NOVA IMS
Information Management School

MDSAA

Master's degree Program in

Data Science and Advanced Analytics

The Impact of Superblocks on Urban Health

Measuring Traffic and Environmental Changes in Lisbon during the Campo de Ourique Superblock Experience

Martim Hernandez Salgado dos Santos

Project Work

Presented as partial requirement for obtaining the Master's degree Program in *Data Science and Advanced Analytics*

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

The Impact of Superblocks on Urban Health

by

Martim Hernandez Salgado dos Santos

Project Work presented as partial requirement for obtaining the Master's degree Program in Data Science and Advanced Analytics, with a specialization in Data Science

Supervisor: Miguel de Castro Neto, PhD, NOVA Information Management School

Co-supervisor: Bruno Jardim, PhD, NOVA Information Management School

July 2024

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 acknowledged the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Lisbon, 13th July 2024

ABSTRACT

This research examines the emerging urban superblock design model in Portugal, focusing on its' first experimental implementation in Campo de Ourique. The superblock concept was first introduced in Barcelona and consists in closing the inner streets of a group of 9 different blocks to vehicular traffic, increasing space for pedestrians, enhancing urban greening and reducing pollution. Although the impact of superblocks has been extensively studied, there is a gap in exploring their feasibility and effectiveness in Portuguese metropolitan areas. This research aims to assess that gap, by studying the impact of the Campo de Ourique's superblock on air quality, noise levels and traffic congestion. By combining the CRISP-DM and the Difference-in-Differences (DiD) methodologies, this study comprehensively analysed traffic data from Waze, and air and noise measurements from quality monitoring stations that are distributed across the city of Lisbon. The DiD methodology, in combination with CRISP-DM, provided a robust framework for estimating the causal impact of the superblock, and comparing the results of the treatment group (Campo de Ourique) and control group (Jardim do Arco do Cego). The results revealed statistically significant variations in nitrogen dioxide (NO2) concentration levels and average traffic queue' lengths. Other important findings included reductions in particulate matter (PM2.5 and PM10) concentrations, further enhancing air quality, and a decrease in average traffic queues' speed, promoting street safety. Considering the specific needs and dynamics of Portuguese cities, this study aims to inform policymakers and stakeholders, facilitating the process of adopting superblocks as a realistic strategy for urban transformation.

KEYWORDS

Superblocks; Air Quality; Traffic Congestion; Noise; Difference-in-Differences

Sustainable Development Goals (SDG):

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AIC	Akaike Information Criterion
BIC	Bayesaian Information Criterion
CI	Confidence Interval
CRISP-DM	Cross-Industry Standard Process for Data Mining
CO	Carbon Monoxide
dB	Decibel
DiD	Difference-in-Differences
GHG	Greenhouse Gas
km/h	Kilometer per hour
Log	Logarithm
NO2	Nitrogen Dioxide
PM	Particulate Matters
PM2.5	Particulate Matters with a diameter less than 2.5 micrometers
PM10	Particulate Matters with a diameter less than 10 micrometers
RCT	Randomized Controlled Trials
SDG	Sustainable Development Goals
VIF	Variance Influence Factors
µm	Micrometer

1. INTRODUCTION

To address the challenges associated with increased urbanization, many cities around the world are undergoing transformative changes and adopting different sustainable strategies. In Barcelona, the concept of superblocks, initially proposed as an innovative urban design strategy (Eggimann, 2022), is serving as a model to reshape the city, street by street. A superblock is composed of nine different city blocks, where their inner roads are closed to vehicular traffic, creating more space for residents and pedestrians, while reducing noise and air pollution (Lopez et al., 2020).

In September 2023, Lisbon started the first experimental test of a superblock, in Campo de Ourique, from the 9th to the 17th (Rodrigues, 2023). During this period, the streets between "Jardim da Parada" and surrounding buildings were temporarily closed to vehicular traffic, creating a superblock in this area. This experience served as a preliminary test, to evaluate the efficacy of this urban design strategy, while increasing air quality, providing more safety and creating more space for visitors and residents.

Despite the existence of multiple studies evaluating the effect of superblocks in Barcelona and other research on health benefits (Li & Wilson, 2023; Mueller et al., 2020), climate change impact (Lopez et al., 2020), road safety improvements and urban green space expansion (Eggimann, 2022), there is significant gap regarding their implications and implementation in Portuguese metropolitan areas.

This study builds upon the implementation of Campo de Ourique superblock to assess its influence on various indicators. It aims to conduct a descriptive analysis and an impact assessment, evaluating how the changes that were imposed by the superblock affected the surrounding areas, utilizing Waze data for traffic analysis and data from sensors that are distributed through the municipality of Lisbon for air quality and noise analysis.

Through this research, we aim to provide insights into the potential benefits of superblocks for residents and visitors. We expect that our findings will guide future urban planning decisions, promote sustainable mobility, improve health outcomes, and create a more environmentally friendly environment.

The upcoming chapter of this master thesis comprises a literature review exploring the concept of superblocks, focusing on how they impact air quality, noise levels, and traffic congestion. The "Data and Methodology" section provides an explanation of the research design and the employed methods, including the CRISP-DM framework, the Difference-in-Differences methodology and an exploratory analysis and description of the used data sources. In the "Results and Discussion" chapter, we present the data analysis outcomes, the findings obtained using the DiD methodology and discuss the implications of these findings for future urban planning and policy. Finally, in the "Conclusions, Limitations and Future

Works” chapter, we summarize our findings, discuss the limitations of the study, and suggest potential directions for future research.

2. LITERATURE REVIEW

In today's world, climate change and global warming stand as one of the most important and threatening challenges faced by societies (Cornell & Gupta, 2020). One of the main causes of this climatic shift is attributed to greenhouse gas (GHG) emissions, which are responsible for inducing global warming effect (Lopez et al., 2020). While there are multiple factors and components that contribute to this environmental problem, transportation stands as one of its' major contributors, with European Environment Agency claiming that this sector is responsible for more than 25% of the European Union's GHG emissions (European Environment Agency, 2023).

Superblocks, and other sustainable urban neighbourhood transformation strategies, are regarded as a way to mitigate these emissions and address the impacts of climate change, by reducing vehicular traffic and introducing green zones in urbanized areas. These implementations do not only improve air quality, but also re-shape neighbourhoods and the spaces where people live - by restricting vehicular access and converting roads into green and secure spaces, superblocks create a safe place for people of all ages, improving air quality, reducing noise pollution (Nieuwenhuijsen et al.,2024) and protecting them from the dangers associated with vehicular traffic.

By analysing data from Waze and air quality and noise sensors, our research aims to assess the impact of the Campo de Ourique superblock experimental implementation, evaluating its' impact on air quality and traffic, intending to determine if the implementation of the superblock has resulted in measurable impacts and improvements.

2.1. SUPERBLOCKS

Barcelona's superblock model proposes a 400 x 400-meter cluster of nine street blocks, limiting traffic on main roads to 50 km/h and closing inner roads to vehicular transportation, aiming to promote diverse use of public space (Müller et al., 2023).

The introduction of superblocks serves as a strategic urban transformation strategy to create pedestrian-centric neighbourhoods (Eggimann et al., 2021) by restructuring the typical urban road network, reducing automobile traffic substantially, and accordingly GHG emissions, while increasing green space in the city and improving the health and quality of life of its residents and visitors (López et al., 2020). López et. al (2020) also claim that superblocks do not require investment in hard infrastructures or involve demolishing buildings or undertaking massive development, being, in fact, a form of "low-tech urbanism".

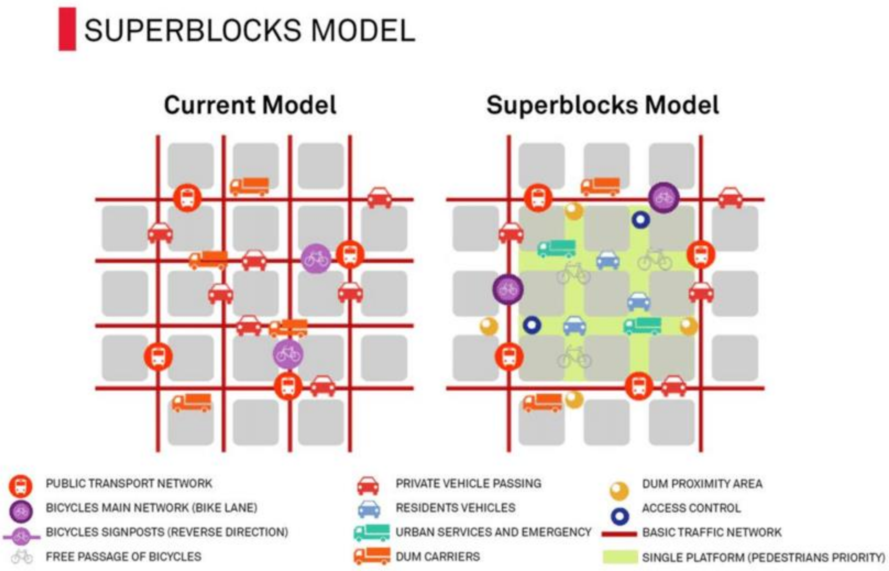

Figure 1 - Superblock Configuration. Taken from (Navarro-Serer, 2021).

As observed in Figure 1, in Barcelona's Superblock model, while inner streets are mainly reserved for cyclist and pedestrians, can, under specific conditions, be used by residential traffic, services, emergency vehicles, and loading/unloading vehicles. Vehicular traffic occurs primarily on the outer boundaries of the superblock, forming the main road system.

By incorporating superblocks into urban landscapes, neighbourhoods can benefit from more liveable, healthy and walkable neighbourhoods (Li and Wilson, 2023). Addressing these transformations holds significant relevance for Lisbon, especially considering its designation as the 2020 European Green Capital. According to de Almeida et al. (2022), despite having received this honour, Lisbon has registered a reduction in green spaces during the past years.

Barcelona is the largest city in the world to have adopted the superblock concept as a guiding principle for the whole city (Amati et al., 2023). Oltra et al.'s (2022) research in Barcelona highlights the economic opportunities associated with this urban model, revealing the establishment of over 20 new local businesses and a 20% increase in economic activity within superblock areas.

The transformation of the Eixample neighbourhood of Barcelona demonstrates the potential of the superblock model to create more sustainable landscapes (Mueller et al., 2020), helping to provide a benchmark for potential outcomes in Lisbon.

. The shift from a car-dependent city planning to a model that prioritizes public space and sustainable mobility aligns with Sustainable Development Goals (SDGs), addressing not only the pressing issues related to climate change but also contributing to other global challenges as poverty reduction, inequality, and environmental degradation.

Despite their numerous advantages, implementing superblocks also faces some challenges and concerns. Filastó (2022) identifies different barriers such as opposition from residents, drivers, and local merchants, as well as issues related to cost recovery, planning for shifted traffic circulation, application, and lack of political support. The author also emphasizes public opposition as a significant obstacle to these sustainable urban neighbourhood transformation strategies. The reduction of available parking spaces, as emphasized by Deisting (2022), also stands out as a significant concern.

Filastó (2022) highlights the importance of policy attributes in public acceptance, noting that policies enhancing public transport are generally more accepted than regulatory or economic measures. The author also emphasizes the relevance of stablishing non-coercive measures over coercive ones, pointing to the perception that the former is more effective, fair, and acceptable. Despite this understanding, the author acknowledges the complexity and context specificity of attitudes towards infrastructural interventions like superblocks, underscoring the need for more systematic evidence.

As Portuguese metropolitan areas consider establishing superblocks or other urban neighbourhood transformation strategies, they will face challenges similar to those described by Filastó (2022). Anticipating and addressing local concerns, particularly related to traffic circulation, parking spaces, cost recovery and public acceptance, is mandatory. By recognizing the importance of a balanced approach, especially in addressing issues like reduced parking spaces, Lisbon can learn from these challenges to inform a more contextually relevant implementation of superblocks, aligning with the city's goal of promoting sustainable urban development and enhancing residents' quality of life.

2.2. Environmental Evaluation

Our respiratory and cardiovascular systems' well-being depends on the quality of the air we breathe. Air pollution has a tremendous impact on people's health, increasing physical discomfort, chronic illnesses, and compromised self-rated health, particularly on those belonging to marginalized segments of society (Liao et al., 2023).

Santibáñez-Andrade et al. (2022) emphasize the connection between the exposure to pollutants like Particulate Matters (PM) with a diameter of less than 10 µm and negative health outcomes, as lung cancer and other lung diseases, due to the disruption of normal cell functions.

Air quality is an extremely important factor that affects our health, lives and the world we live in. Keeping it under control is essential, and an excellent method to do so is through the implementation (and analysis) of distributed sensors that measure air quality across different location. In the specific case of Lisbon, Sarroeira et al. (2023) conducted a characterization of the city, utilizing data from air-quality monitoring stations, green spaces, road infrastructures, and housing density. Their objective was to establish relationships between these characteristics and the distribution of air quality parameters. Over the course of a year, their study examined city zones with elevated pollution levels, pinpointing Avenida da República as the area with the highest concentration of carbon monoxide. The study also examined Waze traffic data, stating that Avenida da República is one of the areas with the highest traffic levels in the city, highlighting and relating the impact of vehicular traffic on air quality.

Noise is another consequence of traffic. Hahad et al. (2023) claim that exposure to noise can impact and damage the central nervous system. Prolonged exposure to noise can lead to permanent threshold changes and hearing loss in certain frequency bands, known as Noise-Induced Hearing Loss, a condition that is one the leading causes of preventable hearing loss, affecting approximately 10 million adults and 5.2 million children in the United States (Seidman & Standring, 2010).

Superblocks are a viable strategy to lower both air and noise pollution, by reducing and restricting vehicular traffic and promoting the creation of more green zones in urbanized areas. Nieuwenhuijsen et al. (2024) state that the superblocks establishments in Barcelona helped to reduce both noise and pollution levels in pilot areas like Poblenou, Sant Antoni, and Horta.

As part of this research, the examination of Campo de Ourique's air quality and noise pollution during the superblock experimental introduction, incorporating data from existing monitoring stations, will be instrumental in assessing the potential impact of superblocks on reducing both noise and air pollution levels, helping to determine whether indicators in this zone exhibited any improvements in the absence of traffic and how people were impacted by it.

2.3. Traffic Evaluation

Vehicular traffic is a significant contributor to urban air pollution, noise, and GHG emissions. The constant flow of vehicles, especially in densely populated cities, releases large amounts of pollutants such as nitrogen dioxide, particulate matter, and carbon dioxide into the atmosphere (Lin et al., 2018). According to Lee et al. (2014), the relationship between air pollution and respiratory diseases is well established, with significant evidence linking air pollution to conditions such as chronic obstructive pulmonary disease and asthma.

In addition to air pollution, vehicular traffic contributes to urban noise pollution, which can cause a range of health problems, including stress, sleep disorders and cardiovascular diseases (Wawa & Mulaku, 2015). Traffic noise is particularly problematic in urban areas with high population density and high traffic volumes, exacerbating the overall impact on residents' quality of life.

As traffic congestion increases, it becomes increasingly difficult to travel quickly, leading to longer journey times, higher fuel consumption and frustration for the people who need to travel. Congestion affects the efficiency of transport systems, but it also has important economic consequences, such as reduced productivity and increased transport costs (Florin & Olariu, 2015).

To address these issues, urban planners and policymakers are exploring innovative strategies to reduce car traffic and mitigate its environmental and health impacts. One promising approach is the introduction of superblocks, but there are also other approaches with a lot of potential.

Mueller et al. (2020) have evaluated the health consequences of adopting the Superblock model in Barcelona and projected that the establishment of 503 additional Superblocks could prevent 667 premature deaths annually. This research has also converted the annually averted premature deaths into a significant economic influence of 1.7 billion euros, along with average life expectancy gains of nearly 200 days attributed to the reduction of harmful environmental exposures. The greatest number of lives preserved was linked to decreases in NO2 levels (291, 95% PI: 0–838), with road traffic noise (163, 95% CI: 83–246) ranking second.

