models.

This last step is focused on giving the proper conclusion of how the plan was executed and to provide a clear overview of the main aspects learned along the research plan. It makes it possible to determine whether the artifact had an impact, to identify and discuss any limitations found and even to propose potential solutions to the existing situation where the main objective is to encourage as many people as possible to use this means of transport so that they consider it an alternative to active transport for daily commuting.

# 4. RESULTS AND DISCUSSION

In this section, a demonstration of the culmination of the research efforts done in the previous steps ables the creation of detailed dashboards and reports. Thus, by employing visualization techniques that express the patterns and insights obtained from the intensive data analysis, a visual narrative that complements the exploration of the data is provided.

The creation of the dashboards and its visuals allows us to conduct a detailed analysis of the GIRA usage and produce insights that wouldn't be understood so easily. This specific way of presenting data is a key differentiator when giving more intel about the work done.

Following, a story through visualizations will be narrated to provide an accurate vision of the performance GIRA had lately and its overall perception by users. Different aspects are highlighted to have a broader knowledge of this subject. At the same time, it is also possible to drill down in some visualizations as well as choosing the specific timeline or location to have a broader knowledge about each report. All reports are interactive, and, therefore the results update depending on the parameters chosen in the slicers or visuals.

After all, four dashboards were developed and organized in the following way:

- Station's Observations - Assessing the performance and demand for individual stations, as identifying underutilized or congested stations.

- Bicycle Counters Observations - Analyze the distribution and intensity of bicycle use in different bicycle paths as well as the overall use in different seasons and times of the day.

- Parish Comparison - Understand variations in usage patterns, infrastructure needs, and user behaviors across different parishes.

- Usage and Trips Behavior- Evaluate the overall perception of users towards this way of transportation while.

## 4.1. STATION'S OBSERVATIONS

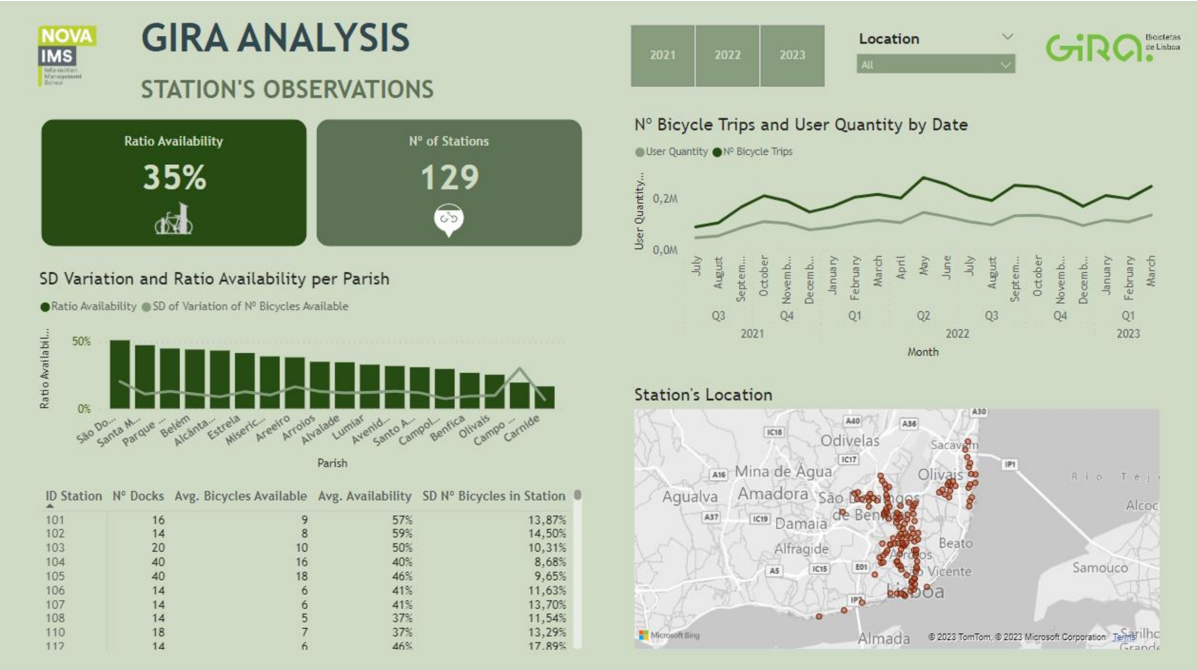


Figure 13 - Dashboard Station's Observations

This dashboard features information regarding the bicycle-sharing stations. Therefore, it shows the overall distribution of bicycle activity in stations across Lisbon, which helps to clarify the locations that are thought to be hubs for the use of bicycles while providing a specific approach to each station and its overall performance. The importance to analyse each station by its availability and variation of bicycles helps to have a better understanding of regions that have high demand or congestion while making well-informed decisions to improve the effectiveness and user experience while resource allocation.

The slicers located at the upper right corner of the dashboard's analysis area may be used to select the time frame for the study and a station's precise location can also be selected for analysis.

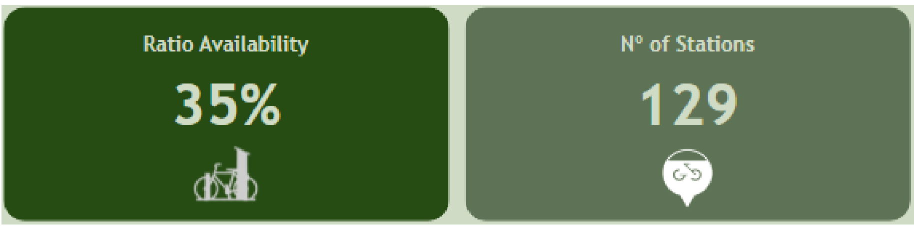


Figure 14 - Card Station's Observations

Fig 14 has the objective of indicating the total number of stations analysed in this study while giving an average availability of bicycles in the stations of the bicycle/sharing system. Every time stations are pinpointed, this card automatically updates to give precise information about one or more stations.

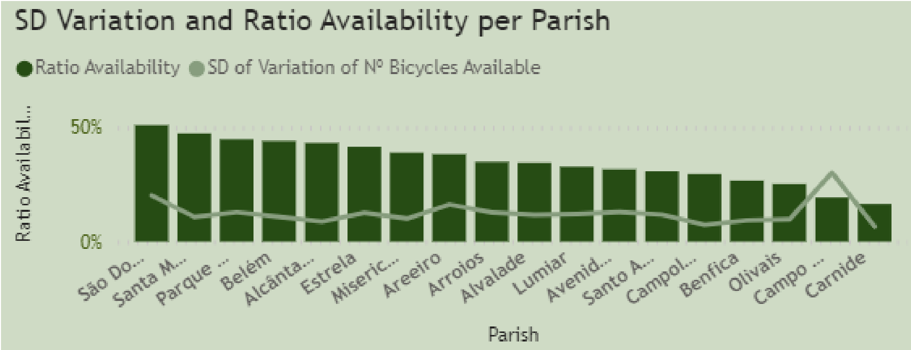


Figure 15 - SD Variation and Ratio Availability by Parish

Above, is possible to show the relationship between ratio availability and the number of bicycles that fluctuate across parishes. Indeed, there is a systemic trend highlighted by the line that shows the variance in bicycles, which stays mostly constant in all parishes. Yet "Campo de Ourique" sticks out with a larger percentage fluctuation.

The bars representing the ratio availability unveil distinct characteristics among parishes. "São Domingos de Benfica" emerges as a standout performer, boasting the highest availability with almost 50%. This indicates a robust and well-maintained bike-sharing infrastructure in this area or low usage by its users, knowing that it only covers one station and is more isolated when compared to others. At the same time, this station only has ten available docks, meaning that the slight variation significantly affects the results. Conversely, "Estrela" and "Campo de Ourique" exhibit lower values, falling below 20%. Given that "Campo de Ourique" also demonstrates higher variation, it prompts a closer examination of the factors influencing availability and the potential impact on user experience, although only having two stations placed in end-points. It is important to highlight that these three parishes have a limited number of stations which ends up not showing an overall result, but an individual analyses. Still, "Santa Maria Maior," which has various stations, gets the second-highest availability number, indicating that the system is operating in good conditions in this specific parish.

| ID Station | Nº Docks | Avg. Bicycles Available | Avg. Availability | SD Nº Bicycles in Station |
|---|---|---|---|---|
| 101 | 16 | 9 | 57% | 13,87% |
| 102 | 14 | 8 | 59% | 14,50% |
| 103 | 20 | 10 | 50% | 10,31% |
| 104 | 40 | 16 | 40% | 8,68% |
| 105 | 40 | 18 | 46% | 9,65% |
| 106 | 14 | 6 | 41% | 11,63% |
| 107 | 14 | 6 | 41% | 13,70% |
| 108 | 14 | 5 | 37% | 11,54% |
| 110 | 18 | 7 | 37% | 13,29% |
| 112 | 14 | 6 | 46% | 17,89% |
| 113 | 14 | 6 | 43% | 12,99% |
| 114 | 10 | 5 | 46% | 13,02% |
| 115 | 18 | 8 | 46% | 13,67% |
| 130 | 20 | 4 | 18% | 7,02% |
| 131 | 21 | 4 | 19% | 9,18% |
| 132 | 25 | 4 | 17% | 6,38% |
| 133 | 18 | 4 | 24% | 8,65% |
| 134 | 20 | 5 | 27% | 7,67% |
| 135 | 20 | 9 | 43% | 13,52% |
| 137 | 33 | 7 | 23% | 6,91% |
| 139 | 25 | 6 | 23% | 8,42% |
| 150 | 17 | 5 | 30% | 11,22% |
| 151 | 24 | 7 | 31% | 15,40% |

Figure 16 - Matrix Table Station's

The matrix table provides a comprehensive overview of the metrics for each station within the bike-sharing system. Therefore, it includes essential information such as the total number of docks, the average number of bicycles available per station, and the ratio availability, offering insights into the station's capacity and accessibility. Additionally, the table features data on the average variation of bicycles for each update, providing a dynamic perspective on how bike availability fluctuates over time. This figure makes it easier to see how each station functions concerning the overall bike-sharing network.