Assessing the impact of urban policy changes, such as the establishment of superblocks, is challenging. For instance, Li et al. (2022), motivated by the deepened environmental issues and need for sustainable urbanization strategies, employed a Difference-in-Differences model to assess the impact of development zones on urbanization in China. Their study highlights the complexity of evaluating urban policies and the need for robust methodologies to measure their impacts accurately.

In addition to Difference-in-Difference models, Randomized Controlled Trials (RCTs) are another used method in urban policy evaluation. Pereira et al. (2022) claim that randomized controlled trials are very rigorous and produce high-quality evidence. The authors also state that smart cities, with their vast amounts of data, offer new opportunities for using RCTs in urban policy evaluation, enabling accurate measurement of policy impacts and identification of favourable and unfavourable effects, leading to better decision-making and policy design.

By combining these improved evaluation methods, policy makers will be better able to more effectively measure the impact of traffic reduction strategies and make informed decisions to improve the living conditions in cities. Combining innovative urban design, such as superblocks, and reliable assessment methods can significantly reduce the harmful effects of vehicle traffic in urban environments, improving residents’ health and life quality, while creating safer spaces for everyone.

3. DATA AND METHODOLOGY

For the implementation of this project, we will adopt the CRISP-DM methodology (Cross-Industry Standard Process for Data Mining), which is a trusted framework for data mining projects. Additionally, we will also employ the Difference-in-Differences (DiD) approach to evaluate the causal impact of the superblock establishment on traffic congestion and air quality indicators.

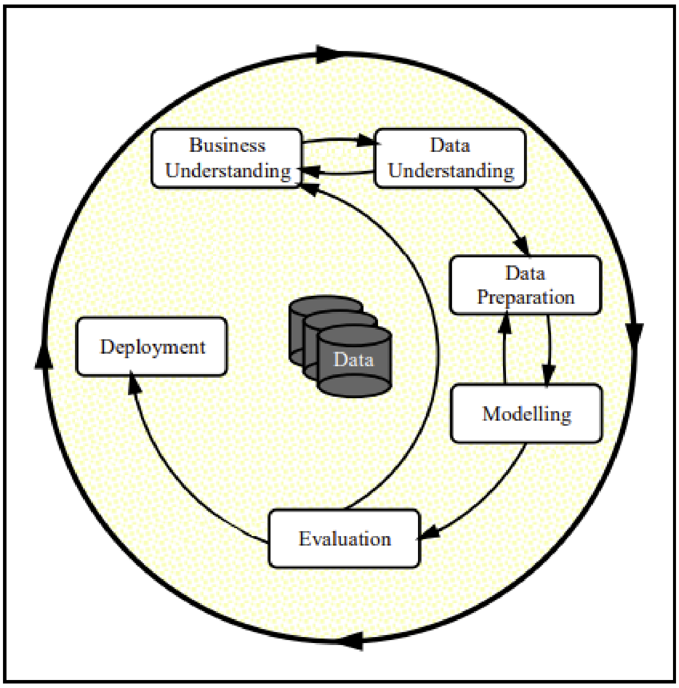

Figure 2 - CRISP-DM Methodology. Taken from (Wirth & Hipp, 2020).

According to Chapman et al. (2000), CRISP-DM provides a structured approach covering six main phases, which are:

1. Business Understanding: in this phase, we began by defining our objectives, focusing on assessing the impact of the superblock implementation in Campo de Ourique both environmentally and functionally, aiming to evaluate changes in air quality, noise levels, and traffic patterns.
2. Data Understanding: we explored and analysed data from Waze and from air quality and noise sensors that are distributed across the Lisbon municipality. This phase involved detailed analysis of the datasets to ensure their accuracy, focusing on key variables such as particulate matter (PM2.5, PM10), nitrogen dioxide (NO_2), and traffic patterns.
3. Data Preparation: in this phase, we extracted and cleaned the data using SQL and Python to remove duplicates, resolve inconsistencies, and focus on relevant streets around Jardim da Parada. Data was summarised by date and location to ensure comparability before and after implementation.

4. Modelling: the Difference-in-Difference (DiD) method was used to analyse the impact of the superblock. By comparing traffic and air quality data in the superblock (treatment group) and a similar control area (Jardim do Arco do Cego), we isolated the intervention’s effects.
5. Evaluation: we evaluated the model through statistical tests of changes in air quality indicators and traffic patterns. The results provide insight into the effectiveness of superblocks in reducing pollution and improving traffic flow.
6. Implementation: we have summarised the results of our analysis in a detailed report and different visualizations. The results highlight the impact of superblocks on air quality and traffic and provide valuable recommendations for Lisbon policymakers and urban planners, intending to support the establishment of similar urban interventions in the future.

Following the CRISP-DM methodology, we systematically guided the project from understanding the business problem to delivering actionable insights. Each stage of the process was executed to ensure the project's success and maximize its impact on event planning and management.

3.1. Business Understanding

According to the 2021 Census conducted by the National Statistics Institute, Lisbon has a population of more than 500,000 people, and over 2,800,000 people live in the Lisbon Metropolitan Area (Instituto Nacional de Estatística, 2021).

Such dense populations raise several concerns and problems, especially related to pollution, noise and traffic. The search for strategies that can mitigate these impacts and improve the population's quality of life is increasingly needed. In this context, the promotion of sustainability is increasing, and the concept of superblocks has gained popularity around the world, especially after the tremendous success seen in Barcelona and other cities.

Recently, in Campo de Ourique, the first experience of establishing a superblock in Portugal was tested, and our study will be assessing its impact. Our study will focus on providing insights to different stakeholders into the effectiveness of this urban model in mitigating the urban challenges faced by cities, as well as contributing to informed decision-making on urban policies and sustainable urban planning.

Key stakeholders include:

- City Councils: this research is very important to the city councils of Portuguese metropolitan areas as it provides a lot of insights and valuable information into the benefits and effectiveness of superblocks in improving urban life’s quality. This information helps city authorities to take data-driven decisions and evaluate the possibility of introducing similar policies.

- Tourists: this study is important for tourists because it can influence their overall travel experience and their decisions before deciding where to travel. According to Luongo et al. (2023) "in 2023, 69% of tourists in 2023 consciously opted for sustainable travel options.". Sustainable cities are becoming more attractive to tourists, as more sustainable urban environments provide better experiences, offering a more enjoyable setting.
- Residents: this study gives residents in Portuguese metropolitan areas hope for a better and healthier life. Reducing pollution, noise and traffic jams actively contributes to improve residents mental and physical health (Wawa & Mulaku, 2015). The creation of more green spaces and pedestrian-friendly areas promotes physical activity and provides a safer environment.

3.2. Data Understanding

In order to better understand the impact of the superblock on the various sectors under study, four different datasets were analysed, by stablishing a direct connection to a SQL server.

Table 1 summarises the traffic jam data from the Waze platform dataset variables. This dataset, composed of data that ranged from April 2019 to November 2023, across all the Portuguese territory (including continental territory and islands), will serve as a basis to analyse traffic during and prior to the superblock implementation in the region of Campo de Ourique.

Table 1 - Description of the characteristics present in the "Waze Jams" dataset.

Variable	**Description**
City	Name of the city where the reported event occurred or where the traffic information was generated.
length	Length of congestion in metres
street	Street name
endNode	Nearest exit to the end of the congestion
type	Type of traffic collection zone
speed	Average traffic speed (m/s)
uuid	Unique identifier assigned to each event or traffic report generated by Waze.
start_time	Start time of the event or traffic condition reported.
country	Country code
turntype	Curve type
level	Traffic congestion level (0 = no traffic, 5 = completely congested)
delay	Traffic delay time compared to the completely free lane
speedKMH	Average traffic speed (km/h)
roadType	Type of road (categorical variable)
pubMillis	Timestamp indicating when the congestion data was published.
blockingAlertUuid	Unique identifier for road block or obstruction alerts
startNode	Initial reference point for locating a traffic jam on the Waze road network.

Data on air quality and noise collected by different sensors in the city of Lisbon were also analysed. This data is subdivided into three tables.

Table 2 contains the values from each air quality and noise indicator. When this dataset was accessed, it was composed of four different columns and 47 million rows, containing data from the Lisbon municipality from June 2021 to January 2024. These values are expected to increase in the future, as the dataset is continuously updated with new data.

Table 2 - Air Quality and Noise Measurements

Variable	Description
id	Unique identifier of the measurement
date	Date
value	Measured Value
pk	Sensor's unique identifier

Table 3 contains descriptive information about the different air quality and noise sensors, as well as details on the various measurements and sensor locations. This dataset originally consisted of 713 rows and 13 columns, but only the 10 described columns were selected for the analysis.

Table 3 - Measurements Details

Variable	Description
lat	Latitude
lon	Longitude
id	Measure identifier
unid	Measurement unit
indicator_id	Indicator's identifier
measure_id	Measure's Identifier
location_id	Location's identifier
measure_description	Measure's description
Indicator_description	Indicator's description
measure_description_secondary	Indicator's name
measure_unit_description	Indicator's measure name

Table 4 also consists of descriptive information. Composed of 43 rows and 4 columns, it contains information about the different evaluated measures, including their unique identifiers and minimum and maximum value.

Table 4 - Measurement Ranges

Variable	Description
Measure_id	Unique identifier of the measure
valor_min	Minimum value for the measure
Valor_max	Maximum value for the measure
classe	Color ranking for the measure’s values

3.3. Exploratory Data Analysis

3.3.1. Air Quality and Noise Sensors Analysis

In the preliminary stage we employed exploratory data analysis methods to help us uncovering potential patterns, trends, and relevant correlations.

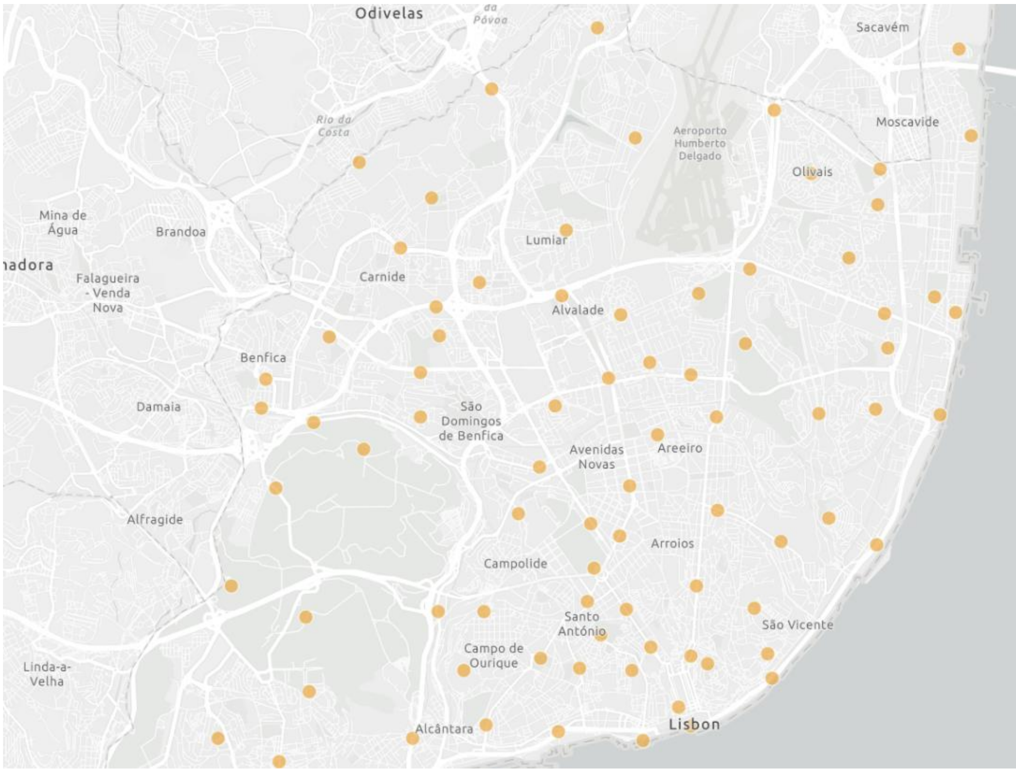

Figure 3 - Air Quality and Noise Sensors

Figure 3 represents a map with the distribution of the points where the various noise and air quality meters are located in the city of Lisbon.

Figure 4 is a map of the Campo de Ourique region representing the four different air quality and noise sensors that were used for the Campo de Ourique’s analysis. The red rectangle shows the boundaries of the superblock, and the green area within it marks the Jardim da Parada boundaries. This map allows us to visualize the streets that were affected by the introduction of the superblock. Although only 9 blocks were not accessible to car traffic, several other streets and traffic dynamics were affected by implemented change, due to different street connections being closed.

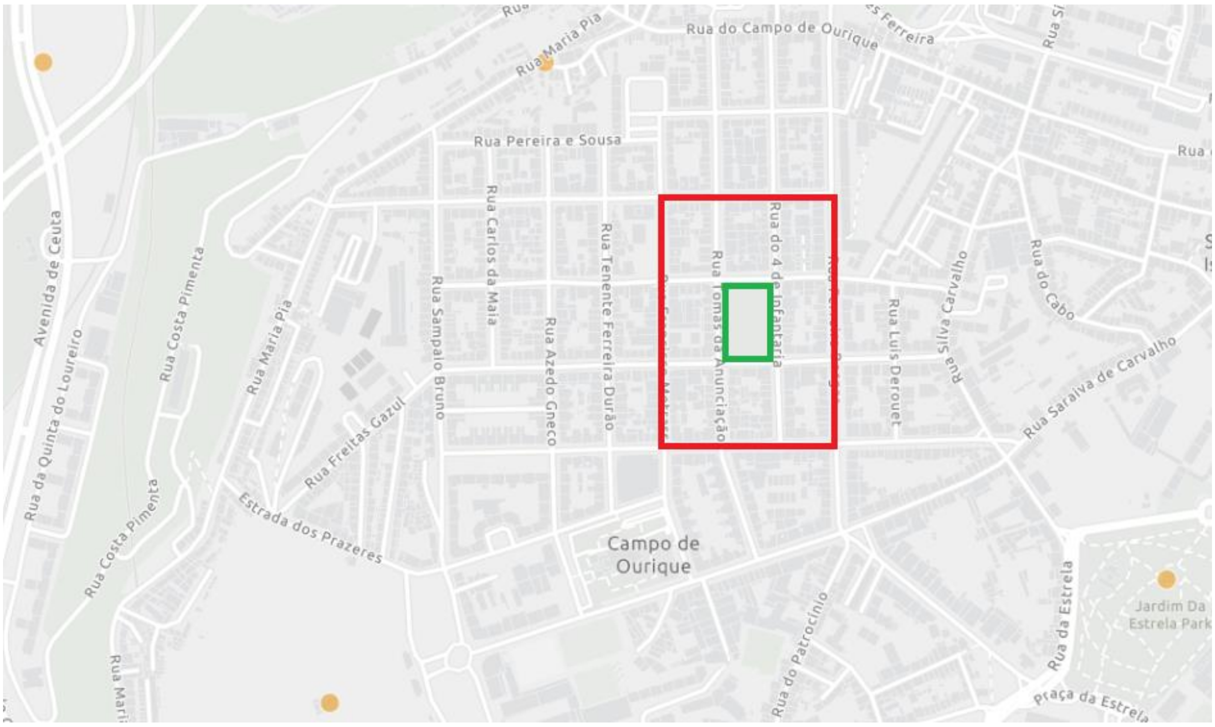

Figure 4 - Superblock Area and Air Quality and Noise Sensors

We have decided to perform a preliminary analysis over a 6-week period to examine the variations in different air quality and noise indicators throughout the days of the week. This time frame was chosen to avoid variations in traffic patterns that could arise from seasonal changes, ensuring that any observed changes were attributable to the superblock implementation rather than other factors. The air quality analysis revealed some patterns, with Carbon Monoxide (CO), particulates < 2.5 µm (PM2.5), particulates < 10 µm (PM10), and Nitrogen Dioxide (NO2) concentrations peaking their concentrations during the middle of the week (between Tuesday and Thursday), and reducing through the weekend, reflecting higher weekday traffic and industrial activity.