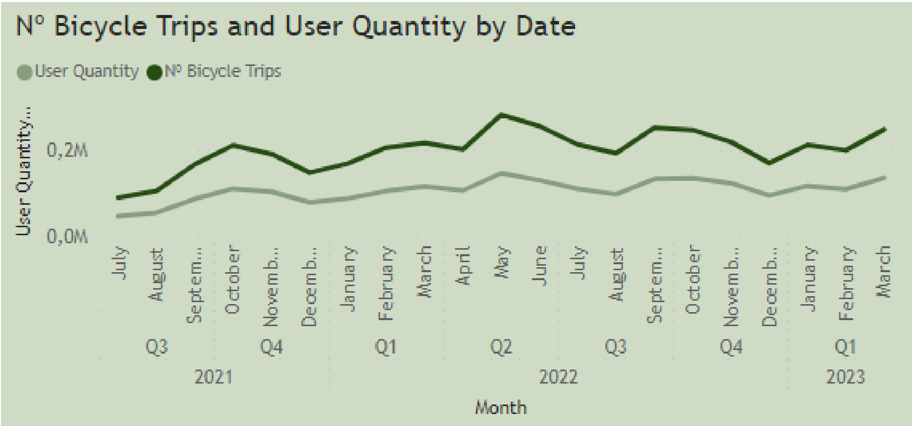


Figure 17 - Nº Bicycles Trips and User Quantity by Date

The line chart of Fig.17 shows that the bicycle trips and user number are aligned consistently, indicating that both features grew simultaneously over the period under observation. Interestingly, there is a slight decline in the number of users and bicycle journeys in the months of August and December. Yet months like May, September, and October continuously rise, indicating higher user engagement and more bicycle trips during these months. Furthermore, there is a noticeable rise in the number of users and bicycle journeys between the same months in subsequent years, highlighting the system's ongoing popularity and expansion over the period in analysis.

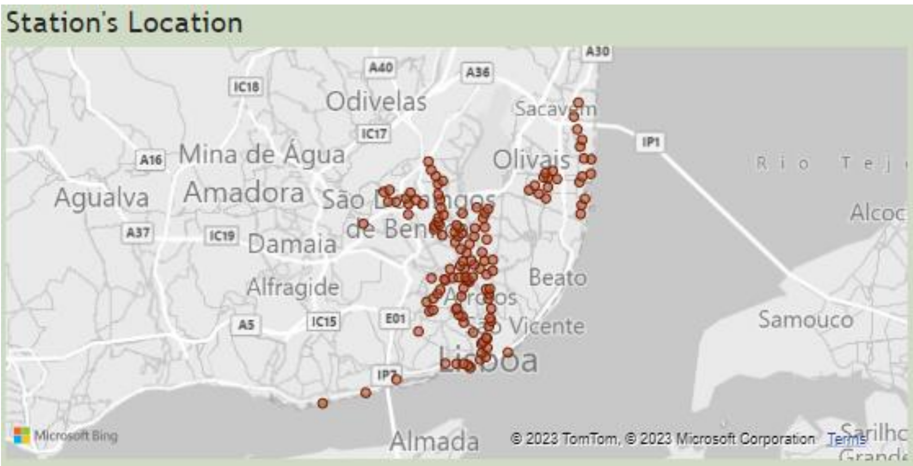


Figure 18 - Station's Location

Figure 18 is a visual representation of the station's distribution across Lisbon that converts geographical data into an engaging and informative visual story. This dynamic graphic adds to a thorough examination by offering subtle insights into the exact positions of each station. The map shows a concentration of stations in the city center and a significant cluster in Parque das Nações and Olivais regions. The identification of clusters that correspond to the main bicycle activity zones in our study improved by this spatial display, which also helps to identify strategic trends and possible areas for system optimization.

## 4.2. Bicycle Counters Observations

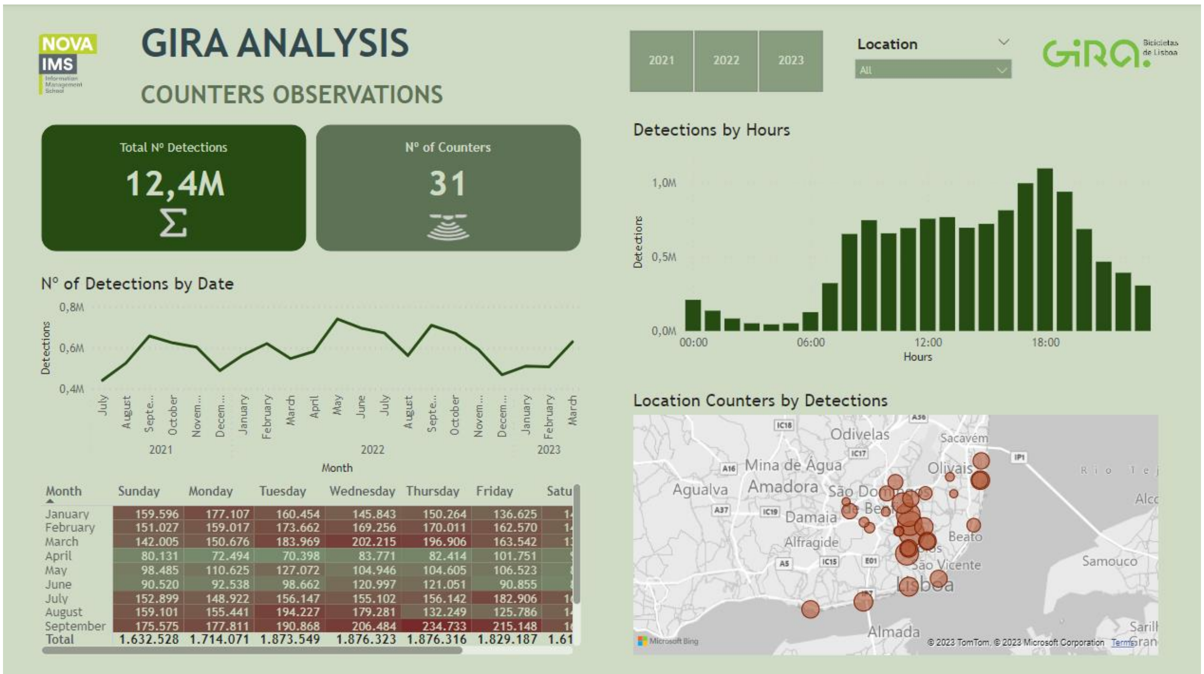


Figure 19 - Dashboard Counters Observations

Comprehending the bicycle traffic's flow is like reading a city's sustainable heartbeat regarding urban mobility. Beyond the perspective of the user, this dashboard explores the factual data collected from counters located across Lisbon. Every detection is a data point and, therefore, the silent symphony of pedal revolutions is captured by this dashboard, offering vital insights into the dynamics and patterns of bicycle movements across metropolitan areas. This dashboard offers a comprehensive understanding of the interactions between cyclists and the developing bicycle-sharing system by providing a temporal analysis, including years, days of the week, and hours, to comprehensively explore patterns and trends over time.

At the same time, it is possible to choose the period for this study using the slicers at the top right of the dashboard's analysis section as well as choosing a specific location of a counter to analyze.

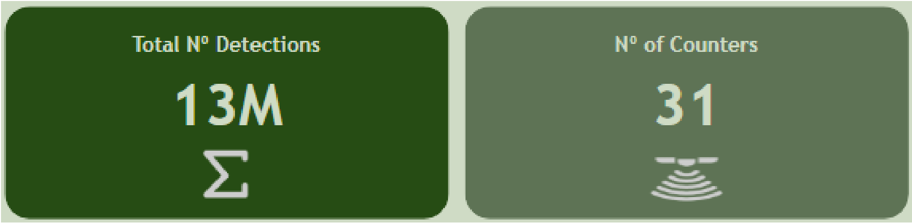


Figure 20 - Cards Counters Observations

This simple card provides information regarding the total number of detections between July 2021 and March 2023 and the total number of counters in this analysis.


Figure 21 - Nº of Detections by Date

The linear graph above shows the number of detections by date, providing information on the dynamics of bicycle movements over time while highlighting some noteworthy trends and anomalies.

The initial quarter of each year consistently exhibits a decrease in bicycle detections compared to other months in the analysis. Notably, the beginning of 2023 recorded a significant drop in detections compared to the general trend observed in the overall months under analysis. Contrastingly, the year 2022 demonstrated an overall growth in bicycle usage, with a steady increase observed across multiple months. It's important to acknowledge that the accuracy of data for August 2022 might be compromised due to data scarcity. This limitation should be considered when interpreting the results for that specific month. In Appendix A, Fig 43 shows the number of detections if the forecast previously explained to August 2022 had not been made.

May is also a particularly noteworthy month due to its spike in bicycle detections, which indicates greater urban mobility and use of the bike-sharing program. A noteworthy trend unfolds post-September, indicating a gradual decrease in the number of detections, whereas December reveals a notable decrease in bicycle detections, going beyond the constant pattern observed.