The noise analysis revealed that noise levels remain constant during the week, but also shows a small reduction over the weekend, indicating reduced human and vehicular activity during the weekend.

While this preliminary analysis only provided a general overview and lacked more detail, it was indispensable in guiding our further exploratory analysis. The results in Figures 5 and 6 confirm that weekdays and weekends exhibit different patterns in air quality and noise levels, which was expected, due to the increased traffic and industrial activity on weekdays. Therefore, to ensure fair and accurate comparisons and provide a more accurate reflection of the data, we decided to analyse weekends and weekdays separately.

Given that the superblock experience lasted for only 9 days, with 4 of those days being weekends, this distinction is particularly important. As an example, if we decided to analyse the superblock as a whole, it could reflect disproportional and skewed results. By analysing weekdays and weekends separately, we guarantee a more accurate and reliable assessment of the superblock implementation's impact.

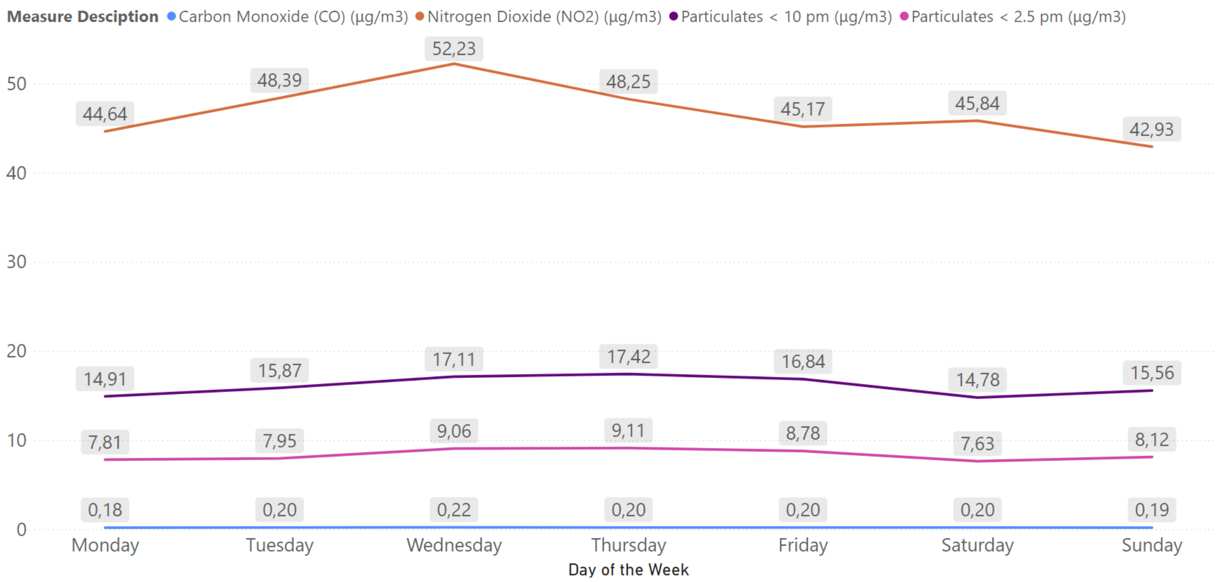

Figure 5 - Average Air Quality Values (µg/m3)

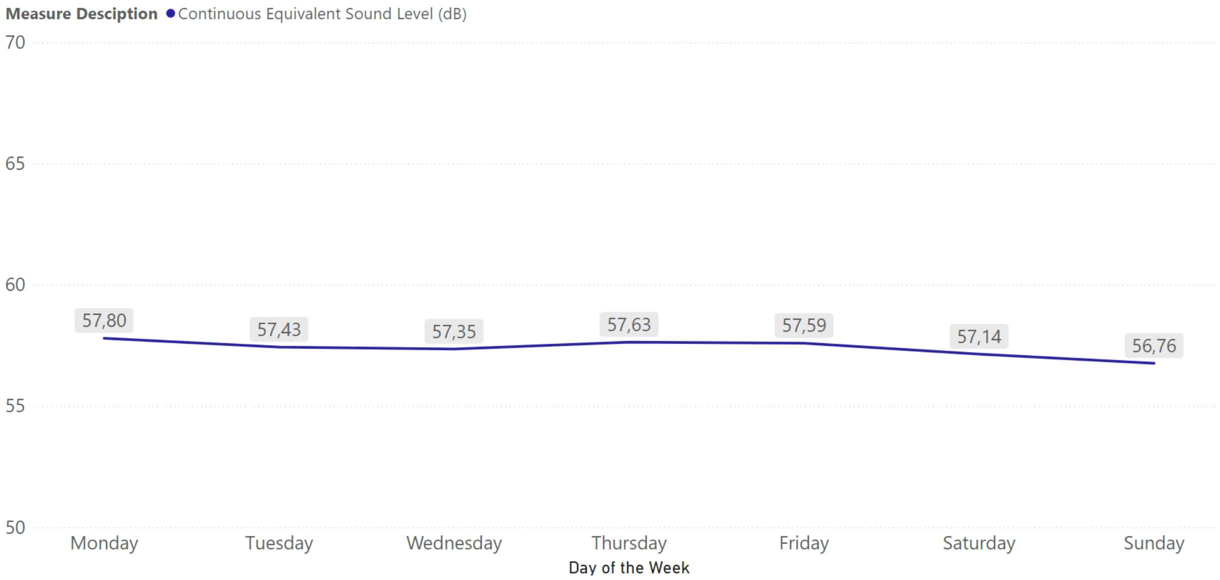

Figure 6 - Average Noise Values (dB)

3.3.2. Traffic Analysis

Similarly to our approach with the air quality and noise data, we performed a weekly analysis on different traffic variables, including delay, length, and average traffic speed. Due to the lack of Waze data for the month of August, our traffic analysis was limited to a 2-week period, unlike the 6-week period that we established in the air quality and noise data analysis.

Despite this shorter time frame, it was enough to uncover significant patterns and changes in traffic behaviour. This analysis was also restricted to Campo de Ourique zone, covering 7 days during the superblock implementation and 7 days before.

In the case of the analysis of the traffic average delay (Figure 7), it was noticeable that the pattern was similar to observed patterns in the air quality and noise analysis – it peaks during mid-week, achieving its highest on Wednesday, and reduces through the weekend, reaching its' lowest on Sunday.

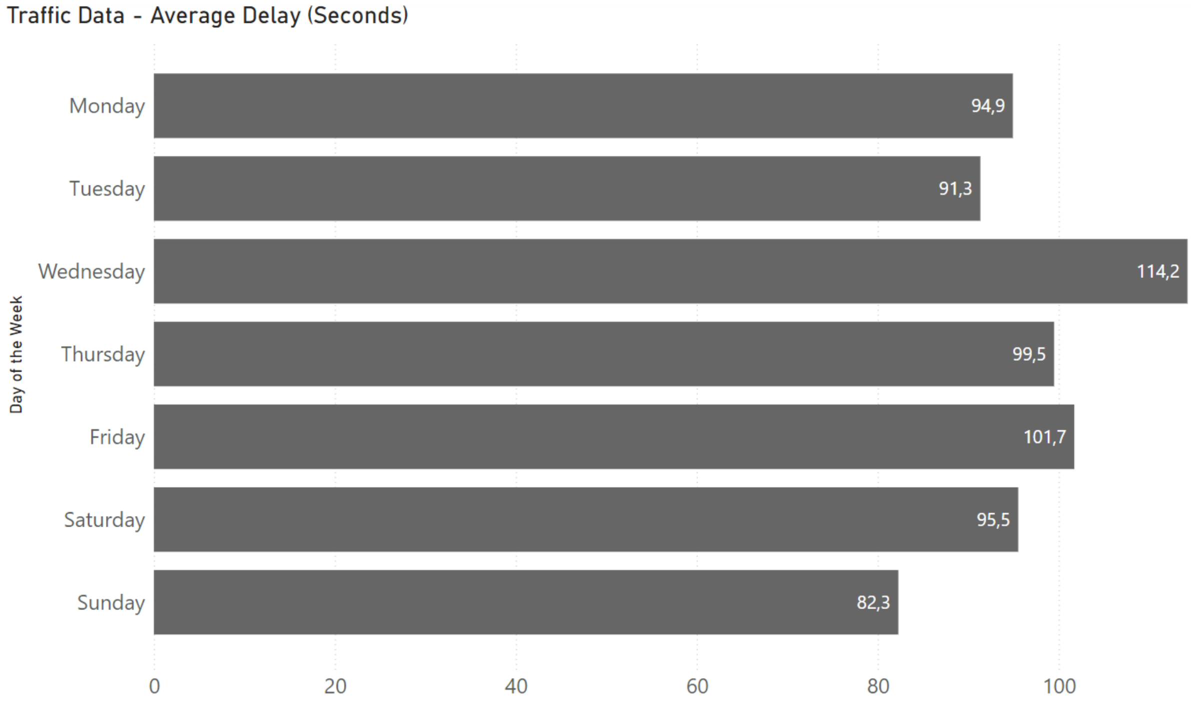

Figure 7 - Traffic Data Average Delay (Seconds)

When analysing the average speed of traffic queues (Figure 8), we noticed a different trend compared to the ones we analysed before. The traffic's speed stayed consistent throughout the weekdays, peaking on Saturday and hitting its lowest point on Sunday. This lack of significant variation and the observed differences between Saturday and Sunday, made it difficult to identify any clear patterns from the data.

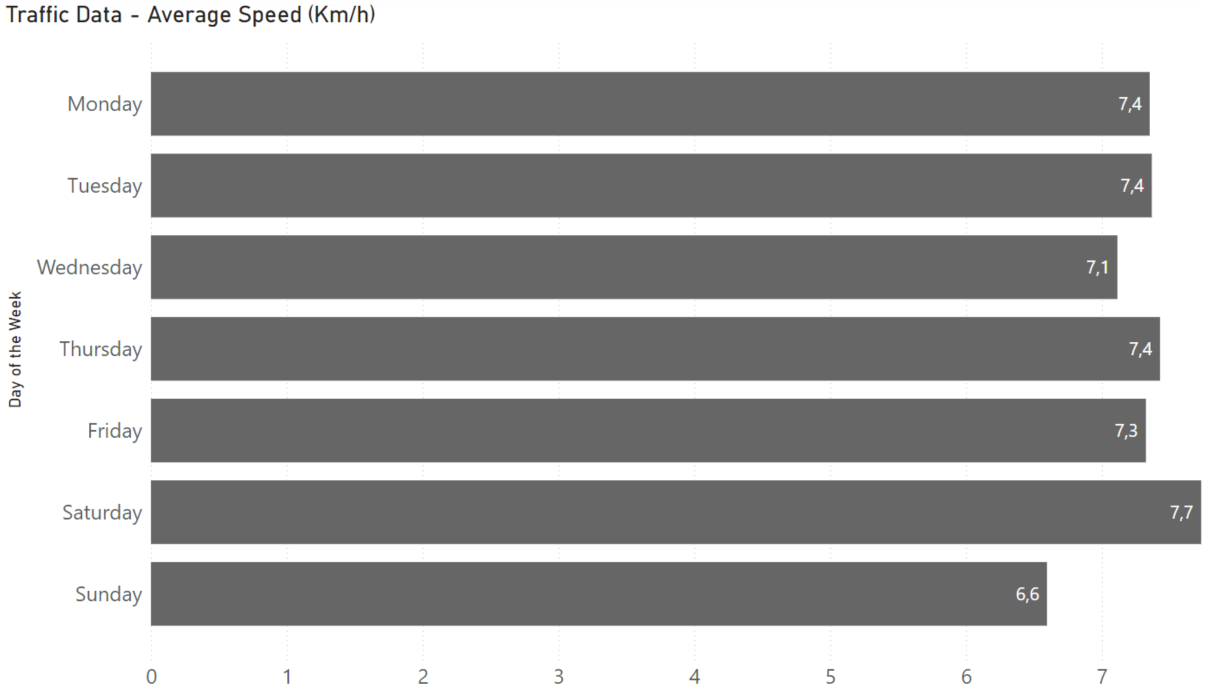

Figure 8 - Traffic Average Speed (km/h)

When analysing the traffic's average length (Figure 9) through the weekdays, we saw a pattern that was identical to the identified pattern in the traffic's average speed, remaining fairly constant through the week, peaking on Saturday, and reaching its lowest value on Sunday. This suggests that traffic stays pretty consistent during the weekdays, but changes on the weekends, probably due to different activities and travel habits.

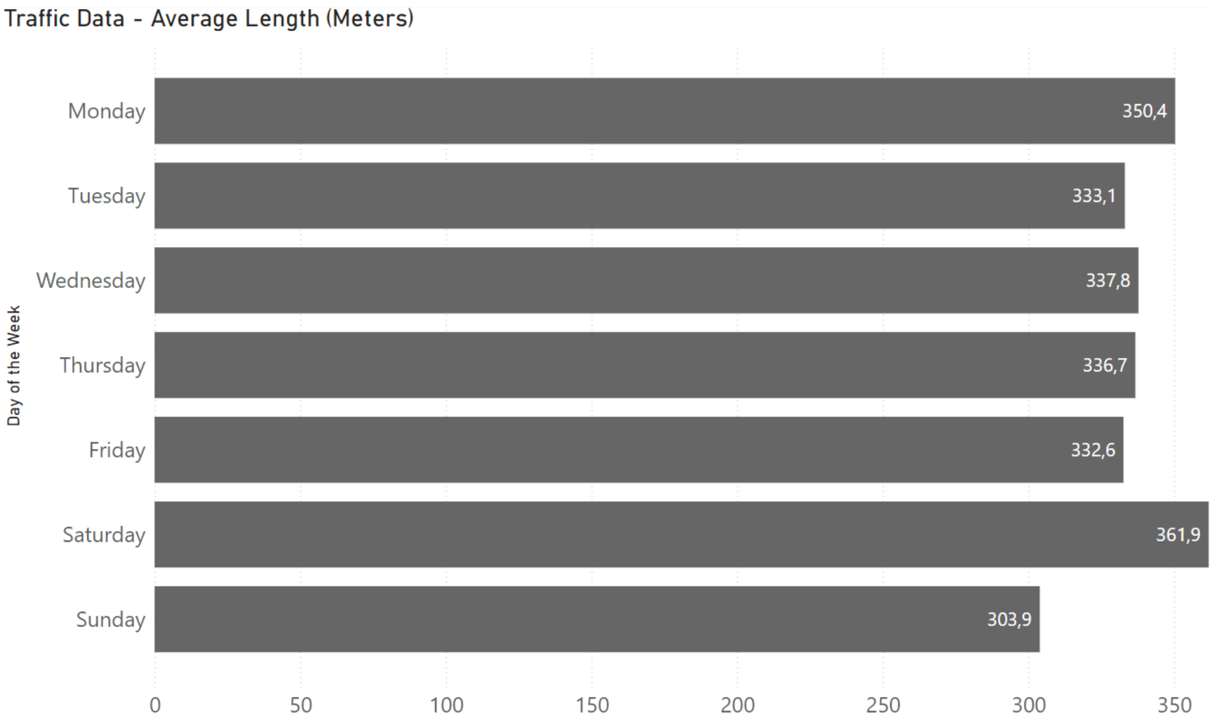

Figure 9 - Traffic Average Length (Meters)

3.4. DATA PREPARATION

In the data preparation stage, different steps were carried out to prepare the dataset for analysis, involving refining, cleaning, transforming and structuring the data to maximize its quality and usability.