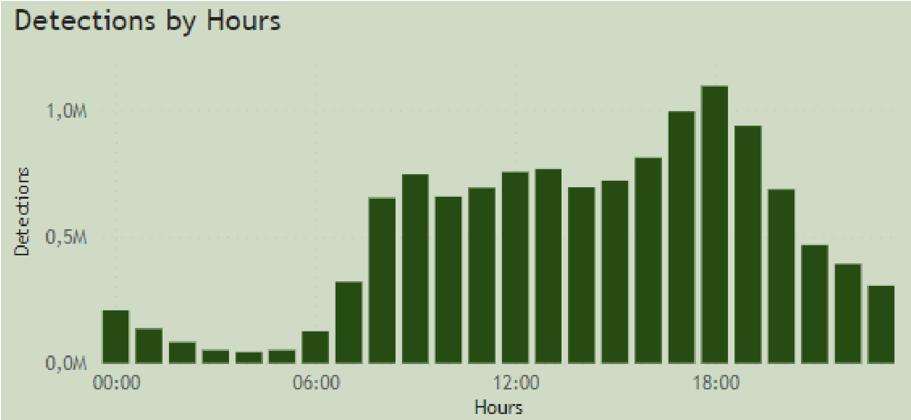


Figure 22 - Detections by Hours

The column chart portraying the number of bicycle detections throughout the day provides a compelling snapshot of the temporal dynamics of urban cycling. Notably, the hours between 17:00 and 19:00 emerge as peak periods. In contrast, the hours from 8:00 to 16:00 depict a consistent and steady line of results. The spike in detections during the late afternoon and early evening hours signals a significant boost in bicycle activity. The pronounced peaks during the evening hours underscore the significance of the bike-sharing system as a reliable mode of commute, particularly during rush hours as individuals navigate their way home from work or engage in recreational cycling. The consistent line observed between 8:00 and 16:00 underlines the versatility of bicycles as a mode of transportation throughout the day.

| Month | Sunday | Monday | Tuesday | Wednesday | Thursday | Friday | Saturday | Total |
|---|---|---|---|---|---|---|---|---|
| January | 159.596 | 177.107 | 160.454 | 145.843 | 150.264 | 136.625 | 146.540 | 1.076.429 |
| February | 151.027 | 159.017 | 173.662 | 169.256 | 170.011 | 162.570 | 141.907 | 1.127.450 |
| March | 142.005 | 150.676 | 183.969 | 202.215 | 196.906 | 163.542 | 137.450 | 1.176.763 |
| April | 80.131 | 72.494 | 70.398 | 83.771 | 82.414 | 101.751 | 91.682 | 582.641 |
| May | 98.485 | 110.625 | 127.072 | 104.946 | 104.605 | 106.523 | 88.809 | 741.065 |
| June | 90.520 | 92.538 | 98.662 | 120.997 | 121.051 | 90.855 | 81.692 | 696.315 |
| July | 152.899 | 148.922 | 156.147 | 155.102 | 156.142 | 182.906 | 161.681 | 1.113.799 |
| August | 159.101 | 155.441 | 194.227 | 179.281 | 132.249 | 125.786 | 140.969 | 1.087.054 |
| September | 175.575 | 177.811 | 190.868 | 206.484 | 234.733 | 215.148 | 167.347 | 1.367.966 |
| October | 179.193 | 189.089 | 192.597 | 178.506 | 181.179 | 194.897 | 179.844 | 1.295.305 |
| November | 139.030 | 166.544 | 204.301 | 190.498 | 170.354 | 182.529 | 142.667 | 1.195.923 |
| December | 104.966 | 113.807 | 121.192 | 139.424 | 176.408 | 166.055 | 135.303 | 957.155 |
| Total | 1.632.528 | 1.714.071 | 1.873.549 | 1.876.323 | 1.876.316 | 1.829.187 | 1.615.891 | 12.417.865 |

Figure 23 - Months per Weekday Detections

This table catalogs bicycle detections, with weekdays adorning the columns and months gracing the rows. It unveils the dynamic pulse of urban mobility and offers valuable insights into the patterns of bicycle usage within the city. From Monday to Friday, weekdays stand out as the dominant players in this urban symphony, portraying how bicycle users integrate this mode of transportation into their daily routines. Nevertheless, Monday still shows a certain difference when compared to the other weekdays. The concentration of detections on weekdays directs attention to the importance of tailoring services and interventions to align with workweek commuting behaviors. September and October months witnessed heightened bicycle activity. Despite only being counted once, May is a spotlight month because of the vast number of detections it contains. At the same time, both months of April and June are also counted only once. Lastly, August shows a difference in behavior patterns with much lower results on Thursday and Friday compared to the rest of the days of the week, despite the possibility that it is the outcome of data scarcity.

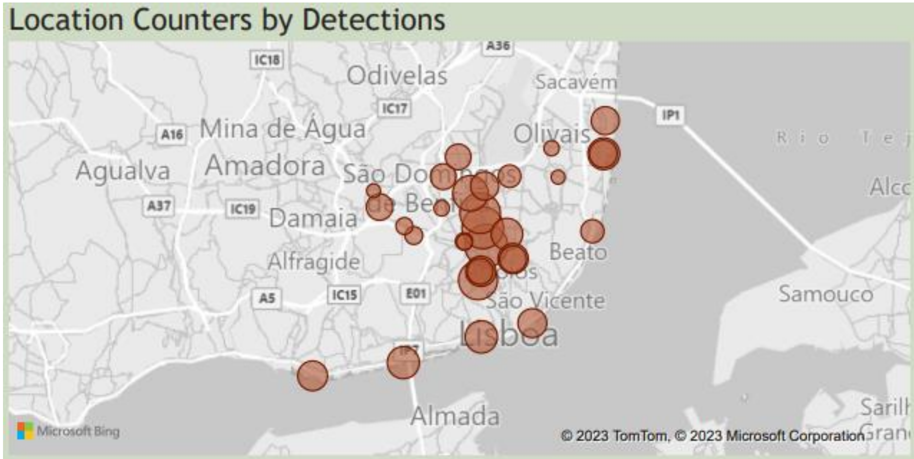


Figure 24 - Location Counters by Detections

Fig. 24 corresponds to a map showcasing the distribution of bicycle counters, where the size of each bubble corresponds to the number of detections, transforming spatial data into a dynamic and insightful visual narrative. This interactive representation offers nuanced insights into the intensity of bicycle activity across each counter of Lisbon. It becomes clear that locations in the city center and near the river tend to have higher bicycle detections.

## 4.3. Parish Comparison

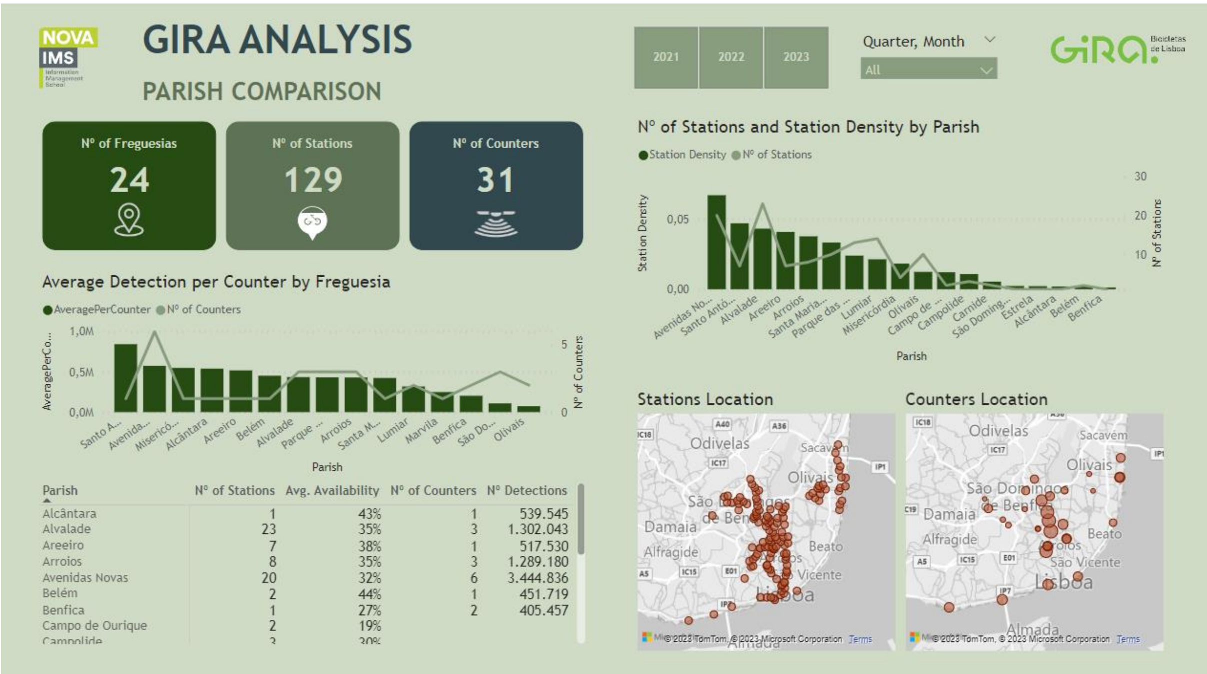


Figure 25 - Dashboard Parish Comparison

The main goal of this dashboard is to explore the interactions between counters and stations in Lisbon's parishes as cities increasingly embrace the significance of cycling for both commuting and recreation. Consequently, understanding the usage patterns aims to expose trends, inequalities, and opportunities, offering important perspectives on how urban transportation is developing in each parish while paving the way for informed decisions to enhance the efficiency and accessibility of bike sharing systems.

This dashboard contains two slicers in the top right corner for choosing the date to be observed depending on the purpose of the analysis.

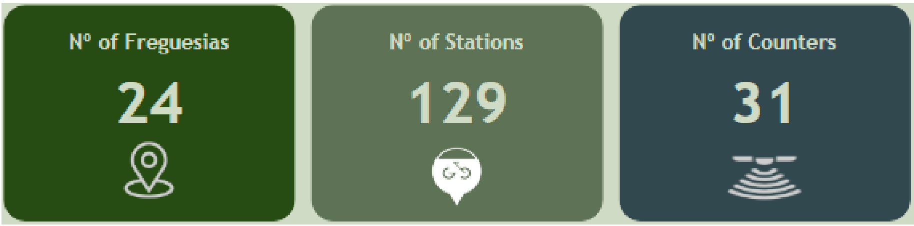


Figure 26 - Cards Parish Comparison

These first visualizations of this dashboard are simple cards that provide an overview of the number of parishes in Lisbon, as well as the total number of stations and bicycle counters to be analysed. It is a relatively basic metric that serves the purpose of indicating the number of stations and counters when a specific parish is selected. Whenever one or more parishes are selected, this card automatically adjusts.