We started the process by combining Python and SQL queries within a Jupyter Notebook environment to extract both sensors and Waze data from a SQL Server. The sensor's data was composed of two descriptive tables and one record table, with the descriptive tables being fully extracted due to their reduced size and relevance for the analysis. On the other hand, the sensor record table had millions of lines, and therefore, was selectively extracted based on sensor IDs and their location. For Waze's data, we identified a specific set of roads to be extracted, which were the roads around Jardim da Parada and the roads intersecting them, as these would be the most impacted roads by the implementation.

After these steps, we proceeded with the data treatment and merged historical and recent traffic jam data, which were originally divided into different datasets. We proceeded with the data treatment by removing duplicate values, inconsistencies, and wrong information from datasets. We have also created new variables, consolidated redundant column categories, and grouped data based on dates and locations.

3.5. MODELLING

Furquim et. al (2020) describe the Difference-in-Differences (DiD) approach as a statistical technique used to estimate causal relationships by comparing the differences in outcomes over time between a treatment group and a control group. This method isolates the effect of the superblock intervention by comparing these two groups, which were subjected to similar conditions during the same periods, except for the superblock implementation in the treatment group.

Difference-in-Differences (DiD) is a widely utilized methodology in impact evaluation studies, renowned for its intuitive approach that combines before-after and treatment-control group comparisons (Fredriksson & Oliveira, 2019). It is particularly valuable in non-experimental settings where random assignment of treatment is not feasible.

The DiD model is specified as:

$$Outcome_{it} = \alpha + \beta_1 Location_i + \beta_2 \text{Treatment}_t + \beta_3 (\text{Location}_i \times \text{Treatment}_t) + \epsilon_{it}\}$$

Where:

- Outcome_{it} represents the outcome of a certain variable for location *i* at time *t*.
- α is the intercept, representing the baseline outcome level at Jardim do Arco do Cego before the treatment.

- Location_i is a binary variable that equals 1 if the observation is from Jardim da Parada and 0 if it is from Jardim do Arco do Cego.
- Treatment_t is a binary variable that equals 1 for the period during the superblock implementation and 0 for the period before.
- $\text{Location}_i \times \text{Treatment}_t$ is the interaction term, capturing the differential impact of the superblock on NO2 levels at Jardim da Parada relative to Jardim do Arco do Cego.
- ϵ_{it} is the error term.

Coefficients are interpreted as follows:

- α: Baseline outcome level at Jardim do Arco do Cego before the superblock implementation.
- β_1: Difference in outcome levels between Jardim da Parada and Jardim do Arco do Cego before the treatment.
- β_2: Change in NO2 levels at Jardim do Arco do Cego after the treatment period starts.
- β_3: Difference-in-Differences estimator, indicating the additional effect of the treatment on outcome levels at Jardim da Parada relative to Jardim do Arco do Cego.

Unlike randomized controlled trials (RCTs), which can directly compare outcomes between randomly assigned treatment and control groups, DiD is designed to estimate causal effects by leveraging existing data from naturally occurring treatment and control groups (Wing et al., 2018).

Ideally, DiD should be applied to two different groups: one that has undergone a specific treatment or intervention and one that has not. The main idea is to measure the difference in outcomes between these groups before and after treatment. By comparing changes in outcomes over time between the treated and untreated groups, researchers can isolate the effect of treatment from other confounding factors that might affect outcomes.

As Roth et al. (2023) state, in Difference-in-Differences, "the key identifying assumption is that the average outcome among the treated and comparison populations would have followed 'parallel trends' in the absence of treatment.". Knowing if that assumption would be true is a hard challenge, and a potential limitation for the Difference-in-Differences methodology. Zhou et al. (2016) say that "While this approach accounts for unobservable variables that are fixed over time, the biggest issue is that it does not account for unobservable variables that are not fixed over time." In most cases, unobservable variables that are not fixed over time will be not easy to control, and therefore pose a significant challenge to the validity of DiD estimates.

4. RESULTS AND DISCUSSION

In this section, we present the findings of our study, focusing on the effects of the superblock implementation, organized into three areas. First, we explore air quality and noise data, assessing variations in pollutant concentrations and noise levels during and prior to the experience, highlighting the environmental impact. Then, we proceed to an analysis of the traffic data, examining traffic patterns before and during the superblock period, which helps us understand changes in congestion and flow. Finally, in the last subsection, we cover model implementation, detailing our methodology and the findings of the Difference-in-Differences (DiD) analysis, providing insights into the statistical significance of the observed changes.

4.1. Air Quality and Noise Exploratory Analysis

In terms of NO2 concentrations, the superblock introduction led to mixed results. While there is an increase in the NO2 concentrations during weekdays, there is also a decrease in these levels during the weekend (Figure 10). The increase in the NO2 concentrations during the weekdays may have been caused by external factors such as industrial activity or may have been a consequence of the implementation, as the traffic changes may have led to increased congestion and higher emissions during peak traffic times on weekdays.

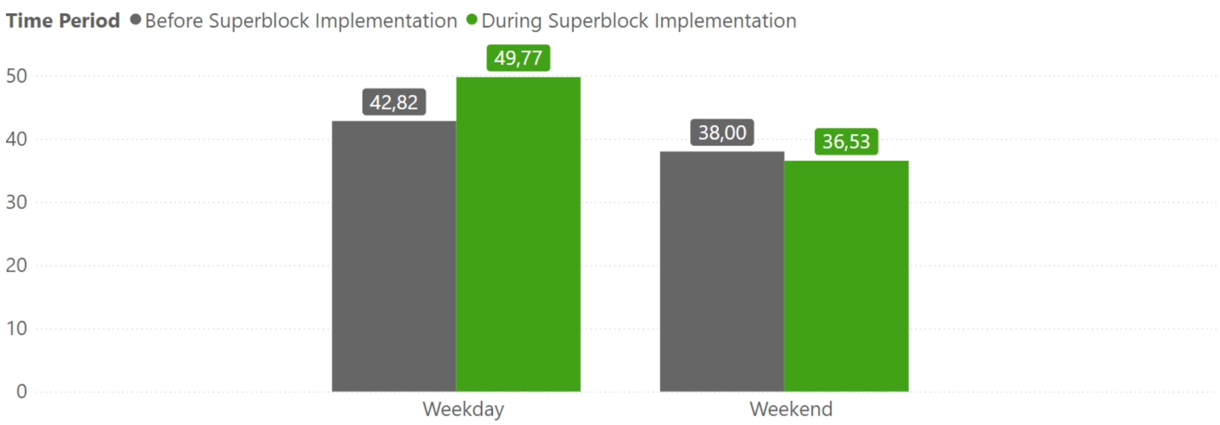

Figure 10 - Average NO2 Concentrations (μg/m³)

A daily analysis (Figure 11) of the superblock period compared with the previous six weeks revealed that the highest NO2 emissions were recorded during the superblock period and that most of the analysed days (5 out of 9) during the superblock period had NO2 concentrations above the average, supporting the findings of the analysis.

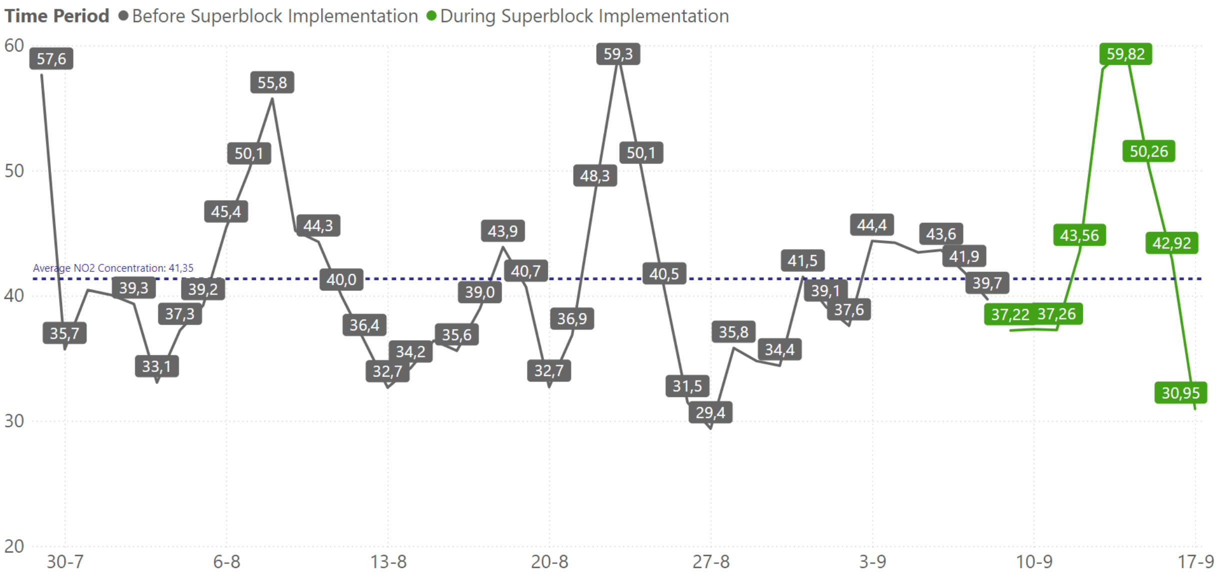

Figure 11 - Daily Average NO2 Concentrations (µg/m³)

A similar pattern was observed in the carbon monoxide (CO) concentrations (Figure 12). While there were no noticeable changes in the gas concentrations over the weekends, a significant increase was detected during the superblock weekdays. As with the increase observed in NO2 concentration, the causes may related with external unknown factors or an increase in traffic flow caused by the changes imposed by the superblock implementation.

The daily analysis (Figure 13) also supports these findings, with 5 out of the 9 superblock days exhibiting higher CO concentrations than the average. As observed with NO2 concentrations, the highest registered CO concentration during the 7-week analysis period was also recorded during the superblock establishment.

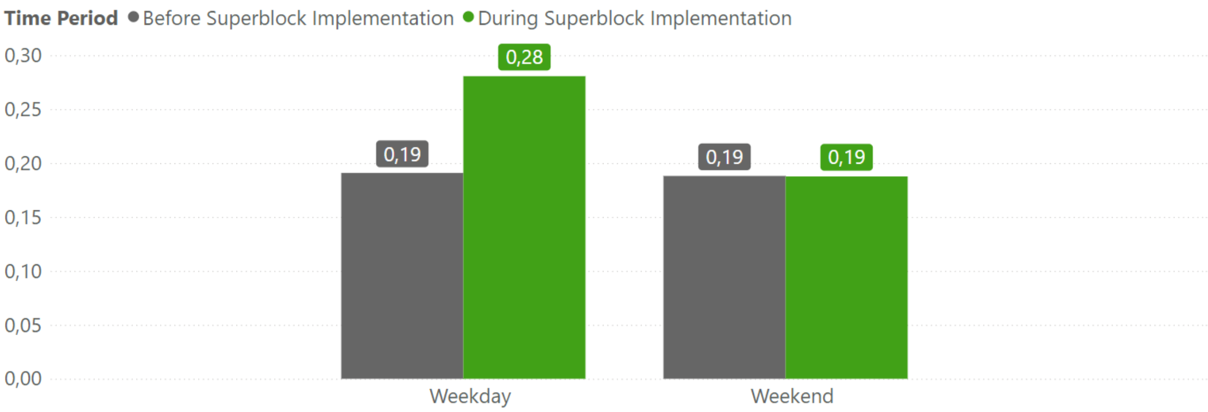

Figure 12 - Average Carbon Monoxide Concentrations (µg/m³)

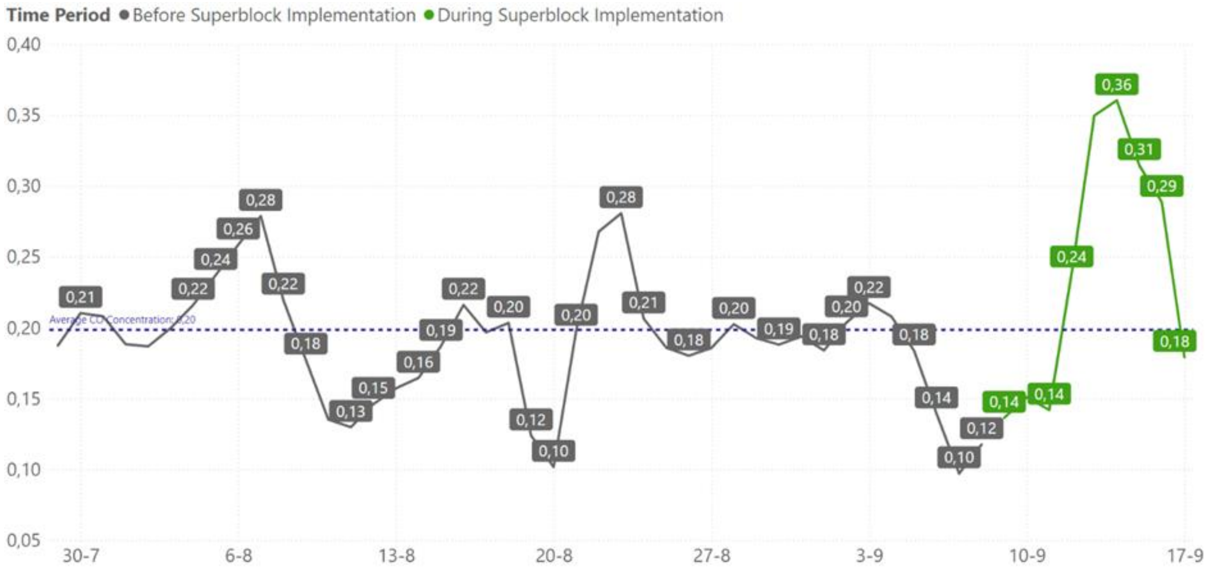

Figure 13 – Daily Average Carbon Monoxide Concentrations (µg/m³)

In the study of particulate matter with diameters smaller than 2.5 micrometres (Figures 14 and 15), there was a substantial reduction in the concentration of these particulates during weekdays suggesting that the implementation of the superblock helped to reduce the concentration of these particulates in the air. This result is not in line with the results of the concentrations of both Carbon Monoxide and Nitrogen Dioxide that increased during the superblock introduction, and the same can be said about the weekend results, where there was a small increase in particulate matter concentrations.