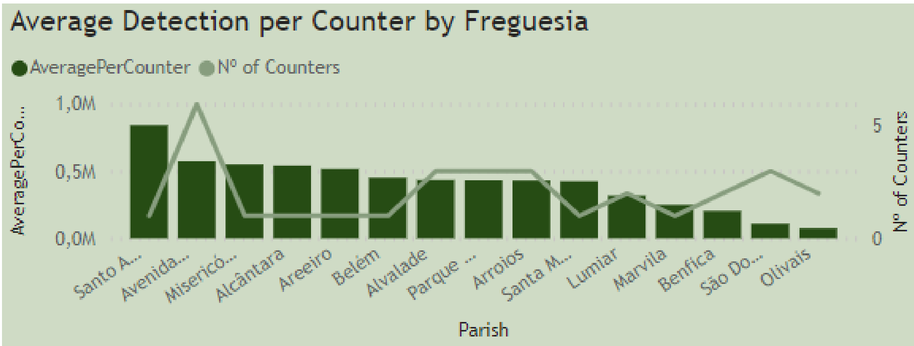


Figure 27 - Average Detection per Parish

This line and column chart illustrates the dynamics between different parishes concerning bicycle detections and the number of counters associated. "Santo Antonio" tops the table with the greatest average number of bicycle detections while only having one counter. This suggests that bicycle users tend to opt for the path where the counter takes place, and, therefore, shows a large affluence of users taking this route when on the move. On the opposite side, "Olivais" records the lowest average number of bicycle detections, although having two counters, which raises questions on the cycling volume activity in this particular parish.

Counters from the city centre and places near the river tend to have a higher number of detections when compared to suburbs (Marvila, Benfica, and São Domingos de Benfica).

While Alvalade, Parque das Nações, and Arroios show a compelling scenario having a high number of average detections per parish regarding three counters each, Alcântara, Areeiro, Belém, and Misericórdia also have the same number of detections when having only one counter each. At the same time "Avenidas Novas" also presents similar values having five counters scattered around the city. This disparity is remarkable and important to highlight in order to allocate resources effectively and empower decision-making as some counters are placed at opposite ends of the same street, to understand the movement of cyclists.

| Parish | Nº of Stations | Avg. Availability | Nº of Counters | Nº Detections |
|---|---|---|---|---|
| Alcântara | 1 | 43% | 1 | 539.545 |
| Alvalade | 23 | 35% | 3 | 1.302.043 |
| Areeiro | 7 | 38% | 1 | 517.530 |
| Arroios | 8 | 35% | 3 | 1.289.180 |
| Avenidas Novas | 20 | 32% | 6 | 3.444.836 |
| Belém | 2 | 44% | 1 | 451.719 |
| Benfica | 1 | 27% | 2 | 405.457 |
| Campo de Ourique | 2 | 19% | | |
| Campolide | 3 | 30% | | |
| Carnide | 2 | 17% | | |
| Estrela | 1 | 42% | | |
| Lumiar | 14 | 33% | 2 | 635.975 |
| Marvila | | | 1 | 248.944 |
| Misericórdia | 4 | 39% | 1 | 548.682 |
| Olivais | 10 | 25% | 2 | 148.106 |
| Parque das Nações | 13 | 45% | 3 | 1.294.280 |
| Santa Maria Maior | 10 | 47% | 1 | 423.660 |
| Santo António | 7 | 31% | 1 | 842.839 |
| São Domingos de Benfica | 1 | 51% | 3 | 325.069 |
| **Total** | **129** | **35%** | **31** | **12.417.865** |

Figure 28 - Matrix Table Parish

Fig. 28 allows a detailed visualization of each parish for a profound exploration of Lisbon's bike-sharing system. The relationship between the number of stations and the average availability in contrast with the number of counters and detections, by each parish, provides a better understanding of the BSS performance. The low rates in Campo de Ourique, Carnide, and Olivais regarding availability suggest that these stations tend to operate with the aim of being stations where users start their journeys. At the same time, parishes like Alvalade, Arroios, and Avenidas Novas show a similar rate of availability and a higher number of detections proving a well-functioning system.

Belém and Parque das Nações have higher rates of availability and a considerable number of detections as these are the parishes furthest away from the centre of Lisbon and therefore have the last stations available to allocate bicycles.

Nevertheless, this table helps understand user behaviour and resource allocation when comparing both metrics. Parishes like Alvalade and Avenidas Novas show a high number of stations when compared with others. At the same time, Alvalade, Avenidas Novas, Lumiar, and Parque das Nações have the highest number of stations and detections which suggests that the highest number of stations in a parish, the higher the acceptance of bicycle-sharing systems as a way of transportation.

Olivais it's the only parish to be considered an outiler and one possible reason for this problem is the low rates regarding availability, therefore, a future look at this situation is necessary.

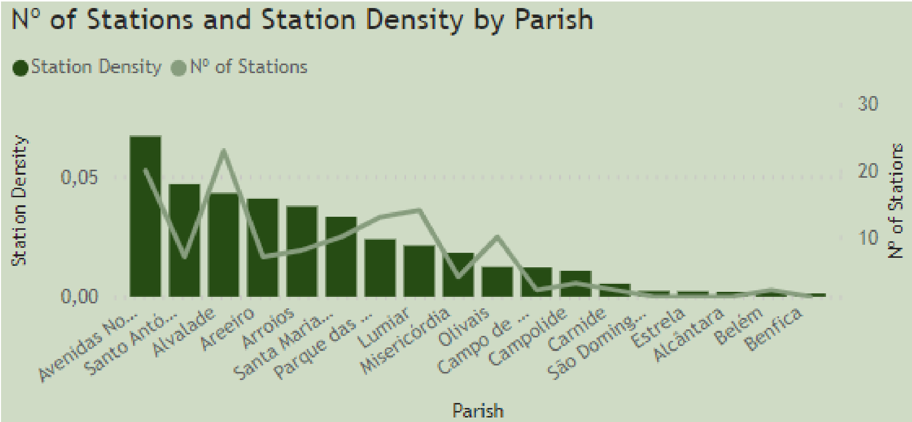


Figure 29 - Nº of Stations and Station Density per Parish

The line and column chart presented above identifies the correlation between the number of stations and the station density each parish has. This visual offers a clear representation of the strategic urban planning regarding the bicycle-sharing system to enhance the service coverage or identify the concentration of infrastructures. Avenidas Novas, Santo António, and Alvalade are identified as the parishes that have a better density regarding their area. Avenidas Novas has the best station density when considering the number of stations, it contains which ultimately represents a good distribution of the placed infrastructures as well as Alvalade. On the other hand, Lumiar, and Parque das Nações have a higher number of stations but a lower density thanks to the territorial area it covers. Is important to highlight that the airport is almost half of the area that Olivais covers, which compromises this metric in this specific case. Lastly, Alcântara, Belém, Benfica, Estrela, and São Domingos de Benfica show a low number of stations in each parish.

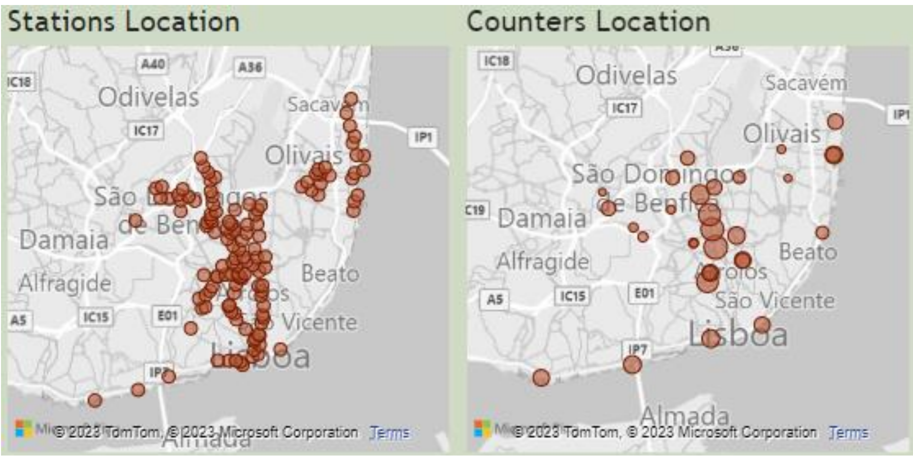


Figure 30 - Counter and Stations Location

The integration of maps relating to the location of each station and counter brings a spatial dimension to our analysis when choosing a specific parish. The map seamlessly adapts to highlight the geographical distribution of both bicycle counters and stations within that chosen parish. This visual representation aids in identifying areas of high or low coverage while providing a holistic perspective on the interaction between counters and bike-sharing stations in the chosen urban context. It is also important to highlight that the bubble size on the counter's location is related to the number of detections.