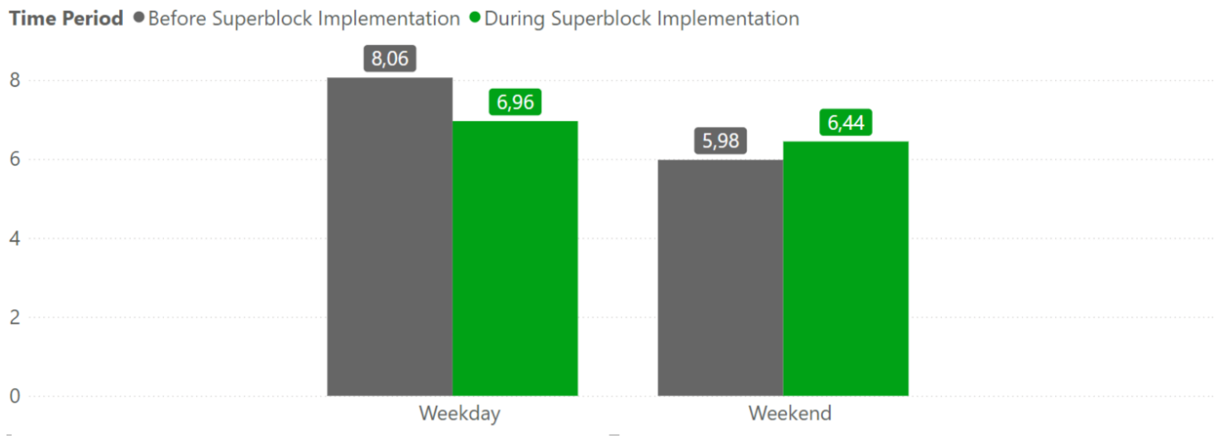

Figure 14 - Average Particulate Matter < 2.5 Micrometers Concentrations (µg/m³)

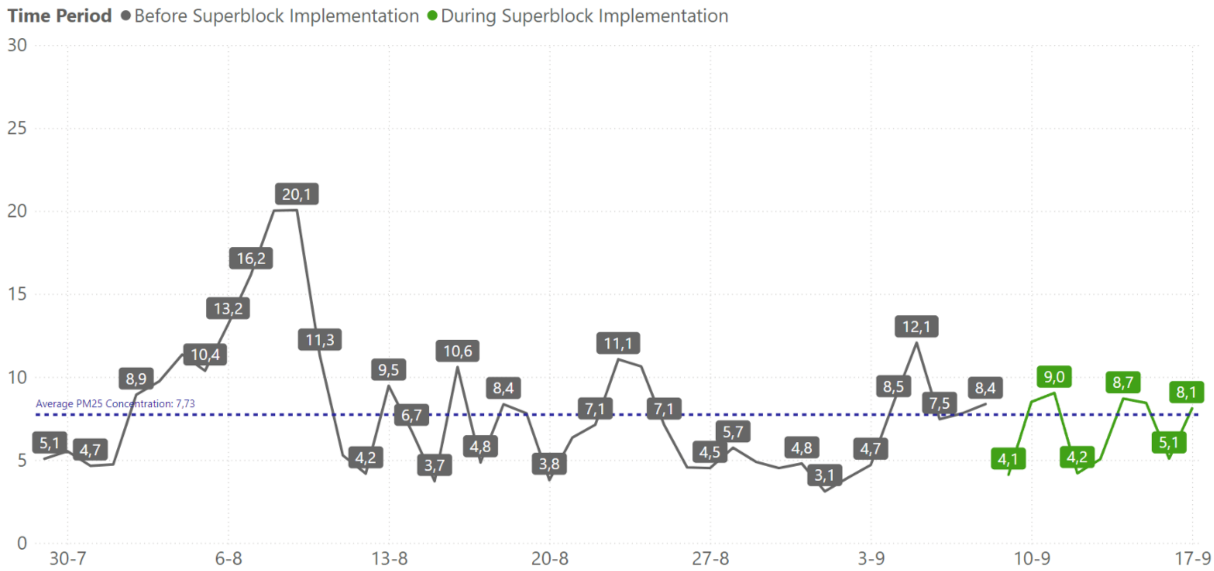

Figure 15 – Daily Average Particulate Matter < 2.5 Micrometers Concentrations (µg/m³)

Regarding the study of particulate matter with diameters smaller than 10 micrometres (Figures 16 and 17), the results were proportional to the ones that were achieved for the concentrations of particulate matter < 2.5 micrometers. While there was a reduction of the concentration of these particles during the weekdays, during the weekend a small increase was observed. The daily analysis also presented a similar pattern, with the days of the superblock showing higher levels of PM10 particles being close to the average, and the remaining days had values well below the baseline.

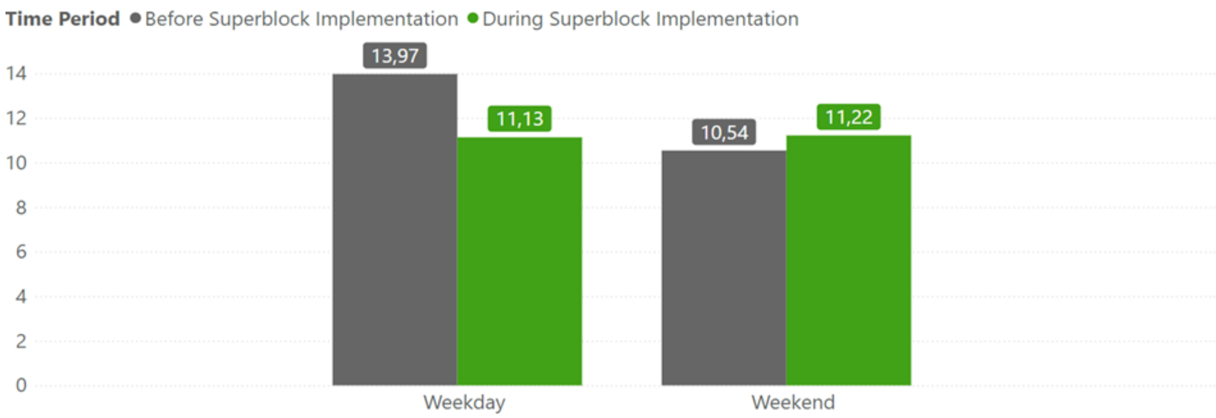

Figure 16 - Average Particulate Matter < 10 Micrometers Concentrations (µg/m³)

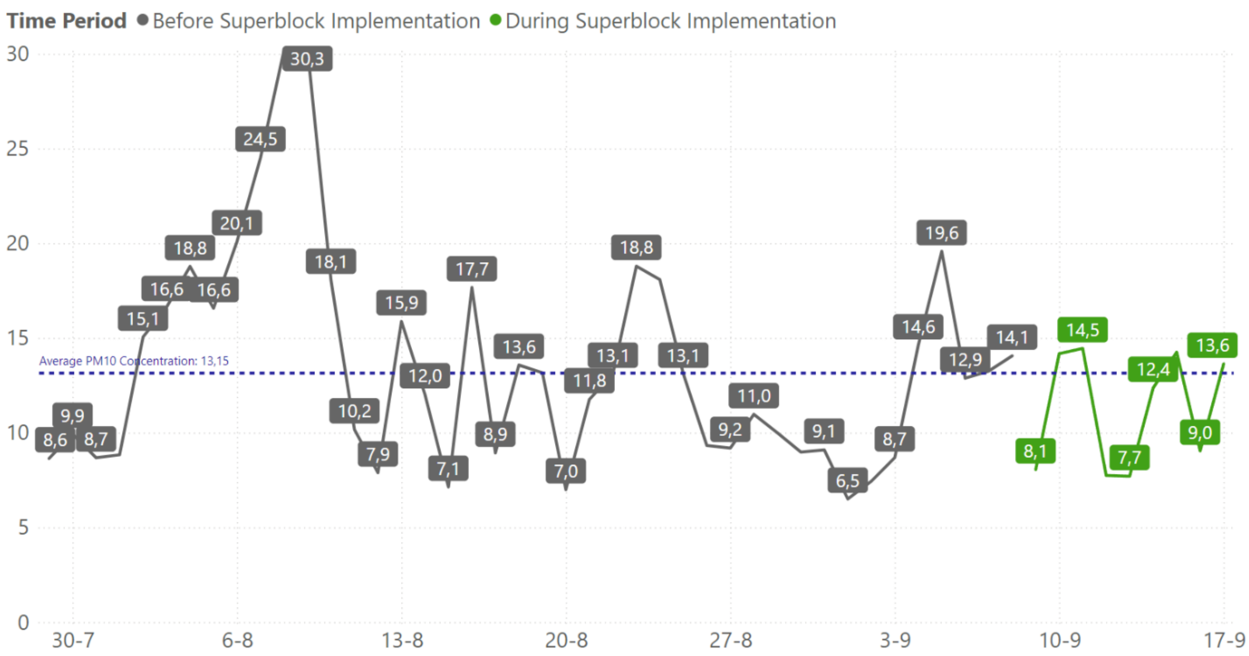

Figure 17 - Daily Average Particulate Matter < 10 Micrometers Concentrations (µg/m³)

Finally, noise levels were also studied in the Campo de Ourique region, but the results were less conclusive. The data (Figures 18 and 19) showed negligible differences between the weekends before and during the superblock experience, making it difficult to draw any definitive conclusions about the impact on noise pollution. The daily analysis revealed that only 1 of the 9 days of the implementation had an average continuous equivalent sound level under the average value, with this day also registering the lowest value of the analysed period.

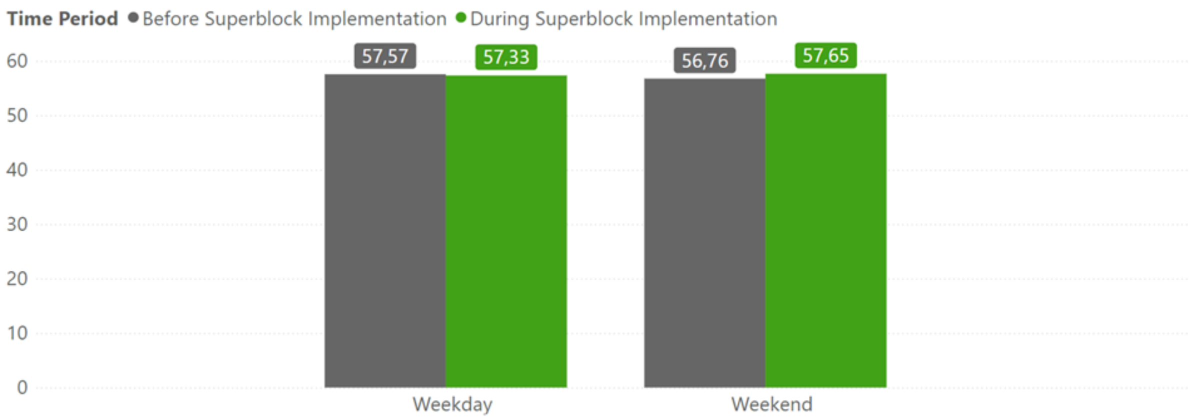

Figure 18 - Average Continuous Equivalent Sound Level (dB)

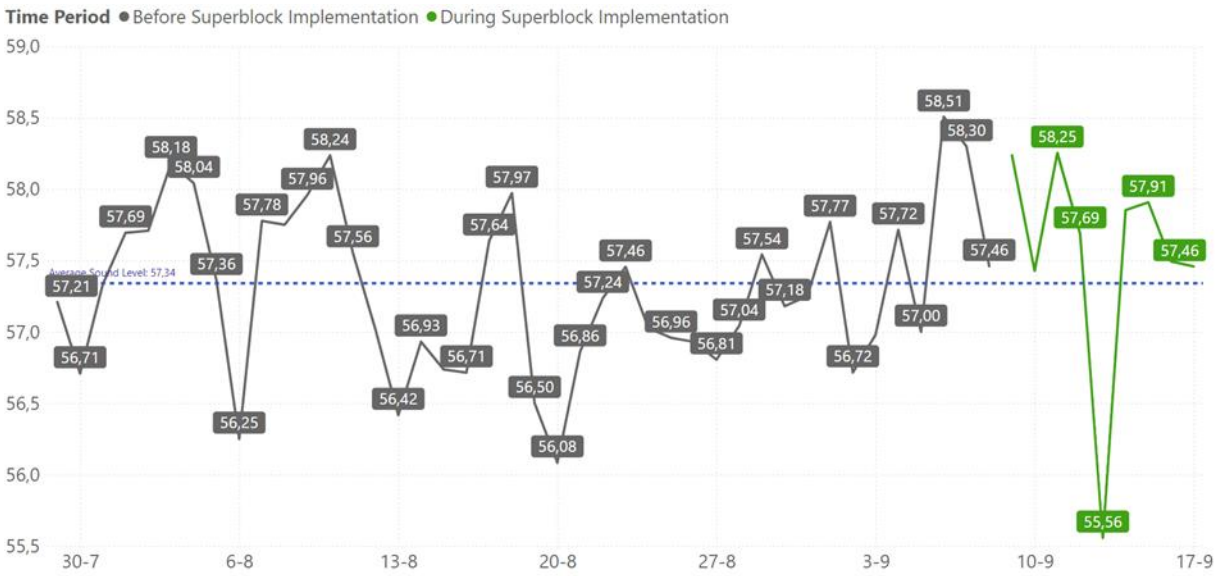

Figure 19 - Daily Average Continuous Equivalent Sound Level (dB)

4.2. Traffic Data Exploratory Analysis

In the traffic data analysis, the analysis was limited to seven days of the superblock experience, due to constraints in data accessibility from Waze's database, which did not include the implementation's first-weekend traffic data (9th and 10th of September). Similarly, there were data limitations for the period before the superblock establishment that we intended to analyse, with the provided dataset including only data after September 1st, allowing us to collect data from only seven different days before the superblock initiation.

Despite the several limitations, the collection of data from both periods still allowed us to perform a structured comparative analysis. By aligning the same number and types of days (i.e., 5 weekdays and 2 weekend days) in both pre and post-implementation stages, we ensured that the comparison was methodologically correct and provided fair and meaningful insights into the traffic pattern changes attributable to the superblock initiative.

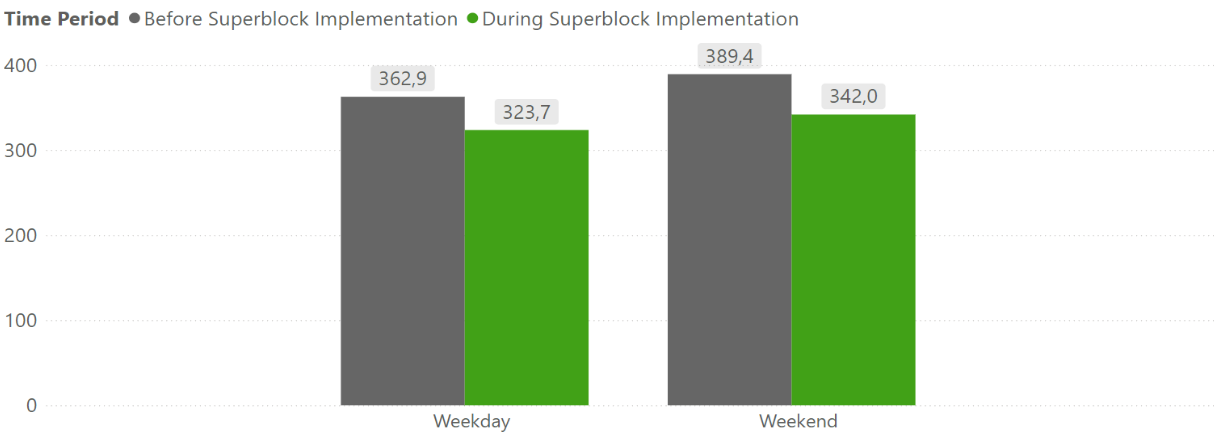

Figure 20 - Average Traffic Queue Length (Meters)

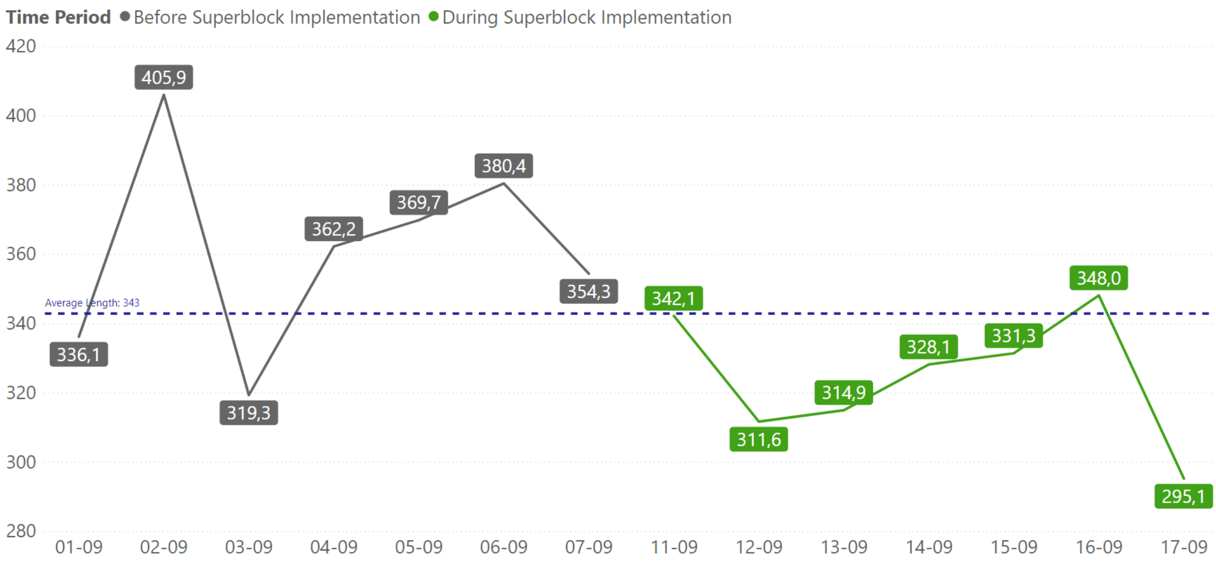

Figure 21 - Daily Average Traffic Queue Length (Meters)

Regarding the analysis of traffic queue lengths before and after the superblock introduction (Figures 20 and 21) several changes were detected. Before the implementation, the average traffic queue length was 362.9 meters on weekdays and 389.4 on weekends, with 5 out of the 7 analysed days having traffic queues exceeding the general average length of 343 meters. In contrast, during the superblock implementation, only 1 of the 7 analysed days had an average queue length above the general average. During this period, the average queue length reduced to 323.7 meters on the weekdays and 342.0 meters on the weekends, representing an 11% decrease for weekdays and a 12% reduction for weekends, which seems to indicate that the superblock introduction effectively reduced traffic congestion, as evidenced by the lower average queue lengths and more stable traffic patterns.