### 4.4. Usage and Trip Behaviour

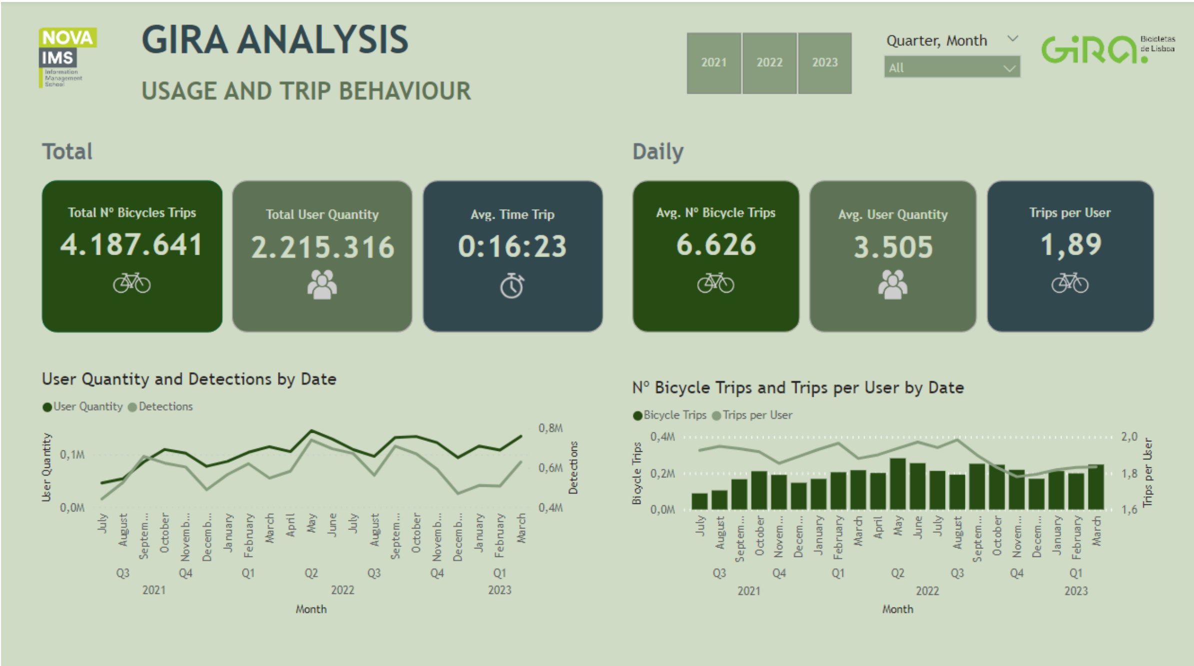


Figure 31 - Dashboard Usage and Trip Behaviour

Lastly, Fig. 31 shows the dashboard that has content related to trips and usage. This last dashboard is less interactive than the others by changing its structure to display visuals more engagingly since it shows mainly cards with values and, therefore, is easier to read and understand. The purpose of this dashboard is to provide a comprehensive view of how users engage with the bike-sharing system while showing bicycle and user activity.

Although changing its format, still contains two slicers for the date in the top right corner as the previous dashboard also had.


Figure 32 - Cards Usage and Trip Behaviour Total

On the one hand, these indicators provide information regarding the total number of bicycle trips, the total number of users, and the average trip duration between July 2021 and March 2023.

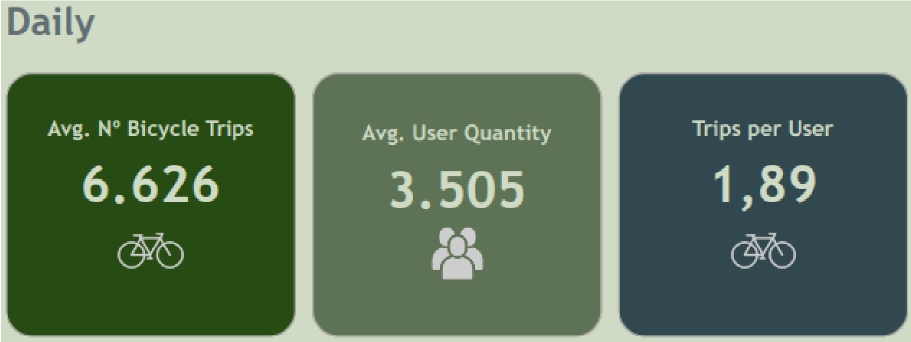


Figure 33 - Cards Usage and Trip Behaviour Daily

On the other hand, it's also possible to make a daily analysis of this indicators that provide information regarding the number of bicycle trips and user quantity but in a daily average value. At the same time, it also visible the number of trips per user value, when considering a daily metric.

Both cards display information about the users' perception towards the system while giving a specific context to the use of bicycles. This cards automatically update according to the date range chosen for the analysis.

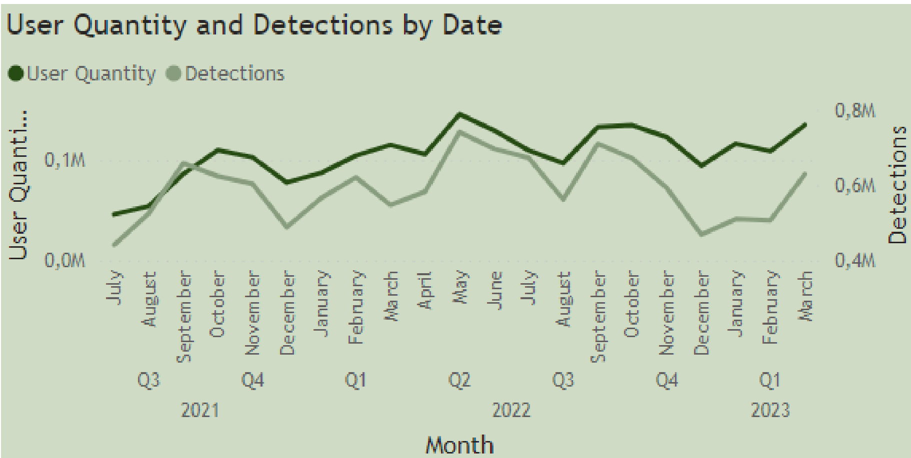


Figure 34 - User Quantity and Detections by Date

This line chart demonstrates the evolution of users and the number of detections between the pre-established date. The upward trend in the graph signifies a consistent growth of users, when analysing the whole spectrum. A correlation emerges between the number of users and the multitude of bicycle detections as both lines tend to follow the same behavioural pattern, suggesting a strong correlation between this metrics. The core of this correlation lies in the understanding that a substantial proportion of these detections are attributed to active users of the bike-sharing system. Moreover, Fig. X reveals distinct patterns in user behaviour over the months.

Notably, the months of September, October and November emerge as periods of greatest affluence of users, contrasting significantly with other times of the year, while the month of May recorded the highest number of users as for detections. During the months of September, its notable an increase in the use of the service as for the use of bicycle around the city. This spike can be attributed to several factors, including the return from the summer holidays, and, therefore the resume of daily activities, as the previous months show a decrease of users. The month of November continues to maintain a positive trend, indicating that the preference for the use of BSS remains robust, however weather changes can be considered as the cause of this decline for the following months. Last but not least, in the last period of 2023, there is a greater dispersion in the correlation between the number of users and detections which indicates a higher variability in user behaviour patterns.

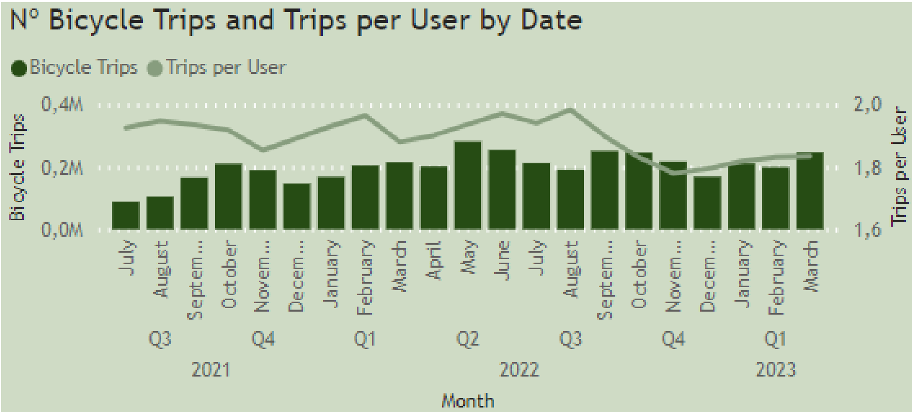


Figure 35 - Nº Bicycle Trips per User by Date

A noteworthy pattern emerges after examining the Fig. 35 depicting the number of bicycle trips and trips per user. Over the observed period, the total number of bicycle trips tends to maintain a consistent trajectory of growth, remaining resilient, which indicates a continuous demand for the bike-sharing system. This shows that the service is being used consistently and effectively. However, a more nuanced insight is revealed when considering trips per user. The line depicting trips per user follows a similar steady course for the majority of the period but experiences a decline in the later observations. This decline suggests a decrease in the average number of trips taken by individual users, signalling a shift in user behaviour. This correlation serves as an indicator for user engagement with the bike-sharing program. Although the number of bicycle trips follows a constant line in the last months, the decrease of trips per users indicates either inefficiencies in user retention, that the system is reaching a saturation point with its current users or even that the system is attracting a larger number of infrequent users.

## 4.5. Discussion

From the analysis of the dashboards developed previously, it is possible to extract some important insights regarding the development of GIRA BSS in Lisbon.

Firstly, the observed continuous growth in the bicycle-sharing system's usage underscores a positive trajectory and increasing acceptance among the city's residents. The rise suggests that more individuals are choosing for the bike-sharing system as a viable and reliable mode of transportation.

This trend may be indicative of shifting preferences towards sustainable and eco-friendly commuting options, reflecting a growing awareness of cycling. Therefore, the public acceptance towards this way of transportation can be seen as an influence for this success (Caggiani et al., 2021). The increasing number of users as detections is evidence of the success of the current system but also a highlight of the potential for further growth and integration of bicycles as a way of urban transportation. Explained in the literature review, the system has an average of 1000 available bicycles to rent daily. This information combined with the average number of bicycles trips per days provides a metric called TDB that represents a way of measuring the success of a BSS. Therefore, the system has an average of 6.6 TDB which is a representation of a good system (Jiménez & Nogal, 2021; Ricci, 2015).

Secondly, when analysing specific months regarding the patterns in bicycle-sharing system utilization May emerges as a standout month, characterized by the highest levels of usage. September showcases significant activity, indicating a surge in usage as individuals return to their routines post-summer. However, the months spanning November to January exhibit a noticeable decline in system utilization due to adverse weather conditions and a possible additional element of reduced daylight hours. As Eren & Uz (2020) and Macioszek et al. (2020) highlight, there's a peak usage on months with higher temperatures contrasting to the levels of usage analysed between November and January.

The analysis of weekly patterns and hourly utilization reveals a distinctive peak in bicycle usage during the hours between 18:00 and 20:00. This suggests a pronounced preference for bicycles during the evening hours. Moreover, a consistent use of bicycles is observed during weekdays from 08:00 to 18:00, signifying their adaptability as a mode of transportation throughout the day. This versatility underscores the suitability of bicycles for various commuting needs, whether for work or other daily activities. Nevertheless, the data highlights that weekdays experience higher bicycle usage compared to weekends, aligning with the conventional commuting patterns observed in others BSS.

Generally, parishes boasting a higher number of stations tend to exhibit more robust bicycle utilization. This implies a positive association between the density of stations and the acceptance of cycling as a preferred mode of commute. The logic follows that the more stations available, the greater the accessibility and convenience for users, fostering a bike-friendly environment as Ricci (2015) previous stated. However, it's noteworthy that Olivais, despite having a substantial number of stations, appears to deviate from this trend. Further investigation into Olivais could provide valuable insights. Additionally, the concentrated usage of bicycles in the city centre and areas along the riverfront underscores the strategic importance of station placement in high-traffic regions.

In conclusion, the careful analysis of station metrics points to a bike-sharing program that, on the whole, provides reliable services. The consistent average availability of bicycles in stations, coupled with a stable variation in the number of bicycles at each station, underlies a system that meets the requirements for satisfactory conditions. This performance highlights an overall availability of bicycles in the system and a general development of cycle mobility in Lisbon, which can accommodate user needs.

# 5. CONCLUSIONS

The development of this study aimed to identify the overall performance that GIRA BSS has in Lisbon regarding users' activity and the bicycle counters spread throughout the city. In doing so, this thesis analyses data collected from open data sources to enable the creation of dashboards that transforms data into easily analysable visuals.

To have a better understanding of the subject in analysis, this study starts by providing information regarding the use of bicycle-sharing systems. An overall perspective towards this sustainable way of transportation is developed highlighting its global popularity as well as the benefits and challenges it faces. Still, to have a deep interpretation of this subject, a section dedicated to the performance of these systems in other countries is also elaborated while giving an introduction to the GIRA bicycle-sharing system since its implementation.

After having a clear perception of the topic to be addressed, a CRISP-DM methodology was applied to the work related with the rest of the project. In this sense, this model conducted a sequence of events that helped in accomplishing the objectives defined in a data mining project by providing an explicit view of each step.

Nevertheless, the creation of four different dashboards, which fulfilled the specific goals defined, display information regarding bicycle detections along the city, station's usage, cycling activity within parishes and the overall behavior of users and bicycles as a whole in a more engaging way that provides insights for a better resource utilization, more informed decision-making regarding bicycle mobility, and its overall development over the years under analysis.

In conclusion, through the analysis made to the GIRA BSS is fair to say that this way of travelling around the city of Lisbon has the potential to be considered a future reliable public transport capable of replacing other means of mobility in Lisbon city as it already fulfills the purpose of several users that use the system constantly. Nevertheless, a continuous future development is needed in order to maximize its full potential and offer better conditions to its users.

## 6. LIMITATIONS AND FUTURE WORK

Although the study on bike sharing systems (BSS) in Lisbon provides insightful information, there are a number of limitations that should be noted in order to direct future research towards a more thorough understanding even though some of them were already previously mentioned when describing the data.

First of all, the reliance on data introduces constraints to analyse a proper report as incomplete datasets and errors pose challenges in achieving an accurate evaluation. The open data source regarding bicycle use is not updated to the present day, which ends up contributing to a lower data quality and insufficient information that is not in real time. At the same time there were several adjustments done regarding data from stations location and data from GIRA stations occupation as a consequence of the lack of data regarding station's dataset which contained a smaller number of stations compared to the actual number that the system currently has. Nevertheless, this dataset doesn't have constant updates for all the period in analysis of the data regarding GIRA which contributes to a less accurate analysis.

The time set for analysis may not capture the full spectrum of seasonal variations since it only comprises data from twenty-one months, making it impossible to provide accurate comparisons when comparing data by year over year. Data from counters is only available from July 2021 and data from GIRA usage is only available until the first quarter of 2023 and, therefore, there is no possible extensive comparison of the bicycle usage that provides a detailed analysis. A more dynamic understanding of BSS usage patterns can be obtained by including real-time data streams into our research, which enables prompt modifications and proactive reactions to evolving circumstances. At the same time, if conducted a longitudinal analysis spanning several years, it would be possible to identify patterns and oscillations that provide a deeper comprehension of the long-term influence of BSS on urban mobility.

Also, a significant constraint of the present investigation concerns the lack of comprehensive user-specific data present in the dataset. The breadth of our analysis is limited by the lack of complete user data, including gender, age gap, and preferences. This intrinsic constraint makes it difficult to understand different aspects of user behaviour. Additional sources of user-specific data could be useful for future research projects in order to improve the level of analysis and offer a more comprehensive view of the user's dynamics regarding bicycle-sharing systems.

Nevertheless, other limitation stems from the assumption that all data obtained from the counters exclusively corresponds only to bicycle usage. While the prevailing perception supports the notion that the majority of recorded occurrences emanate from bicycle users, the absence of precise information introduces a certain level of uncertainty. This limitation poses challenges in ensuring the accuracy of the final analysis. Future work could benefit from refining the precision of bicycles counts to offer a more nuanced understanding of counters data.

The acknowledgment of the potential impact of citywide events prompts a call for more extensive research. City events, ranging from festivals to major gatherings, can significantly alter commuting patterns and user behaviours, affecting the overall utilization of the bike-sharing system. Moreover, the exploration of holidays, notable streaks from public transportation services, and the nuanced interplay with diverse weather conditions adds depth to this consideration. Holidays often reshape travel demands, while streaks from public transportation disruptions might prompt a surge in BSS usage while weather conditions, such as rain, also play a big role when influencing the use of bicycles.

In future research, a crucial exploration lies in understanding the varied impacts of conventional and electric bicycles within bike sharing systems (BSS). By examining how usage patterns differ between conventional and electric bicycles, as well as any perceived advantages or barriers to the adoption of electric bicycles it is easier to forecast bicycle demand.

In conclusion, the acknowledgment of these inherent limitations serves not as a roadblock but as a catalyst for the ongoing evolution of this research. By conscientiously recognizing these constraints and articulating a thoughtful plan for future investigations, this study becomes a dynamic investigation which has several future aspects that can potentially be analysed to keep up with the rapidly changing urban transportation environment, specifically bicycle-sharing systems.

## BIBLIOGRAPHICAL REFERENCES


André, M. R. (2021, abril 23). Ciclovias em Lisboa: Teimosia ou estratégia? *Lisboa Para Pessoas*. https://lisboaparapessoas.pt/2021/04/23/ciclovias-lisboa-estrategia/

André, M. R. (2022a, março 9). O recorde diário da GIRA e mais alguns dados sobre as bicicletas partilhadas de Lisboa. *Lisboa Para Pessoas*. https://lisboaparapessoas.pt/2022/03/09/gira-recorde-diario-dados/

André, M. R. (2022b, dezembro 4). EMEL vai electrificar 400 GIRAs "clássicas" e fazer duas ciclovias em 2023. *Lisboa Para Pessoas*. https://lisboaparapessoas.pt/2022/12/04/emel-gira-ciclovias-2023-lisboa/

Avila-Palencia, I., de Nazelle, A., Cole-Hunter, T., Donaire-Gonzalez, D., Jerrett, M., Rodriguez, D. A., & Nieuwenhuijsen, M. J. (2017). The relationship between bicycle commuting and perceived stress: A cross-sectional study. *BMJ Open*, *7*(6), e013542. https://doi.org/10.1136/bmjopen-2016-013542

Buck, D., & Buehler, R. (2012). *Bike Lanes and Other Determinants of Capital Bikeshare Trips*. 11.

Bullock, C., Brereton, F., & Bailey, S. (2017). The economic contribution of public bike-share to the sustainability and efficient functioning of cities. *Sustainable Cities and Society*, *28*, 76–87. https://doi.org/10.1016/j.scs.2016.08.024

Caggiani, L., Camporeale, R., Hamidi, Z., & Zhao, C. (2021). Evaluating the Efficiency of Bike-Sharing Stations with Data Envelopment Analysis. *Sustainability*, *13*(2), 881. https://doi.org/10.3390/su13020881

Câmara Municipal de Lisboa. (2022, novembro 9). *Rede Gira será gratuita com o Passe Navegante*. MUNICÍPIO de LISBOA. https://www.lisboa.pt/atualidade/noticias/detalhe/rede-gira-sera-gratuita-com-o-passe-navegante

Cavill, N., & Watkins, F. (2007). Cycling and health: An exploratory study of views about cycling in an area of North Liverpool, UK. *Health Education*, *107*(5), 404–420. https://doi.org/10.1108/09654280710778556

Celis-Morales, C. A., Lyall, D. M., Welsh, P., Anderson, J., Steell, L., Guo, Y., Maldonado, R., Mackay, D. F., Pell, J. P., Sattar, N., & Gill, J. M. R. (2017). Association between active commuting and incident cardiovascular disease, cancer, and mortality: Prospective cohort study. *BMJ*, *357*, j1456. https://doi.org/10.1136/bmj.j1456

Chapman, D., & Larsson, A. (2019). Toward an Integrated Model for Soft-Mobility. *International Journal of Environmental Research and Public Health*, *16*(19), Artigo 19. https://doi.org/10.3390/ijerph16193669

Chapman, P. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. https://www.semanticscholar.org/paper/CRISP-DM-1.0%3A-Step-by-step-data-mining-guide-Chapman/54bad20bbc7938991bf34f86dde0babfbd2d5a72

Chemla, D., Meunier, F., & Wolfler Calvo, R. (2013). Bike sharing systems: Solving the static rebalancing problem. *Discrete Optimization*, *10*(2), 120–146. https://doi.org/10.1016/j.disopt.2012.11.005

Clockston, R. L. M., & Rojas-Rueda, D. (2021). Health impacts of bike-sharing systems in the U.S. *Environmental Research*, *202*, 111709. https://doi.org/10.1016/j.envres.2021.111709

Corticelli, R., Pazzini, M., Mazzoli, C., Lantieri, C., Ferrante, A., & Vignali, V. (2022). Urban Regeneration and Soft Mobility: The Case Study of the Rimini Canal Port in Italy. *Sustainability*, *14*(21), Artigo 21. https://doi.org/10.3390/su142114529