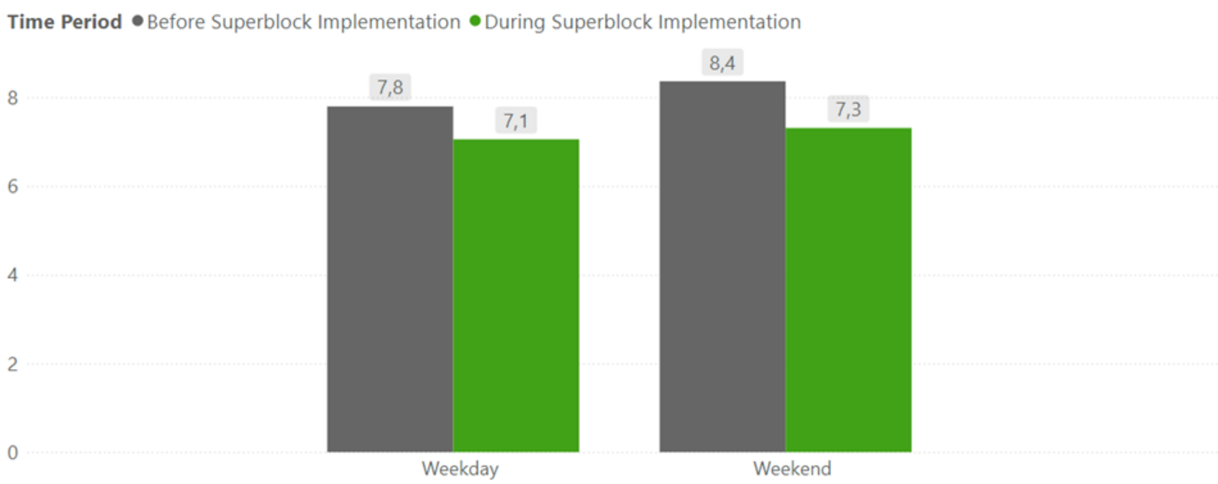

Figure 22 - Average Speed Traffic (km/h)

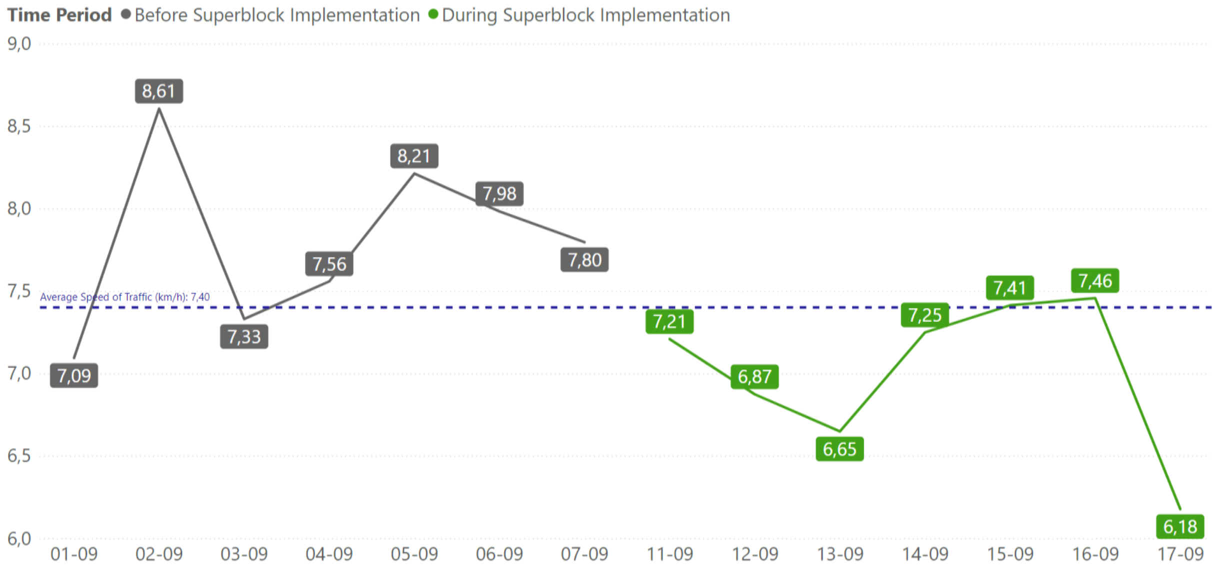

Figure 23 - Daily Average Speed Traffic (km/h)

Upon comparing traffic speed before and after the superblock introduction (Figures 22 and 23), several changes were also detected. The average speed of the queues had a 9% reduction on weekdays and a 13% reduction on weekends during the experimental implementation. This reduction could be attributed to traffic calming measures aimed at improving pedestrian safety and reducing vehicular traffic within the superblock area. This decrease in speed, along with the previously noted reduction in queue lengths, suggests a shift towards more stable and controlled traffic patterns.

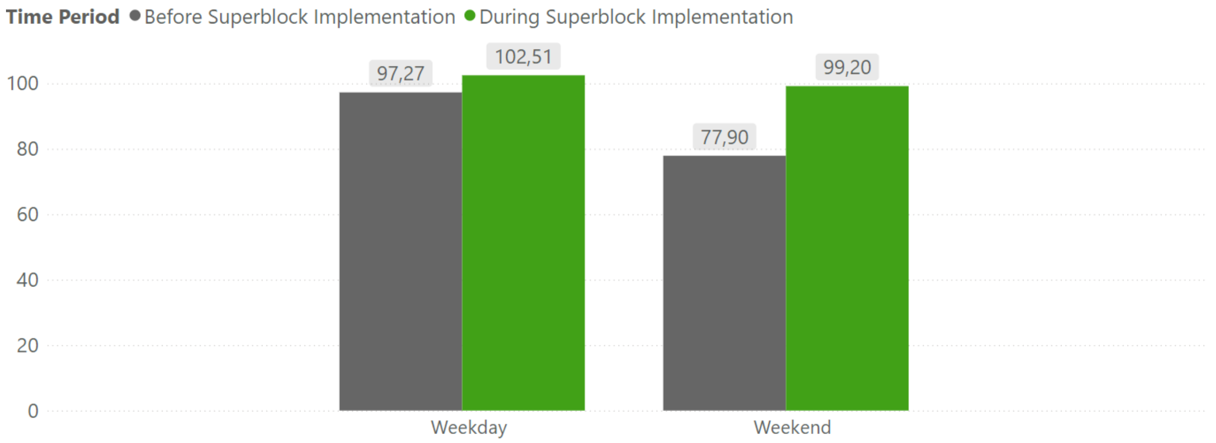

Figure 24 - Average Traffic Delay (Seconds)

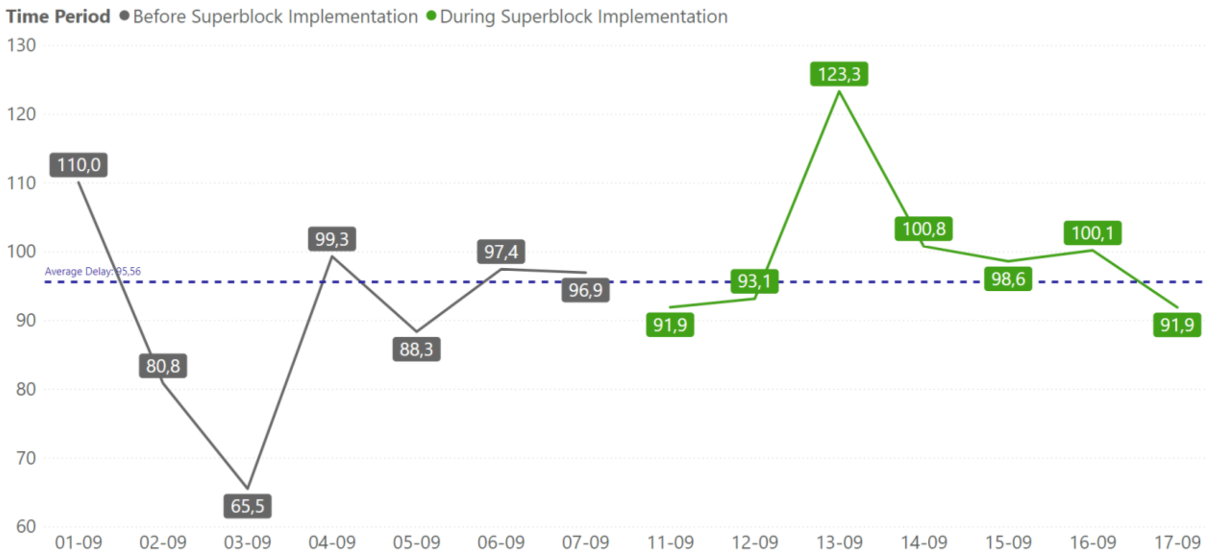

Figure 25 - Daily Average Traffic Delay (Seconds)

During the analysis of the average traffic delay, the results were once again, different for both periods (Figures 24 and 25). During the superblock experience, the average traffic delay increased to 102.51 seconds on weekdays and 99.2 seconds on weekends, representing increases of 5.4% for weekdays and 27.3% for weekends. Once again, the bigger delays could be attributed the traffic calming measures and the redistribution of traffic flows within the superblock area, which may have increased waiting times at certain intersections or areas.

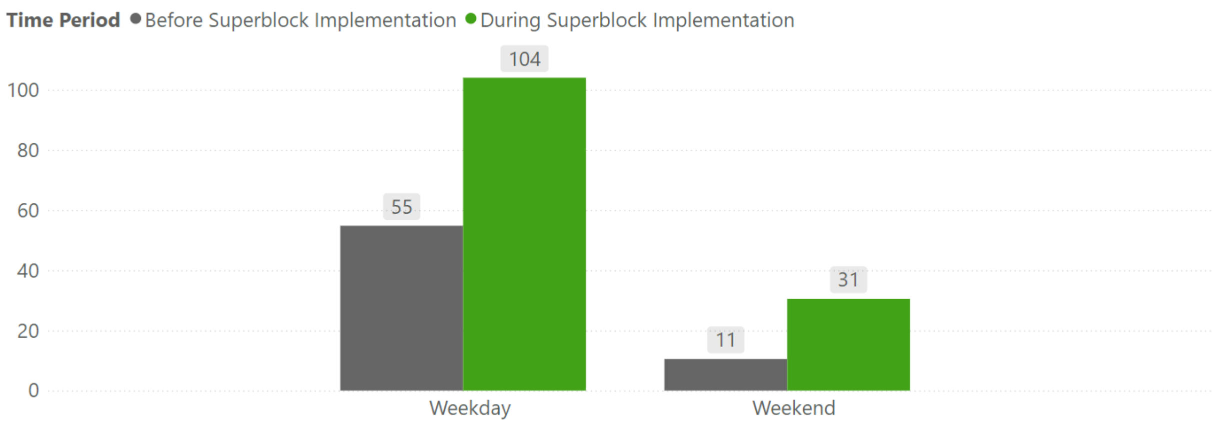

Figure 26 - Average Intensive Traffic Reports

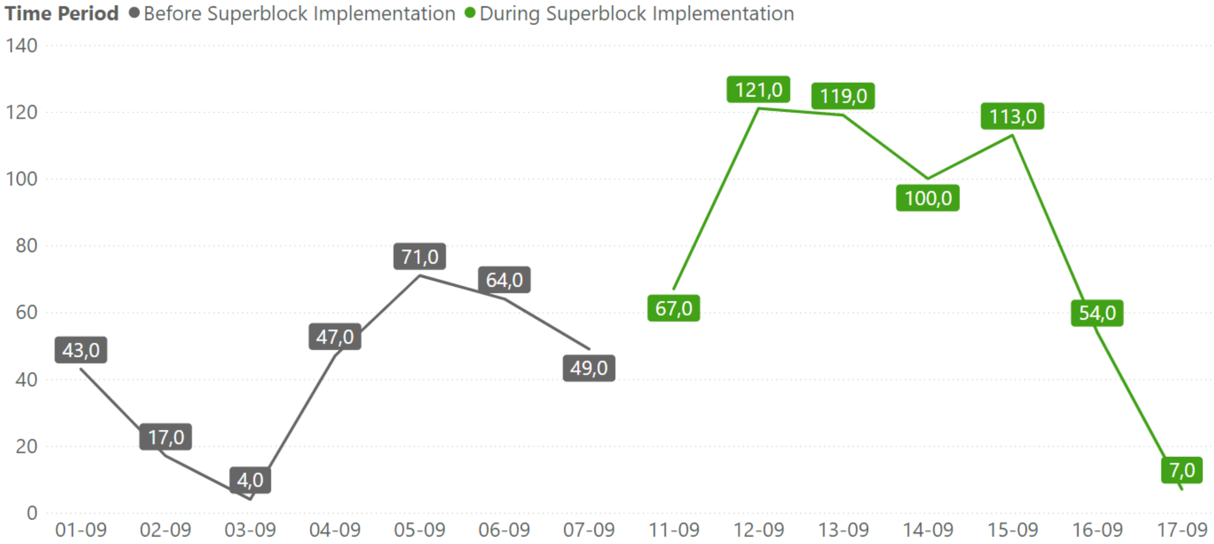

Figure 27 - Daily Average Intensive Traffic Reports

The establishment of the superblock led to an increase in the average number of traffic queue occurrences (Figures 26 and 27), confirming that the changes that were imposed on traffic flow dynamics and road closures led to an increase in traffic congestion, nearly doubling the occurrence frequency. This increase was expected, as the closure of certain roads forced all traffic onto the remaining open roads, inevitably causing more congestion than usual.

Figure 28 – Total Accident Reports in Campo de Ourique

Campo de Ourique is not an extensive region. While it experiences a lot of traffic during the week, accidents are not an everyday occurrence. Due to this, the number of accidents detected during the analysed periods before and during the superblock establishment (7 days each) is low (Figure 28), with no significant difference that can be extrapolated to broader conclusions. However, a reduction in the number of accidents was observed, and no major accidents occurred during the implementation days.

For comprehensive analysis and comparison, we delved into the impacts of the superblock on traffic patterns and air quality and analysed data encompassing traffic incidents and air pollution levels during the implementation period by making a deep dive analysis and comparing it with the weekends immediately preceding and succeeding it, over six weeks before and six weeks after the weekends of September 9th/10th and 16th/17th.

4.3. MODEL IMPLEMENTATION

In this section, we present the methodology and detailed results of the Difference-in-Differences (DiD) analysis used to assess the impact of the superblock on air quality and traffic congestion in Campo de Ourique. We have divided this analysis into two different sections: in the first one, we will focus on changes in the concentrations of different air pollutants, namely Carbon Monoxide (CO), Particulate Matter (PM2.5 and PM10), and Nitrogen Dioxide (NO2). In the second section, we will focus on traffic congestion indicators before and during the introduction of the superblock.