D'Almeida, L., Rye, T., & Pomponi, F. (2021). Emissions assessment of bike sharing schemes: The case of Just Eat Cycles in Edinburgh, UK. *Sustainable Cities and Society*, *71*, 103012. https://doi.org/10.1016/j.scs.2021.103012

DeMaio, P. (2009). Bike-sharing: History, Impacts, Models of Provision, and Future. *Journal of Public Transportation*, *12*(4), 41–56. https://doi.org/10.5038/2375-0901.12.4.3

do Nascimento Liberto, L. F. (2018). *O Potencial dos Sistemas de Bicicletas Partilhadas: Uma contribuição para a construção do panorama português*. https://repositorio-aberto.up.pt/bitstream/10216/111382/2/259610.pdf

EMEL. (2019). *Estatutos EMEL*. https://www.emel.pt/fotos/gca/bm_11_04_2019_3.%C2%BA_sup_1312_estatutos_da_emel_1_8_19_19683457435d7b4fcb5645f.pdf

EMEL. (2021). *GIRA - Termos e Condições*. https://www.gira-bicicletasdelisboa.pt/wp-content/uploads/2022/05/GIRA_Termos_e_Condicoes_20220601.pdf

Eren, E., & Uz, V. E. (2020). A review on bike-sharing: The factors affecting bike-sharing demand. *Sustainable Cities and Society*, *54*, 101882. https://doi.org/10.1016/j.scs.2019.101882

Félix, R., Cambra, P., & Moura, F. (2020). Build it and give 'em bikes, and they will come: The effects of cycling infrastructure and bike-sharing system in Lisbon. *Case Studies on Transport Policy*, *8*(2), 672–682. https://doi.org/10.1016/j.cstp.2020.03.002

Ferrari, A., & Russo, M. (2016). *Introducing Microsoft Power BI*. Microsoft Press.

Few, S. (2006). *Information dashboard design: The effective visual communication of data*. https://www.academia.edu/1380138/Information_dashboard_design_The_effective_visual_communication_of_data

Fishman, E. (2016). Bikeshare: A Review of Recent Literature. *Transport Reviews*, *36*(1), 92–113. https://doi.org/10.1080/01441647.2015.1033036

Fishman, E., Washington, S., & Haworth, N. (2013). Bike Share: A Synthesis of the Literature. *Transport Reviews*, *33*(2), 148–165. https://doi.org/10.1080/01441647.2013.775612

Fishman, E., Washington, S., Haworth, N., & Watson, A. (2015). Factors influencing bike share membership: An analysis of Melbourne and Brisbane. *Transportation Research Part A: Policy and Practice*, *71*, 17–30. https://doi.org/10.1016/j.tra.2014.10.021

Gallo, M., & Marinelli, M. (2020). Sustainable Mobility: A Review of Possible Actions and Policies. *Sustainability*, *12*(18), Artigo 18. https://doi.org/10.3390/su12187499

Holden, E., Gilpin, G., & Banister, D. (2019). Sustainable Mobility at Thirty. *Sustainability*, *11*(7), Artigo 7. https://doi.org/10.3390/su11071965

Hosford, K., Lear, S. A., Fuller, D., Teschke, K., Therrien, S., & Winters, M. (2018). Who is in the near market for bicycle sharing? Identifying current, potential, and unlikely users of a public bicycle share program in Vancouver, Canada. *BMC Public Health*, *18*(1), 1326. https://doi.org/10.1186/s12889-018-6246-3

Høyer, K. G. (1999). *Sustainable Mobility—The Concept and its Implications*.

Hunt, J. D., & Abraham, J. E. (2007). Influences on bicycle use. *Transportation*, *34*(4), 453–470. https://doi.org/10.1007/s11116-006-9109-1

Jacobsen, P. L. (2015). Safety in numbers: More walkers and bicyclists, safer walking and bicycling. *Injury Prevention*, *21*(4), 271–275. https://doi.org/10.1136/ip.9.3.205rep

Jensen, P., Rouquier, J.-B., Ovtracht, N., & Robardet, C. (2010). Characterizing the speed and paths of shared bicycle use in Lyon. *Transportation Research Part D: Transport and Environment*, *15*(8), 522–524. https://doi.org/10.1016/j.trd.2010.07.002

Jiménez, P., & Nogal, M. (2021). Analysis of real experiences using different sized bike sharing schemes in Irish cities. *Transportation Research Procedia*, *58*, 37–44. https://doi.org/10.1016/j.trpro.2021.11.006

Kou, Z., Wang, X., Chiu, S. F. (Anthony), & Cai, H. (2020). Quantifying greenhouse gas emissions reduction from bike share systems: A model considering real-world trips and transportation mode choice patterns. *Resources, Conservation and Recycling*, *153*, 104534. https://doi.org/10.1016/j.resconrec.2019.104534

Leyland, L.-A., Spencer, B., Beale, N., Jones, T., & Reekum, C. M. van. (2019). The effect of cycling on cognitive function and well-being in older adults. *PLOS ONE*, *14*(2), e0211779. https://doi.org/10.1371/journal.pone.0211779

Li, Q., Fuerst, F., & Luca, D. (2022). *Do Shared E-Bikes Reduce Urban Carbon Emissions?* (SSRN Scholarly Paper 4233639). https://doi.org/10.2139/ssrn.4233639

Machado, P., Ferreira, J., Fernandes, T., & Simas, A. (2020). *MOVE Lisboa: Visão estratégica para a mobilidade 2030*. Câmara Municipal de Lisboa.

Macioszek, E., Świerk, P., & Wyderka, A. (2020). The Bike-Sharing System as an Element of Enhancing Sustainable Mobility—A Case Study based on a City in Poland. *Sustainability*, *12*, 3285. https://doi.org/10.3390/su12083285

Médard de Chardon, C. (2019). The contradictions of bike-share benefits, purposes and outcomes. *Transportation Research Part A: Policy and Practice*, *121*, 401–419. https://doi.org/10.1016/j.tra.2019.01.031

Médard de Chardon, C., Caruso, G., & Thomas, I. (2017). Bicycle sharing system 'success' determinants. *Transportation Research Part A: Policy and Practice*, *100*, 202–214. https://doi.org/10.1016/j.tra.2017.04.020

Midgley, P. (2011, fevereiro 28). *Bicycle-sharing Schemes:Enhancing Sustainable Mobility in Urban Areas*. https://scholar.googleusercontent.com/scholar?q=cache:SdZSTf72zwAJ:scholar.google.com/+BICYCLE-SHARING+SCHEMES:+ENHANCING+SUSTAINABLE+MOBILITY+IN+URBAN+AREAS&hl=en&as_sdt=0,5

Moreira, J. (2022, novembro 10). *Bicicletas Gira vão ser gratuitas para quem tem passe Navegante em 2023*. Time Out Lisboa. https://www.timeout.pt/lisboa/pt/noticias/bicicletas-gira-vao-ser-gratuitas-para-quem-tem-passe-navegante-em-2023-111022

Nations, U. (2017). *Sustainable Mobility for All. 2017. Global Mobility Report 2017: Tracking Sector Performance. Washington DC, License: Creative Commons Attribution CC BY 3.0*. https://sdgs.un.org/sites/default/files/publications/2643Global_Mobility_Report_2017.pdf

Oja, P., Titze, S., Bauman, A., de Geus, B., Krenn, P., Reger-Nash, B., & Kohlberger, T. (2011). Health benefits of cycling: A systematic review. *Scandinavian Journal of Medicine & Science in Sports*, *21*(4), 496–509. https://doi.org/10.1111/j.1600-0838.2011.01299.x

Olmos, L. E., Tadeo, M. S., Vlachogiannis, D., Alhasoun, F., Espinet Alegre, X., Ochoa, C., Targa, F., & González, M. C. (2020). A data science framework for planning the growth of bicycle infrastructures. *Transportation Research Part C: Emerging Technologies*, *115*, 102640. https://doi.org/10.1016/j.trc.2020.102640

O'Sullivan, F. (2022, fevereiro 26). The Radical Roots of Bikesharing. *Bloomberg.Com*. https://www.bloomberg.com/news/features/2022-02-26/the-dutch-anarchists-who-launched-a-bikesharing-revolution

Otero, I., Nieuwenhuijsen, M. J., & Rojas-Rueda, D. (2018). Health impacts of bike sharing systems in Europe. *Environment International*, *115*, 387–394. https://doi.org/10.1016/j.envint.2018.04.014

Parkes, S. D., Marsden, G., Shaheen, S. A., & Cohen, A. P. (2013). Understanding the diffusion of public bikesharing systems: Evidence from Europe and North America. *Journal of Transport Geography*, *31*, 94–103. https://doi.org/10.1016/j.jtrangeo.2013.06.003

Patel, S. J., & Patel, C. R. (2020). A stakeholders perspective on improving barriers in implementation of public bicycle sharing system (PBSS). *Transportation Research Part A: Policy and Practice*, *138*, 353–366. https://doi.org/10.1016/j.tra.2020.06.007

Penela, R. (2022, novembro 8). *Lisboa. Bicicletas Gira gratuitas no início de 2023 só para quem já tem passes grátis*. Observador. https://observador.pt/2022/11/08/lisboa-bicicletas-gira-gratuitas-no-inicio-de-2023-so-para-quem-ja-tem-passes-gratis/

Programme, U. N. E. (2010). *Investment in walking and cycling road infrastructure share the road*. UNEP. https://wedocs.unep.org/xmlui/handle/20.500.11822/7890

Ricci, M. (2015). Bike sharing: A review of evidence on impacts and processes of implementation and operation. *Research in Transportation Business & Management*, *15*, 28–38. https://doi.org/10.1016/j.rtbm.2015.03.003

SAPO. (2022, set). *Andar a pé, é a estratégia nacional para reduzir o sedentarismo em 15% até 2030*. SAPO 24. https://24.sapo.pt/atualidade/artigos/andar-a-pe-e-a-estrategia-nacional-para-reduzir-o-sedentarismo-em-15-ate-2030

Shaheen, S. A., Guzman, S., & Zhang, H. (2010). Bikesharing in Europe, the Americas, and Asia: Past, Present, and Future. *Transportation Research Record*, *2143*(1), 159–167. https://doi.org/10.3141/2143-20

Shaheen, S. A., Zhang, H., Martin, E., & Guzman, S. (2011). China's Hangzhou Public Bicycle: Understanding Early Adoption and Behavioral Response to Bikesharing. *Transportation Research Record*, *2247*(1), 33–41. https://doi.org/10.3141/2247-05

Shaheen, S., Guzman, S., & Zhang, H. (2012). *Bikesharing across the Globe*. https://escholarship.org/uc/item/0qm296pf

Shi, J., Si, H., Wu, G., Su, Y., & Lan, J. (2018). Critical Factors to Achieve Dockless Bike-Sharing Sustainability in China: A Stakeholder-Oriented Network Perspective. *Sustainability*, *10*(6), Artigo 6. https://doi.org/10.3390/su10062090

SIC Notícias. (2022, setembro 22). *Bicicletas GIRA: Uma alternativa sustentável cada vez mais procurada em Lisboa*. SIC Notícias. https://sicnoticias.pt/pais/2022-09-22-Bicicletas-GIRA-uma-alternativa-sustentavel-cada-vez-mais-procurada-em-Lisboa-526d1139

Urban Solutions, P. (2021). *The Meddin Bike-sharing World Map*. PBSC Urban Solutions. https://bikesharingworldmap.com/reports/bswm_mid2021report.pdf

Vallez, C.M. (2021, março 1). *Bike-Sharing Systems*. https://encyclopedia.pub/entry/7646

Vasconcelos, B. (2021). *Sistemas de Bicicletas Partilhadas—O CASO DA GIRA*. https://lisboaenova.org/images/stories/Conferencias/2021/Sharing_Cities_Conf_Final/Bruno_Vasconcelos_Maia_EMEL.pdf

Willberg, E., Salonen, M., & Toivonen, T. (2021). What do trip data reveal about bike-sharing system users? *Journal of Transport Geography*, *91*, 102971. https://doi.org/10.1016/j.jtrangeo.2021.102971

Wirth, R., & Hipp, J. (2010). *CRISP-DM: Towards a Standard Process Model for Data Mining*.

Woodcock, J., Tainio, M., Cheshire, J., O'Brien, O., & Goodman, A. (2014). Health effects of the London bicycle sharing system: Health impact modelling study. *BMJ*, *348*, g425. https://doi.org/10.1136/bmj.g425

Zee, R. van der. (2016, abril 26). Story of cities #30: How this Amsterdam inventor gave bike-sharing to the world. *The Guardian*. https://www.theguardian.com/cities/2016/apr/26/story-cities-amsterdam-bike-share-scheme

Zhang, L., Zhang, J., Duan, Z., & Bryde, D. (2015). Sustainable bike-sharing systems: Characteristics and commonalities across cases in urban China. *Journal of Cleaner Production*, *97*, 124–133. https://doi.org/10.1016/j.jclepro.2014.04.006

# 7. APPENDIX A

| ID Counter | Postal Code | Latitude | Longitude | Location | Identifier | Freguesia |
|---|---|---|---|---|---|---|
| 2 | 1500-069 | 38.749622968 | -9.191152508 | Avenida do Colégio Militar | #18 | Benfica |
| 8 | 1500-006 | 38.755009563 | -9.193723551 | Avenida Lusíada (Entrada - Benfica) | #11 | Benfica |
| 3 | 1600-616 | 38.759682478 | -9.164044385 | Telheiras - Colégio Alemão | #01 | Lumiar |
| 25 | 1750-148 | 38.766217494 | -9.157662731 | Alameda das Linhas de Torres | #16 | Lumiar |
| 4 | 1700-111 | 38.75984155 | -9.135774987 | Avenida do Brasil - Mata de Alvalade | #02 | Alvalade |

Figure 36 - DIM_CountersLocation rows

| Latitude | Longitude | Location | Postal Code | Opening Date | ID Station | Freguesia |
|---|---|---|---|---|---|---|
| 38,778466 | -9,097068 | Rua de Moscavide | 1900-238 | 21/06/2017 | 110 | Benfica |
| 38,747751 | -9,136681 | Rua Teixeira de Pascoais / Rua Dr. Gama Barros | 1000-009 | 18/11/2017 | 452 | Lumiar |
| 38,75176 | -9,13749 | Av. Rio de Janeiro / Parque de Jogos 1º de Maio | 1700-330 | 13/12/2017 | 460 | Alvalade |
| 38,735188 | -9,148443 | Av. Duque de Ávila / Av. Conde Valbom | 1000-141 | 06/12/2017 | 415 | Alvalade |
| 38,735352 | -9,142703 | Av. Duque de Ávila / Jardim Arco Do Cego | 1000-141 | 06/12/2017 | 417 | Avenidas Novas |

Figure 37 - DIM_StationsLocation rows

| Dicofre | Parish | Concelho | Distrito | TAA | AREA_EA_Ha |
|---|---|---|---|---|---|
| 110660 | Estrela | LISBOA | LISBOA | ÁREA PRINCIPAL | 460,4 |
| 110661 | Misericórdia | LISBOA | LISBOA | ÁREA PRINCIPAL | 219,22 |
| 110665 | Santa Maria Maior | LISBOA | LISBOA | ÁREA PRINCIPAL | 301,16 |
| 110658 | Belém | LISBOA | LISBOA | ÁREA PRINCIPAL | 1042,76 |
| 110667 | São Vicente | LISBOA | LISBOA | ÁREA PRINCIPAL | 198,58 |

Figure 38 - DIM_Parish rows

| Date | Year | Quarter | Month | Day | Day of Week | ID Date | Week Day | MonthNumber |
|---|---|---|---|---|---|---|---|---|
| 01/07/2022 | 2022 | Q3 | Julho | 1 | Sexta-Feira | 20220701 | 6 | 7 |
| 02/07/2022 | 2022 | Q3 | Julho | 2 | Sábado | 20220702 | 7 | 7 |
| 03/07/2022 | 2022 | Q3 | Julho | 3 | Domingo | 20220703 | 1 | 7 |
| 04/07/2022 | 2022 | Q3 | Julho | 4 | Segunda-Feira | 20220704 | 2 | 7 |
| 05/07/2022 | 2022 | Q3 | Julho | 5 | Terça-Feira | 20220705 | 3 | 7 |

Figure 39 - DIM_Date rows

| ID Station | Nº Bicycles | Nº Docks | Time | Date | ID Date |
|---|---|---|---|---|---|
| 428 | 2 | 12 | 00:00:00 | 23/02/2022 | 20220223 |
| 427 | 3 | 12 | 00:00:00 | 15/06/2022 | 20220615 |
| 464 | 7 | 14 | 00:00:00 | 21/04/2022 | 20220421 |
| 107 | 8 | 14 | 00:00:00 | 21/04/2022 | 20220421 |
| 472 | 10 | 14 | 00:00:00 | 15/06/2022 | 20220615 |

Figure 40 - Fact_CountersData rows

| Counts | ID Counter | Direction | Date | Hours | ID Date |
|---|---|---|---|---|---|
| 7 | 19 | 1 | 08/07/2021 | 21:00:00 | 20210708 |
| 26 | 21 | 0 | 08/07/2021 | 22:00:00 | 20210708 |
| 39 | 21 | 0 | 08/07/2021 | 20:00:00 | 20210708 |
| 56 | 21 | 0 | 08/07/2021 | 19:00:00 | 20210708 |
| 42 | 21 | 0 | 08/07/2021 | 17:00:00 | 20210708 |

Figure 41 - Fact_GiraData rows

| Date | Bike Use per Day in Seconds | User Quantity | Avg Trip by User in Seconds | Avg Trip Seconds | Nº Bicycle Trips | ID Date | Duration Avg Trip by User | Duration Avg Trip |
|---|---|---|---|---|---|---|---|---|
| 08/07/2021 | 4661401 | 2489 | 1872,800723 | 1000,086033 | 4661 | 20210708 | 00:31:13 | 00:16:41 |
| 09/07/2021 | 4290296 | 2288 | 1875,12937 | 984,690383 | 4357 | 20210709 | 00:31:16 | 00:16:25 |
| 10/07/2021 | 3530996 | 1460 | 2418,49041 | 1302,469937 | 2711 | 20210710 | 00:40:19 | 00:21:43 |
| 11/07/2021 | 3468043 | 1384 | 2505,811416 | 1314,150435 | 2639 | 20210711 | 00:41:46 | 00:21:55 |
| 12/07/2021 | 4186384 | 2168 | 1930,988929 | 999,614135 | 4188 | 20210712 | 00:32:11 | 00:16:40 |

Figure 42 - Fact_Trips&Usage rows

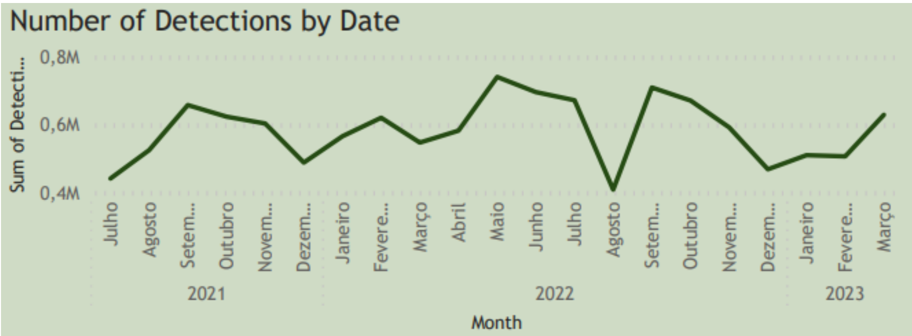


Figure 43 - Number Detections without Forecast