4.3.1. Air Quality

The data for this analysis were collected from air quality sensors located in Jardim da Parada (treatment group) and Jardim do Arco do Cego (control group), with measurements taken before and during the superblock implementation.

To apply the difference-in-differences methodology, we compared the differences in air quality indicators for the treatment group (Jardim da Parada) and the control group (Jardim do Arco do Cego) before and during the superblock implementation period.

Table 5 - Carbon Monoxide (CO) Concentrations

Location	Jardim da Parada	Jardim do Arco do Cego
Before Treatment (0)	0.20013	0.54787
During Treatment (1)	0.25230	0.57200
Difference-in-Differences		0.02804

The results for CO concentrations (Table 5) showed a slight increase in CO levels in Jardim da Parada relative to the control location.

Table 6 - Particulate Matter (PM2.5) Concentrations

Location	Jardim da Parada	Jardim do Arco do Cego
Before Treatment (0)	7.480152	5.309174
During Treatment (1)	7.518000	6.303111
Difference-in-Differences		-0.956089

The results for PM2.5 concentrations (Table 6) indicated a reduction in PM2.5 levels in Jardim da Parada compared to the control location, with Jardim da Parada registering an increase of 0.5% in this substance levels during the superblock establishment, which was considerably less than the increase of 18.7% registered in Jardim do Arco do Cego during the same period.

Table 7 - Particulate Matter (PM10) Concentrations

Location	Jardim da Parada	Jardim do Arco do Cego
Before Treatment (0)	13.439326	14.639652
During Treatment (1)	12.192200	17.160000
Difference-in-Differences		-3.767474

The results for PM10 concentrations (Table 7) showed a significant reduction in PM10 levels in Jardim da Parada relative to the control location. This was particularly interesting as the concentration values of this substance reduced by 9.3% during the superblock implementation in the treatment area, while in the control area the opposite was verified: the concentration of these particulates increased by 17.3%.

Table 8 - Nitrogen Dioxide (NO2) Concentrations

Location	Jardim da Parada	Jardim do Arco do Cego
Before Treatment (0)	41.160326	58.833500
During Treatment (1)	43.316800	76.044556
Difference-in-Differences		-15.054582

The results for NO2 concentrations (Table 8) highlighted a significant improvement in NO2 levels in Jardim da Parada compared to the control location. Despite verifying a small increase (5%) in the concentration of this substance during the superblock experience in Jardim da Parada, this increase was particularly reduced when compared to the increase that was registered in Jardim do Arco do Cego (29,2%). This substantial reduction in the difference-in-differences values underscores the effectiveness of the superblock in mitigating air pollution. As such, we will proceed by implementing an Ordinary Least Squares (OLS) regression on the NO2 data (described as follows).

$$NO2 \sim Location + Treatment + Location * Treatment$$

The developed model is statistically significant, having an F-statistic of 38.40, with a p-value of 6.32e-17. The model accounts for approximately 51.8% of the variability in the NO2 concentration (R-squared = 0.518). The adjusted R-squared has a similar value (0.505), which means that the model remains robust even when accounting for the number of predictors. Finally, the model's information criteria, AIC and BIC, are 846.1 and 856.9, respectively, suggesting a good fit.

Table 9 - OLS Regression Results for NO2

Coefficient	Estimate	Std. Error	t-value	p-value	[0.025,0.975]
Intercept	59.1390	1.567	37.732	0.000	[56.032,62.246]
Location	-18.0073	2.217	-8.124	0.000	[-22.401,-13.613]
Treatment	17.2621	4.110	4.200	0.000	[9.115,25.409]
Treatment*Location	-14.6882	5.672	-2.589	0.011	[-25.933,-3.444]

To ensure the robustness of our model, several diagnostic tests were performed (Tables 10 and 11).

Table 10 - OLS Regression Diagnostic Tests

Test	Statistic	p-value	Conclusion
Breusch-Pagan Test	2.185	0.535	No heteroscedasticity.
Durbin-Watson Statistic	1.335	NA	Some positive autocorrelation.
Jarque-Bera Test	13.807	0.001	Residuals are not normally distributed.

Table 11 - OLS Regression VIF Test

Coefficient	VIF
Location	1.191489
Treatment	2.125000
Treatment*Location	2.316489

The Breusch-Pagan test's null hypothesis wasn't rejected, confirming the absence of heteroscedasticity, as indicated by the p-value being above 0.05. While there may be some positive autocorrelation (with a value below 2 typically indicating positive autocorrelation and a value above 2 indicating negative autocorrelation), Field (2009) suggests that a Durbin-Watson statistic between 1 and 3 is usually not problematic, and our value of 1.335 falls between that range. The Variance Inflation Factor (VIF) values are all below 5, indicating no severe multicollinearity among the predictors. Despite that, the Jarque-Bera test, with a p-value under 0.05, confirms that the residuals are not normally distributed.

To address the non-normality of residuals, we decided to log-transform the dependent variable, NO2, as follows:

$$Log\ (NO2) \sim Location + Treatment + \ Location\ *\ Treatment$$

This model registered lower R-squared and adjusted R-squared values (0.505 and 0.491) than the original model, indicating that this model accounts for a smaller percentage of the variability of the (log-transformed) NO2 concentrations. Having an F-statistic of 36.41 and a p-value of 2.68e-16, the overall model is also statistically significant. The model's information criteria, AIC and BIC, are -23.81 and -12.97, respectively, indicating a strong fit.

Table 12 - Log-transformed OLS Regression Results

Coefficient	Estimate	Std. Error	t-value	p-value	[0.025,0.975]
Intercept	4.0816	0.031	131.030	0.000	[-4.020,4.143]
Location	-0.3751	0.044	-8.124	0.000	[-0.462,-0.288]
Treatment	0.2626	0.082	3.228	0.002	[0.102,0.426]
Treatment*Location	-0.1866	0.113	-1.655	0.101	[-0.410,0.037]

To assess the robustness of our model, we undertook multiple diagnostic tests:

Table 13 - Log-transformed OLS Regression Diagnostic Tests

Test	Statistic	p-value	Conclusion
Breusch-Pagan Test	8.156	0.043	Heteroscedasticity.
Durbin-Watson Statistic	1.357		Some positive autocorrelation.
Jarque-Bera Test	4.061	0.131	Residuals are normally distributed.

Table 14 - Log-transformed OLS Regression VIF Test

Coefficient	VIF
Location	1.191489
Treatment	2.125000
Treatment*Location	2.316489

When looking at the model's assumptions, we can observe that it’s safe to say that the log-transformed OLS regression didn’t show enough evidence to prove that the residuals don’t follow a normal distribution (Table 13). Regardless, when looking at the interaction terms, it is safe to claim that it performed worse than the non-logarithmized version, by failing to provide enough evidence to conclude that there is a statistically significant interaction effect between the location and the treatment period on log-transformed NO2 levels (Table 13). It also showed evidence of heteroscedasticity (Table 13). Therefore, with this analysis, and the fact that this model had a smaller R-squared than the original model, we conclude that this model turned out to perform worse than the original one. Given that both models break classical assumptions and don’t pass all tests, we will not implement either of the models, yet several conclusions can still be drawn from the Difference-In-Difference Analysis and from our attempts at model implementation.

4.3.2. Traffic Patterns and Characteristics

To apply the DiD approach to study of changes in traffic patterns and characteristics, we collected data from the Waze platform. This data was used to create two groups: the treatment group, consisting of traffic data from the Jardim da Parada area, and the control group, consisting of traffic data from the Jardim do Arco do Cego area.

To apply the difference-in-differences methodology, we have established comparisons between the differences in traffic data indicators for the treatment group (Jardim da Parada) and the control group (Jardim do Arco do Cego) before and during the superblock implementation period.

Table 15 - Traffic Queues’ Average Speed (km/h)

Location	Jardim da Parada	Jardim do Arco do Cego
Before Treatment (0)	7.797072	6.562360
During Treatment (1)	7.003575	7.127791
Difference-in-Differences		-1.35893

The results displayed in Table 15 show a decrease in the average speed of traffic queues in Jardim da Parada relative to the control location. This reduction aligns with one of the superblocks’ main objectives, which is to make streets a safer space for everyone. This decrease in average speed was likely caused by changes in traffic flows, which were intensified on roads that were not accessible to cars. This shift may have resulted in altered traffic dynamics, contributing to the observed reduction in speed.

Table 16 - Traffic Queues’ Average Length (meters)

Location	Jardim da Parada	Jardim do Arco do Cego
Before Treatment (0)	361.127092	297.664083
During Treatment (1)	324.456204	326.567844
Difference-in-Differences		-65.57465

In Table 16, we observe substantial changes in the average length of traffic queues. While in Jardim do Arco do Cego, there was a small increase of 9% in the average length of traffic queues, Jardim da Parada experienced a reduction of 10% in this measure.

Table 17 - Traffic Queues’ Average Delay (meters)

Location	Jardim da Parada	Jardim do Arco do Cego
Before Treatment (0)	91.183938	90.635473
During Treatment (1)	99.938537	93.147271
Difference-in-Differences		6.2428

The increase in the traffic delay registered in Jardim da Parada’s area (9.6%) was higher than the increase registered in the control area (2.8%). The increased delay could also be attributed to the changes in traffic patterns caused by the implementation of the superblocks, which may have led to adjustments in traffic flow and congestion levels.

Overall, the Difference-in-Differences analysis showed that the results had a greater impact on the average traffic queue length and average speed compared to the average traffic delay, where the results were more discrete. With the results being more evident in the case of the traffic queues' average length, we decided to apply an Ordinary Least Squares (OLS) regression to the traffic queues' average length data (described as follows):

$$Queues'\ Average\ Speed \sim Location + Treatment +\ Location\ *\ Treatment$$

Table 18 - OLS Regression Results (Predicting Queue’s Average Speed)

Coefficient	Estimate	Std. Error	t-value	p-value	[0.025,0.975]
Intercept	6.5624	0.298	22.027	0.000	[5.947,7.177]
Location	1.2347	0.421	2.930	0.007	[0.365,2.104]
Treatment	0.5654	0.421	1.342	-0.192	[-0.304,1.435]
Treatment*Location	-1.3589	0.596	-2.281	0.032	[-2.589,-0.129]

This model's F-statistic achieved a value of 2.94, with a p-value of 0.0535, indicating that the model is not statistically significant at the conventional 5% significance level. The model also achieved an R-squared of 0.269, indicating that it accounts for only 26.9% of the variability in the dependent variable. The adjusted R-squared is even smaller at 0.177, showing a reduction in explained variability when considering the number of predictors. While the "Location" and the interaction term were classified as significant, the "Treatment" variable was not, with a p-value of -0.192 (Table 18). These results suggest that the model may not provide a strong explanation of the variations in the traffic queues' average speed. As such, we decided to try a log-transformed version of this model:

$$Log(Queues'\ Average\ Speed) \sim Location + Treatment +\ Location * Treatment$$

Table 19 - OLS Regression Results (Predicting Log-Transformed Queue's Average Speed)

Coefficient	**Estimate**	**Std. Error**	**t-value**	**p-value**	**[0.025,0.975]**
Intercept	2.0129	0.041	49.038	0.000	[1.928,2.098]
Location	0.1601	0.058	2.757	0.011	[0.040,0.280]
Treatment	0.0766	0.058	1.320	0.199	[-0.043,0.196]
Treatment*Location	-0.1712	0.082	-2.085	0.048	[-0.341,-0.002]

The log-transformed model exhibited worse values in almost every parameter: the R-squared was only 0.243, the adjusted R-squared was 0.148, and the F-statistic was 2.562 with a p-value of 0.0785. Additionally, the "Treatment" variable was again not significant at the 5% level. Given these values, we decided to reject this model (Table 19).

As we failed to create this model, and we still believed it could be useful to create a model to predict one of the traffic data variables, we decided to try a new attempt and to predict the average queues' length. As such, we defined the model as:

$$Queues'\ Average\ Length \sim Location + Treatment +\ Location\ *\ Treatment$$

Table 20 - OLS Regression Results (Predicting Queue's Average Length)

Coefficient	**Estimate**	**Std. Error**	**t-value**	**p-value**	**[0.025,0.975]**
Intercept	297.6641	11.163	26.666	0.000	[274.626,320.703]
Location	0.1601	15.786	4.020	0.001	[30.882,96.044]
Treatment	0.0766	15.786	1.831	0.080	[-3.678,61.485]
Treatment*Location	-0.1712	22.325	-2.937	0.007	[-111.652,-19.498]

The developed model exhibited better performance than previously generated models (which aimed to predict the queues' average speed), achieving an R-squared of 0.404 and an adjusted R-squared of 0.330. The model was also considered statistically significant, by achieving a F-statistic of 5.43 with a p-value of 0.00537, being relevant at a 5% level. As in the previously generated models, the location and the interaction between treatment and location significantly affect the queues' length, but that does not happen with the treatment variable alone (Table 20). This means that, while treatment alone has a positive impact, it is not significant unless considering its interaction with the location.

Table 21 - OLS Regression Diagnostic Tests (Predicting Queue's Average Length)

Test	Statistic	p-value	Conclusion
Breusch-Pagan Test	4.365	0.043	No significant heteroscedasticity.
Durbin-Watson Statistic	1.556	NA	Some positive autocorrelation.
Jarque-Bera Test	6.218	0.0446	Residuals are not normally distributed.

Table 22 - Log-transformed OLS Regression VIF Test

Coefficient	VIF
Location	2.0
Treatment	2.0
Treatment*Location	3.0

To ensure the model's robustness, we performed several diagnostic tests (Tables 21 and 22). The Breusch-Pagan test for heteroscedasticity showed no evidence of heteroscedasticity and the Durbin-Watson statistic was 1.556, revealing some positive autocorrelation in the residuals, which, according to Field (2009), is not considered problematic. Additionally, we examined the Variance Inflation Factors (VIFs) to assess multicollinearity, and all VIF values were low, indicating that multicollinearity is not a concern in this model. The only test that was negative for the model was the Jarque-Bera test, which indicated that the residuals didn't follow a normal distribution. To address the non-normality of the residuals, we will proceed by log-transforming the dependent variable.

Table 23 - OLS Regression Results (Predicting Log-Transformed Queue's Average Length)

Coefficient	Estimate	Std. Error	t-value	p-value	[0.025,0.975]
Intercept	5.6890	0.038	151.543	0.000	[5.611,5.766]
Location	0.2004	0.033	3.774	0.001	[0.091,0.310]
Treatment	0.1009	0.053	1.901	0.069	[-0.009, 0.211]
Treatment*Location	-0.2064	0.075	-2.749	0.011	[-0.361-0.051]

The new model (Table 23) has shown similar results to those achieved with the non-logarithmized version, indicating overall statistical significance with an F-statistic of 0.00969 (< 0.05), but not demonstrating significance for the treatment variable, showing significance only for the location and interaction term. Compared to the previous model, it achieved a lower R-squared (0.373) and adjusted R-squared (0.294).

Table 24 - OLS Regression Diagnostic Tests (Predicting Queue's Average Length)

Test	Statistic	p-value	Conclusion
Breusch-Pagan Test	4.736	0.043	No significant heteroscedasticity.
Durbin-Watson Statistic	1.556	NA	Some positive autocorrelation.
Jarque-Bera Test	20.269	0.000	Residuals are not normally distributed.

Table 25 - Log-transformed OLS Regression VIF Test

Coefficient	VIF
Location	2.0
Treatment	2.0
Treatment*Location	3.0

To verify the model's reliability, we performed several diagnostic tests (Tables 24 and 25). The Breusch-Pagan test showed no heteroscedasticity, and the Durbin-Watson test indicated acceptable autocorrelation levels. All VIF values were low, suggesting no evidence of multicollinearity. Despite our efforts to address non-normal residuals, the Jarque-Bera test revealed that residuals are not normally distributed.

4.4. DISCUSSION

In this subsection, we summarize our models and their implications for the study. We identified both significant and non-significant models, providing insights into the effects of the superblock implementation.

4.4.1. Air Quality Model

Regarding NO2 concentrations, we developed two OLS regression models, and both were statistically significant, demonstrating that the superblock implementation succeeded in reducing NO2 concentrations in the studied area. The main difference between the models was that the second model used a logarithmic transformation for the predicted NO2 values, while the first model did not.

The first model (non-logarithmized version) achieved better performance, and we adopted it as our best model. This model's predictive variables were all statistically significant, indicating that the location, the treatment, and the interaction term between these variables were significant for the changes in NO2 concentrations, underscoring the success of the superblock introduction in reducing NO2 levels.

The model's variables coefficients indicated that while there was an increase in NO2 levels during the treatment period (with a positive estimate of 17.2621), the location itself showed significantly lower NO2 levels (estimate of -18.0073), and the superblock's implementation contributed to further reductions in NO2 concentrations (interaction term was -14.6882).

4.4.2. Traffic Model

Four different OLS regression models were created to analyse the traffic data. The first two models were based on the average speed of the vehicles but were not statistically significant. Although these two models had two significant variables (location and interaction term), from which we could still draw some conclusions, we decided to discard them due to their low level of overall significance. As it was not possible to create a model predicting the average speed of vehicles, we created a similar model to predict the average length of queues.

The two models developed to predict average queue length showed better results than the model developed to predict average queue speed and were both statistically significant. The first model used the original measurements of queue length and showed a R-squared value of 0.404, which means that it explains around 40.4 percent of the variation in average queue length. The second model, which transforms the average queue length data logarithmically, has a slightly lower value (0.373).

Both models revealed a significant effect of location and the interaction between location and treatment, but no significant effect of treatment alone. This suggests that although the development of superstreets has an impact on traffic behaviour, their success depends very much on specific local condition.

The non-logarithmic model performed best in predicting average queue length. The coefficients in this model show that the introduction of a superblock (treatment) is associated with an increase in average queue length of 28.9038 metres, although this increase is not statistically significant. The coefficients on the location variable show that queue lengths at superblock locations were initially 63.46 metres longer than the mean, reflecting the traffic conditions at superblock location prior to the superblock implementation. The interaction between the intervention and the specific location resulted in a decrease in the average queue length by 65.5746 metres.

5. CONCLUSIONS, LIMITATIONS AND FUTURE WORKS

5.1. Conclusions

Our research provides an assessment of the first experimental implementation of a superblock in Portugal. Through the integration of the CRISP-DM and DiD techniques, we evaluated the effects of this experimental establishment on traffic congestion, noise levels, and air quality.

The exploratory investigation of the air quality indicators during the superblock implementation yielded mixed results. These results revealed that particulate matter concentration levels were lower during this period than during the preceding analysis period, suggesting possible benefits from the experience. In contrast, our analysis of the CO and NO2 concentrations levels revealed that these pollutants concentrations were present at higher amounts than before the superblock trial. Given the short period analysis, it's possible that these findings were influenced by external factors (e.g., rural fires), which could have biased the results. Therefore, we decided to use a different kind of analysis: the Difference-in-Differences approach. This approach is more suitable for such periods as it compares two different locations during the same period, under the same external conditions.

The traffic exploratory analysis revealed that the traffic suffered deep changes that were naturally expected due to the closure of some roads, and understanding these changes is important to make a fair assessment of the superblock introduction. Speed traffic was 9% lower during the implementation period, which aligns with one of the superblocks objectives, making streets a safer place for everyone. Some other changes were the occurrence of more frequent traffic queues, which had a bigger delay than in the period before, even being smaller (in length) than usual. This is understandable if we think that drivers and pedestrians had to adapt to the new characteristics of the roads during the experience period and that the closure of the superblock inner roads led more drivers to the outside streets, causing more frequent and intense traffic.

The superblock implementation led to noticeable changes in both traffic flows and air quality in the treatment region. The Difference-in-Differences analysis showed an improvement in air quality indicators, particularly regarding NO2 concentrations, which were confirmed as statistically significant. While particulate matter (PM2.5 and PM10) concentrations showed improvement both visually and numerically, this improvement lacked statistical confirmation.

The DiD analysis of traffic flows revealed a small reduction in average traffic speed, consistent with the superblock concept's purpose of making streets safer for everyone, although this reduction was not statistically significant. It was also noticeable that traffic flows experienced an increase in delay times, indicating that further optimization of traffic management within the superblock is needed to mitigate these delays and enhance overall efficiency.

In summary, our study highlights the advantages of the implementation of superblocks, such as creating more green spaces, increasing residents' life quality, and improving air quality, but also underscores the importance of planning and consideration of traffic patterns to address any consequences. These findings may be particularly valuable and useful for city planners and policymakers to support the adoption of superblocks as a viable strategy for sustainable urban transformation.

5.2. Limitations and Future Work

While our research has provided significant insights, some areas require further investigation to address the remaining challenges based on our findings.

One of the main limitations was that the air quality and noise sensors were distant from the superblock's implementation area. Our approach included using the nearest sensors to the area, but even those were slightly distant from Jardim da Parada, with their distances ranging from 350 meters to 1.1 kilometres.

Other significant limitations were the data, as potential errors could affect the study's results, and analysing a single location. Introducing these urban solutions in different areas could help in understanding the impact in diverse urban contexts. Extending these implementations to another Portuguese city like Porto would provide insights about the superblocks' impact in different regions, that are subject to different environmental and climacteric conditions.

While we were able to extract and provide valuable insights from the Campo de Ourique's experience, it would be important to repeat this experience during a more extended period to make a deeper analysis, as the nine-day duration was restrictive and a strong limitation for our research.

Superblocks have a strong impact in people's lives, especially on those who live within or near them. Conducting an assessment to understand how people feel about the implementation would be particularly important, considering that one of the superblocks purposes is to improve the quality of residents' lives. Evaluating the economic impact caused by these implementations would also be valuable before turning them into a definitive solution. Changing traffic flows can lead to a decrease or an increase in some businesses' economic activity, and it is important to understand how it may affect the local economy.

BIBLIOGRAPHICAL REFERENCES

Amati, M., Stevens, Q., & Rueda, S. (2023). Taking Play Seriously in Urban Design: The Evolution of Barcelona's Superblocks. *Environment and Planning C: Politics and Space, 41*(1), 1–18. doi:10.1177/12063312231159229

de Almeida, J.; Alves, G. B. M., Nunes, R. d. O., & Dias, T. (2022). Despite Being Distinguished as the 2020 European Green Capital, Lisbon Has Lost Public Green Areas over the Previous Decade. *Sustainability, 14*(19), 12112. https://doi.org/10.3390/su141912112

Cornell, S. E., & Gupta, A. (2020). Is climate change the most important challenge of our times? In M. Hulme (Ed.), Contemporary climate change debates (pp. 6-20). Routledge. https://doi.org/10.4324/9780429446252-2

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. The CRISP-DM consortium. Retrieved from https://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/14.2/es/CRISP-DM.pdf

Deisting, K. (2022). Grätzltransformation: Implementing the Superblock 2.0 (Master's thesis). Lund University, Faculty of Engineering, Department of Architecture and Built Environment. Retrieved from http://lup.lub.lu.se/student-papers/record/9101830

European Environment Agency. (2023). Transport and Mobility. Retrieved from https://www.eea.europa.eu/en/topics/in-depth/transport-and-mobility

Eggimann, S., Lischer, P., & Bolliger, J. (2021). Evaluating superblock design to enhance urban greening. *Journal of Physics: Conference Series, 2042*(1), 012005. doi:10.1088/1742-6596/2042/1/012005

Eggimann, S. (2022). Expanding urban green space with superblocks. *Land Use Policy, 117*, 106111. doi:10.1016/j.landusepol.2022.106111

Field, A. (2009). *Discovering statistics using SPSS (and sex and drugs and rock 'n' roll)* (3rd ed.). SAGE Publications.

Furquim, F., Corral, D., & Hillman, N. (2020). A primer for interpreting and designing difference-in-differences studies in higher education research. In L. Perna (Ed.), *Higher education: Handbook of theory and research* (Vol. 35, pp. 245-274). Springer, Cham. https://doi.org/10.1007/978-3-030-31365-4_5

Filastò, B. N. (2022). Are Superblocks the super solution to social problems? Assessing social sustainability of the Superblock pilot project in Vienna (Master of Science thesis). Department of Environmental Sciences and Policy, Central European University, Vienna.

Florin, R., & Olariu, S. (2015). A survey of vehicular communications for traffic signal optimization. *Vehicular Communications.* https://doi.org/10.1016/j.vehcom.2015.03.002

Fredriksson, A., & Magalhães de Oliveira, G. (2019). Impact evaluation using Difference-in-Differences. RAUSP Management Journal, 54(4), 519-532. https://doi.org/10.1108/RAUSP-05-2019-0112

Hahad, O., Kuntic, M., Al-Kindi, S., Kuntic, I., Gilan, D., Petrowski, K., Daiber, A. & Münzel, T. (2024). Noise and mental health: evidence, mechanisms, and consequences. Journal of Exposure Science & Environmental Epidemiology. https://doi.org/10.1038/s41370-024-00642-5

Instituto Nacional de Estatística. (2021). Censos 2021 resultados definitivos. Retrieved from https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes

Lee, B.-J., Kim, B., & Lee, K. (2014). Air pollution exposure and cardiovascular disease. *Toxicological Research, 30*(2), 71-75. https://doi.org/10.5487/TR.2014.30.2.071

Li, K., & Wilson, J. P. (2023). Modeling the health benefits of Superblocks across the City of Los Angeles. *Applied Sciences, 13*(4), 2095. doi:10.3390/app13042095

Li, K., Gao, S., Liao, Y., Luo, K., & Wang, S. (2022). The impact of development zones on China's urbanization from the perspectives of the population, land, and the economy. *Land, 11*(10), 1726. https://doi.org/10.3390/land11101726

Liao, L., Du, M., & Chen, Z. (2023). Environmental pollution and socioeconomic health inequality: Evidence from China. *Sustainable Cities and Society, 95,* 104579. https://doi.org/10.1016/J.SCS.2023.104579

Lin, J. C., Mitchell, L., Crosman, E., Mendoza, D. L., Buchert, M., Bares, R., Fasoli, B., Bowling, D. R., Pataki, D., Catharine, D., Strong, C., Gurney, K. R., Patarasuk, R., Baasandorj, M., Jacques, A., Hoch, S., Horel, J., & Ehleringer, J. (2018). CO_2 and carbon emissions from cities: Linkages to air quality, socioeconomic activity, and stakeholders in the Salt Lake City urban area. *Bulletin of the American Meteorological Society, 99*(11), 2325-2339. https://doi.org/10.1175/BAMS-D-17-0037.1

López, I., Ortega, J., & Pardo, M. (2020). Mobility infrastructures in cities and climate change: An analysis through the superblocks in Barcelona. *Atmosphere, 11*(4), 410. doi:10.3390/atmos11040410

Luongo, S., Napolano, E., & Gul, K. (2023). Sustainable tourism intentions: Extending the theory of planned behavior. *Turistica - Italian Journal of Tourism, 32*(3), 31-60. https://doi.org/10.1234/turistica.v32i3.28

Mueller, N., Rojas-Rueda, D., Khreis, H., Cirach, M., Andrés, D., Ballester, J., Bartoll, X., Daher, C., Deluca, A., Echave, C., Milà, C., Márquez, S., Palou, J., Pérez, K., Tonne, C., Stevenson, M., Rueda, S., & Nieuwenhuijsen, M. (2020). Changing the urban design of cities for health: The superblock model. *Environmental International, 134,* 105132. doi:10.1016/j.envint.2019.105132

Müller, J., Straub, M., Stubenschrott, M., & Graser, A. (2023). Simulation of a full-scale implementation of Superblocks in Vienna. In 15th *ITS European Congress, Lisbon, Portugal, 22-24 May 2023.*

Navarro-Serer, B. (2021). Healthier streets: How superblocks and congestion pricing can help reduce air pollution in Washington, DC. *Journal of Science Policy & Governance, 18*(1). https://doi.org/10.38126/JSPG180110

Nieuwenhuijsen, M., de Nazelle, A., Cirach Pradas, M., Daher, C., Dzhambov, A. M., Echave, C., Gössling, S., Iungman, T., Khreis, H., Kirby, N., Khomenko, S., Leth, U., Lorenz, F., Matković, V., Müller, J., Palència, L., Barboza, E. P., Pérez, K., Tonne, C., & Mueller, N. (2024). The Superblock model: A review of an innovative urban model for sustainability, liveability, health and well-being. *Environmental Research, 251*, 118550.

Oltra, C., Sala, R., López, S., & Germán, S. (2022). An unexpected polarization: Individual-level factors behind the public acceptance of the superblocks of Barcelona. https://doi.org/10.31235/osf.io/9dwgk

Pereira, M., Neves, F. T., & Neto, M. D. C. (2022). The application of randomised controlled trials for urban policy evaluation in smart cities: a systematic literature review. In *Livro de Resumos da Conferência do Projeto de Investigação Científica "Fatores de Transformação Urbana (DRIVIT-UP)" em conjunto com I Conferência sobre Ciência de Dados para Ciências Sociais e VI Conferência de Planeamento Regional e Urbano.* [Abstract book from the Conference of the Scientific Research Project "Drivers of urban transformation (DRIVIT-UP)" a joitly event with I Conference on Data Science for the Social Sciences And VI Conference on Regional and Urban Planning] (pp. 38-40). UA Editora. https://doi.org/10.48528/pkzd-wz70

Rodrigues, N. J. (2023, September 8). Lisboa testa superquarteirão em Campo de Ourique. *Smart Cities*. https://smart-cities.pt/noticias/lisboa-testa-superquarteirao-em-campo-de-ourique-08-09/

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature

Santibáñez-Andrade, M., Sánchez-Pérez, Y., Chirino, Y. I., Morales-Bárcenas, R., Quintana-Belmares, R., & García-Cuellar, C. M. (2022). Particulate matter (PM10) destabilizes mitotic spindle through downregulation of SETD2 in A549 lung cancer cells. *Chemosphere, 133900*. https://doi.org/10.1016/j.chemosphere.2022.133900.

Sarroeira R, Henriques J, Sousa AM, Ferreira da Silva C, Nunes N, Moro S, Botelho MDC. Monitoring Sensors for Urban Air Quality: The Case of the Municipality of Lisbon. *Sensors (Basel), 23*(18) doi: 10.3390/s23187702. PMID: 37765759; PMCID: PMC10537901.

Seidman, M. D., & Standring, R. T. (2010). Noise and Quality of Life. *International Journal of Environmental Research and Public Health, 7*(10), 3730-3738. doi:10.3390/ijerph7103730

Wawa, E. A., & Mulaku, G. C. (2015). Noise pollution mapping using GIS in Nairobi, Kenya. *Journal of Geographic Information System, 7*(5), 486-493. https://doi.org/10.4236/jgis.2015.75039

Wing, C., Simon, K., & Bello-Gomez, R. A. (2018). Designing difference in difference studies: Best practices for public health policy research. *Annual Review of Public Health, 39*, 453-469. https://doi.org/10.1146/annurev-publhealth-040617-013507

Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. In Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining (pp. 29-39). Manchester, UK. Retrieved from https://cs.unibo.it/~danilo.montesi/CBD/Beatriz/10.1.1.198.5133.pdf

Zhou, H., Taber, C., Arcona, S., & Li, Y. (2016). Difference-in-differences method in comparative effectiveness research: Utility with unbalanced groups. *Health Services and Outcomes Research Methodology, 16*(3-4), 155-174. https://doi.org/10.1007/s10742-016-0158-2

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa