NOVA IMS Information Management School

MGI

Mestrado em Gestão de Informação
Master Program in Information Management

A Walkability Simulator using a BI approach: a Lisbon Case Study

Liu Fangxiao

Project Work presented as partial requirement for obtaining the Master's degree in Information Management

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

2020

Title: A Walkability Simulator using a BI approach: a Lisbon Case Study

Liu Fangxiao

MGI

NOVA Information Management School

Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa

A WALKABILITY SIMULATOR USING A BI APPROACH: A LISBON CASE STUDY

by

Liu Fangxiao

Project work presented as a partial requirement for obtaining the master's degree in Information Management, with a specialization in Knowledge Management and Business Intelligence

Co Advisor: Miguel de Castro Neto

Co Advisor: Miguel André Oliveira

May 2022

ACKNOWLEDGEMENTS

In the first place, I would like to express my gratitude to my supervisor, Professor Miguel de Castro Neto, and Professor Miguel André Oliveira, who guided me to complete this project. Always available and patient to clarify my doubts with decisive answers. I would also like to acknowledge the help provided by Jônatas Augusto Manzolli. Thank you for sharing the valuable information and all the advice given during the project development.

Additionally, this endeavor would not have been possible without the generous support from the Climate Driven Technologies for Low Carbon Cities project (C-TECH), which is co-financed by the European Regional Development Fund (ERDF) through the Operational Program for Competitiveness and Internationalisation (COMPETE 2020), the North Portugal Regional Operational Program (NORTE 2020), and the Portuguese Foundation for Science and Technology (FCT under MIT Portugal). Thanks for giving me this golden opportunity to develop this project.

Appreciate my family for all the support and goods provided, which allowed me to reach this point. A special thanks to my parents, Liu Tongzhuo and Sun Yan. I cannot make myself who I am today without your unconditional and selfless love. Thank you for supporting me, motivating me, and believing in me.

I wish to extend my special thanks to my direct superior, Gonçalo Alexandre Gomes Cavaco, a decisive person in my life who led me to the world of Business Intelligence, helping me to start my first internship and the full-time job of my life. At the time, I had no idea what Business Intelligence was. With his knowledge and leadership in it, I started to love this area, so I enrolled in this master's degree to expand my knowledge about it. I would also appreciate Ana Filipa Matoso Ramos Rosario and Pedro Miguel Loureiro Meias, my colleagues with whom I had worked and contributed to this fantastic experience. Furthermore, I wish to express my appreciation to Henrique Miguel Rodrigues Carvalho, an excellent friend who has always been by my side, supporting and advising me during the most critical period of my life.

ABSTRACT

Urban pedestrian environment quality improvement is one of the city planners' main focuses in achieving sustainable city planning. Further, physical activities (PA) such as walking may provide advantages to health, social and environmental aspects. Considering the factors aforementioned, this project explores the idea of walkability. This concept measures if an urban environment is favorable to pedestrians. Based on a solid framework for the walkability evaluation previously developed using a multi-criteria decision analysis approach (MCDA), this project proposes implementing a business intelligence (BI) solution using Power BI. According to the selected evaluation process and the 5Cs layout, which defends that a walkable environment has to be Connected, Convenient, Comfortable, Convivial, and Conspicuous, the nineteen most suitable criteria encompassing route condition, route safety, and route characteristics were defined.

A conceptual model was created first, consisting of three main steps: overall requirements identification, data model architecture proposal, and dashboard proposal. An intensive data collection will be performed from multiple data sources, and several transformations will be done to prepare the data collected for further uses. In this project, three parishes are selected, namely Marvila, Beato, and Parque das Nações.

Finally, a data visualization platform and a walkability simulator will be developed. The main idea is to give a clear view of each environment's walkability level, improve the decision-making process, and provide accurate results in a fast and efficient manner. Analyze the cost-efficiency and effectiveness of an action that intends to improve the walkability of a particular street/ Area and help the city planners to build a more pedestrian-friendly environment.

KEYWORDS

Walkability; city planning; Business intelligence; data visualization; micro-mobility; Simulation

INDEX

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

DMs	Decision-Makers
BI	Business Intelligence
UNDP	United Nations Development Programme
PA	Physical Activity
BI&A	Business Intelligence and Analytics
MCDA	Multi-Criteria Decision Analysis
C-TECH	Climate Driven Technologies for Low Carbon Cities
DSR	Design Science Research
DS	Design Science
IS	Information Systems
DSRM	Design Science Research Methodology
MWM	Measurement Assessment Framework
NEWS-Y	Neighborhood Environment Walkability Scale-Youth
NEWS	Neighborhood Environment Walkability Scale
AURIN	Australian Urban Research Infrastructure Network
IAAPE	Indicators of Accessibility and Attractiveness of Pedestrian Environments
FAHP	Fuzzy Analytic Hierarchy Process
PROMETHEE	Preference Ranking Organization Method for Enrichment of Evaluations
ERDF	European Regional Development Fund
NORTE 2020	North Portugal Regional Operational Program
FCT	Foundation for Science and Technology
MIT	Massachusetts Institute of Technology
KPI	Key Performance Indicator
SSDT	SQL Server Data Tools
IoT	Internet Of Things
IDV	Interactive Data Visualization

GLD	Greater London Authority
ROI	Return of Investments
GIS	Geographic Information System
Esri	Environmental Systems Research Institute
PGIL	Plataforma de Gestão Inteligente de Lisboa

1. INTRODUCTION

Walking is one of the most critical aspects that must be considered when decision-makers (DMs) deal with sustainable planning. This attribute is closely linked to improving communities' quality of life, such as health, social, economic, and environmental benefits (Manzolli, Oliveira, & Neto, 2021). Cities that encourage individuals to walk can improve public health, reduce transportation costs, and promote a better environment ecosystem.

Some of the biggest challenges faced by city planners and DMs are basically how the term walkability can be measured, visualized, and used. This project aims to present a possible solution using Business Intelligence (BI) to solve their needs. After proposing the conceptual model, its implementation will be conducted concerning three selected parishes in Lisbon, namely Marvila, Beato, and Parque das Nações.

1.1. BACKGROUND

According to the United Nations Development Programme (2022) (UNDP), Cities account for 60 to 80 percent of energy consumption and at least 70 percent of carbon emissions, with only 3% of land use in the entire world. From another hand, more than two-thirds of all humanity will be urban. 6.5 billion people will live in a city by 2050. Making cities sustainable is now imminent. It means "creating career and business opportunities, safe and affordable housing, and building resilient societies and economies. It involves investment in public transport, creating green spaces and improving urban planning and management in participatory and inclusive ways". The increase of the urban population across the globe links the logistics of transporting people and goods with smart and sustainable city planning, contributing to how society will face mobility in the future. Just like Paiva, Ahad, Tripathi, Feroz, & Casalino (2021) mentioned, new optimization algorithms for transport are necessary as the global population increases exponentially.

Moreover, due to the COVID-19 pandemic, mobility choices have been profoundly affected and disturbed(Campisi et al., 2020; Manzolli et al., 2021). Long-distance travel, especially on an international scale, has been banned or restricted to restrict the virus's spread. Mobility has also been controlled to contain the virus on a short-distance scale. Furthermore, Lisbon is one of the most impacted cities in Portugal. Thus, individual mobility is vital to avoid the spread of the disease at a local level. As a result, many city dwellers have chosen more active and non-motorized transportation modes as the principal method to move (Manzolli et al., 2021).

Many studies demonstrate the importance of walkability to sustainable city planning and social development. It improves the quality of life, reduces transportation costs, reduces negative environmental impacts, and offers more equity of the access to urban activities (Ruiz-Padillo, Pasqual, Larranaga Uriarte, & Cybis, 2018). Moreover, intensive face-to-face contact may foster knowledge spread, social capital, a sense of community, and a closer relationship with neighbors (Dovey & Pafka, 2020). Just like Fancello, Congiu, & Tsoukiàs (2020) mentioned, walkability is a crucial component for improving urban life quality, allowing citizens to use and explore the urban environment concerning their needs and values. Furthermore, some articles illustrate the relationship between walking and positive impacts on the economy. For example, improving walkability can encourage customers to purchase more local goods and promote economic resilience (Gilderbloom, Riggs, & Meares, 2015).

Nowadays, diseases such as hypertension, stroke, diabetes, obesity, osteoporosis, depression, and some types of cancer are associated with inadequate physical activity (PA). A high proportion of the population lacks regular PA. However, there are many ways to be physically active, where walking plays a vital role. It tends to be particularly important for elderly, disabled, and lower-income people who have few opportunities to participate in sports or formal exercise programs. Thus, neighborhoods that facilitate people walking are crucial to improving public health (Litman, 2003a).

Emerging from the previously mentioned aspects, this project focuses on creating a BI solution to measure walkability. During the last two decades, BI and related areas have grown significantly (Chen, Chiang, & Storey, 2012), both in the number of products and services offered and the adoption of these technologies (Chaudhuri, Dayal, & Narasayya, 2011). This considerable growth has been fueled by the low cost of software acquisition and the capability of storing substantial amounts of data from multiple sources (Chaudhuri et al., 2011). It proposes a set of techniques and methodologies aimed at enabling DMs to make better and faster decisions at the right time. Just like the Gartner (2022) mentioned in their information technology glossary about analytics and business intelligence (BI&A), it “is an umbrella term that includes the applications, infrastructure and tools, and best practices that enable access to and analysis of information, to improve and optimize decisions and performance.”

1.2. PROBLEM IDENTIFICATION

In this project, the evaluation process created by Manzolli et al. (2021) was selected. It was created based on a Multi-Criteria Decision Analysis (MCDA) approach with the objective to improve the methodological background of walkability assessment, providing a solid MCDA framework that can be replicated and used as a support tool for the design of urban public policies worldwide.

However, the process created requires much effort, from data collection to walkability calculation. It is time-consuming, manual, and not easy to interpret. Since the main goal of this process is to be used as a support tool for the design of urban public policies, how the data can be managed and presented to the end-users is what matters. Furthermore, it is hard to prioritize the attributes selected. In the article, all the criteria are equal-weighted. There is no possibility of having a different scenario depending on the users' interests. Moreover, it is not easy to figure out the weakness of a certain street and evaluate the effectiveness and efficiency when an action intends to improve walkability occurs.

1.3. MAIN GOAL AND SPECIFIC OBJECTIVES

The main goal of this project is to create a walkability simulator using a BI approach. By implementing BI solutions, dashboards are created that allow users, especially city planners, to visualize and analyze the walkability level of each alternative on the map. Moreover, enabling the simulation capability is also one of the objectives and requirements. It allows users, especially city planners and DMs, to figure out the weakness of a certain street and evaluate the effectiveness and efficiency of an action that intends to improve walkability. Firstly, a conceptual model and project planning were created. Then, a case study in Lisbon, Portugal, will be conducted, and three parishes will be selected.

Starting from these considerations, the objectives of this work are twofold.

1 – Conceptual model and project planning design

- Business needs identification
- Data model proposal
- Dashboard proposal

2 – Conceptual model application to the Lisbon case study

- Study area selection
- Software selection
- Data preparation
- Data model creation and application
- Dashboard development

3- Results evaluation and discussion

1.4. Study Relevance and Importance

Firstly, this project was developed as a contribution to the Climate Driven Technologies for Low Carbon Cities (C-TECH). This research aims to identify the possibility of simulating the walkability, with an objective to help end-users make decisions, provide valuable results, and ensure an action that intends to improve the walkability of a certain street or area cost-efficient and effective. From other perspectives, create a data visualization platform that allows end-users to view each street/area's walkability level, recognize its weaknesses, and figure out the best criteria to improve it. Even though each city is different, and the way to improve walkability may differ from case to case, this project intends to propose a dimensional model that can be replicated and used in other cities. The dashboard requires some knowledge about the concepts presented and some know-how about the tool. However, it will be available for all kinds of users.

1.5. Methodology

The Design Science Research (DSR) Methodology, a methodology developed for Design Science (DS) for Information Systems (IS), was selected during the project development. "Design Science Research Methodology (DSRM) incorporates principles, practices, and procedures required to carry out such research and meets three objectives: it is consistent with prior literature, it provides a nominal process model for doing DS research, and it provides a mental model for presenting and evaluating DS research in IS" (Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007).

DSR pretends to develop valid and reliable knowledge for designing solutions and solving problems (van Aken, 2004), intending to create new knowledge and reality (Iivari & Venable, 2009) where it fits perfectly into our project. As mentioned previously, the project pretends to develop a BI solution trying to answer the needs of city planners and DMs once they are dealing with sustainable city planning. Innovative artifacts and new knowledge will be created during the development. The DSRM is what is needed and should be applied.

According to Peffers et al. (2007), the DSRM process generally includes six core activities (Figure 1.1).

1. Problem identification and motivation
2. Definition of objectives for a solution
3. Design and development of artifacts
4. Demonstration by using the artifact to solve the problem
5. Solution evaluation
6. Communication

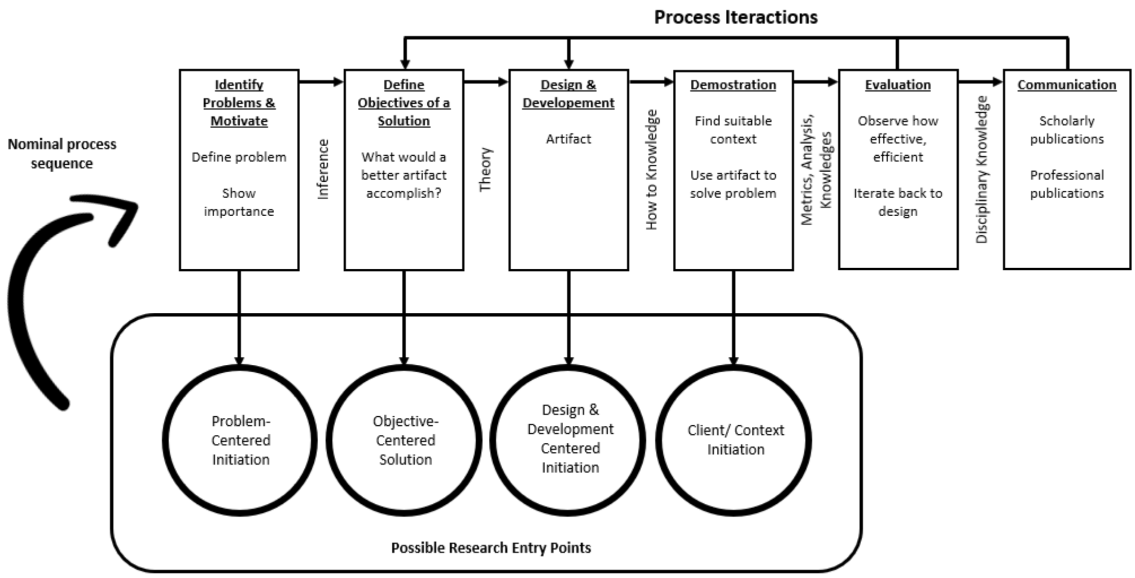

Figure 1.1 – DSRM Process Model. Adopted from (Peffers et al., 2007)

2. LITERATURE REVIEW

This section presents the literature review developed as a basis for the study. First, a study of walkability and its benefits is conducted. Then, pertinent work concerning the evaluation of walkability is presented. In the end, relevant topics regarding BI will be addressed. Furthermore, A review of best practices to implement a BI to satisfy the business's needs is managed.

2.1. WALKABILITY

Enormous work and research were conducted to understand and institutionalize transport space design for motorized vehicular modes over the last decades. Pedestrian transportation, however, is a much more recent addition to planning processes and is still addressed with far less intensity, seriousness, and funding (Lo, 2009). The benefits of having high-quality walkable streets and public spaces are underestimated due to more extensive problems that need to be addressed first, and seemingly more severe solutions need to be applied (Porębska, Rizzi, Otsuki, & Shirotsuki, 2019). Moreover, walking is the cheapest and the universal mode of transportation (Papageorgiou, Maimaris, Efstathiadou, & Balamou, 2018). Thus, it tends to be undervalued in conventional transport planning. The term walkability has only started to be identified as an essential component of efficient, accessible, equitable, sustainable, and livable communities in the post-modernist planning era (Lo, 2009).

The term walkability refers to a sustainable place for walking (Rebecchi et al., 2019). It measures whether an urban environment is favorable to pedestrians and has been wildly used in several studies to describe the quality of walking conditions, including safety, comfort, and convenience (Manzolli et al., 2021). Considering the positive correlation between the built environment factors and walking behavior (Moura, Cambra, & Gonçalves, 2017), developing a sustainable urban environment that encourages citizens to walk is unquestionable (Fancello et al., 2020). City planners may address the pedestrian environment's quality by measuring it, supporting more objective, effective and comprehensive walking-related strategies, and interventions (Moura et al., 2017).

Several studies pointed out that the term walkability influences public health. Thus, sustainable city design and public health have always been drawn together (Ewing & Handy, 2009). In fact, the built environment is an essential contributor to the persistence of inequalities in public health. Thus, It is pertinent to consider which aspects of the built environment encourage or discourage pedestrians from walking and which are most closely associated with relevant health outcomes. A study conducted by Glazier et al. (2014) tried to determine the relationship between built environment characteristics and active transportation behaviors, overweight, obesity, and diabetes. They concluded that individuals living in more walkable areas were more than twice as likely to use active transportation such as walking and cycling than those living in less walkable places. Furthermore, PA is the fourth leading cause of death worldwide. Thus, it is crucial to mark the promotion of PA as a primary target for the public health policy for disease prevention and health promotion (Rebecchi et al., 2019).

Besides the advantage of walkability to public health, having a walkable neighborhood also benefits social, environmental, and economic aspects. In the first place, active transportation modes, such as walking and cycling, is the healthiest, cheapest, and most universal mode of transportation (Papageorgiou et al., 2018). By encouraging citizens to choose active modes of transportation, the use of cars or other types of vehicles can be minimized, reducing CO2 emissions and pollution. Just like Zook, Lu, Glanz, & Zimring (2012) mentioned, walking represents an environmentally friendly and most sustainable transportation mode from an environmental point of view.

Furthermore, residents tend to walk through their neighborhood more frequently with highly walkable areas. This fosters good relationships with neighbors as residents and provides more opportunities for social interactions (Götz, Yoshino, & Oshio, 2020). Just like Abass & Tucker (2018) mentioned, it "promotes informal interactions, social capital and sense of community." In addition, a walkable environment facilitates and promotes knowledge spread, network routines, and face-to-face communications (Adler, Florida, King, & Mellander, 2019; Götz et al., 2020). Thus, it often has unique creative, and innovative sociocultural vibes (Florida, Mellander, & Stolarick, 2010; Götz et al., 2020). Moreover, a growing body of work demonstrates the inherent economic value of walkable neighborhoods by encouraging economic transactions and social exchange (Gilderbloom et al., 2015; Leinberger & Alfonzo, 2012; Litman, 2003b).

2.2. Walkability Measurement

The concept of walkability does not have a straightforward evaluation process defined yet due to its complexity and subjectivity. Depending on the objectives, the evaluation process may change from case to case and the technology used to support it. Just like Manzolli et al. (2021) mentioned, even for experts, defining an appropriate way to evaluate walkability is a challenging task. Moreover, many attributes can impact the citizens' choice of transportation modes, such as social condition, population age, health condition, and people categories (Manzolli et al., 2021). The attribute selection varies depending on the objectives of the study, and both quantitative and qualitative data can be used (Boongaling, Luna, & Samantela, 2021; Talen, 2002). Also, it is hard to prioritize the attributes to use (Zook et al., 2012).

However, to measure it, it is essential to consider how pedestrians are defined and the discourses that are shaping pedestrian space development (Lo, 2009). Then, it is important to identify the key attributes that can influence walkability in the first place. In a study carried out by the London Planning Advisory Committee, the multidimensional 5C's layout, has been widely adopted for this purpose. According to this schema, a walkable, pedestrian-friendly environment has to be Connected, Convenient, Comfortable, Convivial, and Conspicuous (Moura et al., 2017).

Connected: the extent to which the pedestrian network connects to the main origins and destinations of the journey, as well as the extent of the connection between the different routes of the network (Manzolli et al., 2021; Moura et al., 2017).

Convenient: Analyze how convenient is a street to pedestrians. The extent to which walking can compete with other means of transportation in terms of efficiency, such as time, money, and space (Manzolli et al., 2021; Moura et al., 2017).

Comfortable: Analyze how a street is designed and built to provide a comfortable walking environment for all types of pedestrians (Manzolli et al., 2021; Moura et al., 2017).

Convivial: Streets provide a better conviviality in which walking is a pleasant activity, in terms of interaction with people, buildings, and natural environments, and encompassing social and recreation activities (Manzolli et al., 2021; Moura et al., 2017).

Conspicuous: Streets that provide clear and legible signage and information in terms of social legibility, complexity, and coherence, resulting in walking routes and public spaces that are discernible and encourage pedestrians to walk (Manzolli et al., 2021; Moura et al., 2017).

A study done by Rebecchi et al. (2019) in Milan pointed out the importance of PA, such as walking and cycling, to an individual's well-being. An evidence-based assessment framework was built to measure walkability and evaluate cyclist-pedestrian accessibility, The Milan Walkable Measurement Assessment Framework (MWM). An innovative assessment framework that investigates both scales of the city, Macro and Micro, addresses the urban features to promote a healthier behavior of the population. The macro analysis is quantitatively related to urban planning and is divided into three criteria: density, diversity, and design. The micro analysis is qualitatively related to urban and street design and is divided into four criteria: usefulness, safeness, comfort, and aesthetics. Furthermore, Rosenberg et al. (2009) examined the psychometric properties of the Neighborhood Environment Walkability Scale-Youth (NEWS-Y) and explored its associations with context-specific and overall PA among youth. The Neighborhood Environment Walkability Scale (NEWS) was developed to determine how the environmental factors hypothesized to influence PA and demonstrated good reliability among adults. Based on that, they created a new version of NEWS relevant to youth age and reached a similar result. To complete it, parents of children ages 5–11, adolescents ages 12–18, and adolescents ages 12–18 from Boston, Cincinnati, and San Diego, USA, were studied.

Giles-Corti et al. (2014) demonstrated the possibility of creating an automated walkability index based on open-source software architecture within the Australian Urban Research Infrastructure Network (AURIN) framework. Nevertheless, it has not been possible to create a national walkability measure due to the lack of consistent national data measuring land use mix. The walkability index was created based on three environmental characteristics: street connectivity, residential or dwelling density, and land use mix. The main objectives of the research are: (i) To develop, trial, and validate an automated open-source tool capable of creating walkability indices at user-specified scales. (ii) To create a flexible tool that would enable walkability to be measured using existing data available within the AURIN portal or enable users to upload their own detailed data. (iii) To create a tool that would allow researchers to upload geocoded addresses from survey data to create user-specified service areas around these addresses and to download associated geospatial data for further interrogation.

In a different approach, a participatory assessment method was created by Moura et al. (2017) to measure walkability. The framework considers travel purposes (e.g., utilitarian, leisure) and different pedestrian groups (adults, children, seniors, and pedestrians with reduced mobility). A seven key dimensions approach (7Cs layout) was developed and implemented to express the walkability. Based on the conceptual framework, a methodology of the walkability evaluation process through GIS-based and street auditing indicators is presented. The process is examined in a central area in Lisbon, Portugal, and the result demonstrates a clear difference in walkability scores in different pedestrian groups. In addition, a tool called IAAPE (Indicators of Accessibility and Attractiveness of Pedestrian Environments) was developed based on local circumstances and expertise (Moura et al., 2017).

Ruiz-Padillo et al. (2018) present a study that aims to determine a weighted walkability index constructed based on the relative importance of their attributes. To do so, a robust multi-criteria method called Fuzzy Analytic Hierarchy Process (FAHP) was implemented to reach the goal. Moreover, a comparison between the FAHP method and other more straightforward methods was established, and a chi-square test for homogeneity was conducted. Then, a case study was managed in Porto Alegre, Brazil, to assess the effect of changes on attributes in walkability. The result indicates that the three most crucial walkability attributes are: Public Security, Traffic Safety, and Pavement Quality.

2.3. Business Intelligence

Just like Chen et al. (2012) stated, BI&A has emerged as an important area of study in both the academic and the business communities during the last decades. Computer-based decision support continues to evolve from its early structured reporting roots. It turns from “nice to have” to a “competitive necessity” and has become widely used (Watson, 2009). One of the primary objectives is to close the gap between current performance and the desired one. The term BI was introduced by Howard Drenser of the Gartner Group to describe a set of concepts and methods to improve business decision-making using fact-based, computerized support systems in 1989 (Watson, 2009). According to Gartner (2022), it “is an umbrella term that includes the applications, infrastructure and tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance.” It is a combination of applications, technologies, and processes for gathering, storing, accessing, and analyzing data to help business users while making decisions (Watson, 2009).

Moreover, large data volumes are generated on a daily basis from heterogeneous sources at an unprecedented rate (e.g., health, government, social networks, marketing, financial) due to many technological trends, including the Internet Of Things (IoT), the proliferation of cloud computing, and the spread of smart devices. (Oussous, Benjelloun, Ait Lahcen, & Belfkih, 2018). Just like Vercellis (2009) stated, "The adoption of low-cost massive data storage technologies and the wide availability of Internet connections have made available large amounts of data that have been collected and accumulated by various organizations over the years." Therefore, how the data can be managed and used to obtain helpful insights is a complex challenge in today's business. The enterprises capable of transforming data into information and knowledge can make quicker and more effective decisions and achieve a competitive advantage. Likewise, in public administration, the analysis of available information enables the development of better and innovative services for citizens. These ambitious goals are technical, however complex, and cannot be achieved alone without the support of competent minds and advanced analytical methods (Vercellis, 2009). There is where BI takes place. It can take advantage of opportunities presented by the abundant data and domain-specific analytics required in many critical and high-impact application areas (Chen et al., 2012). In this case, It can be defined as a set of mathematical models and analytical methods that systematically utilize available data to retrieve information and knowledge capable of supporting complex decision-making processes (Vercellis, 2009).

2.3.1. Data Preparation

Transforming the data into usable information and knowledge to obtain reliable insights is a fundamental step. As its name indicates, data preparation is a process that prepares the data for future uses. Just Like Craig Stedman, Ed Burns, & Mary K. Pratt (2022) stated, it is a process of gathering, combining, structuring, and organizing data to be used in BI, analytics, and data visualization platforms. The primary purpose of data preparation is to ensure the raw data being readied for processing and analysis is accurate and consistent so the results of BI and analytics applications will be valid. It also involves finding relevant data to guarantee that BI applications deliver meaningful information and actionable insights to help DMs while making decisions. However, low data quality may lead to misunderstanding the data and making wrong decisions.

According to David Stodder (2016); Mariana Bento (2021), it is necessary to understand and assimilate the analysis objectives to proceed with data preparation. With well-defined objectives, data can be treated appropriately and prepared for further steps, such as reporting. Firstly, the process integrates the ingestion of data sources, and then the tools and methodologies are applied to transform the data according to the objectives defined previously. Moreover, it is also essential to integrate with governance and administration to monitor and improve the data used.

A well-defined data preparation process has several benefits to an organization: (1) Able to ensure the data quality and result reliability produced by analytic tools/ BI applications. (2) Capable of identifying and fixing data issues that otherwise might not be detected. (3) DMs have more reliable information while making decisions. (4) Reduce the data management and analytics costs. (5) Avoid duplication of effort in preparing data for use in multiple applications. (6) Increases the Return on Investment(ROI) from BI and analytics initiatives (Craig Stedman et al., 2022).

According to Mariana Bento (2021); Patnaik & Jain (2017), data preparation has four fundamental steps:

- Data cleaning: A process that cleans the data. It detects and corrects the corrupted or inaccurate records from data.
- Data integration: A process that aims to combine multiple data sources, such as databases, data cubes, or flat files, into a single unified view.
- Data transformation: A process that applies necessary changes to ensure that the data is in a format suitable for the type of analysis desired.
- Data reduction: A process that allows for obtaining a smaller sample of the data without compromising the quality of the analysis results and improves classification performance by eliminating redundant or irrelevant records.

2.3.2. Information Visualization and Dashboard Design

According to Ware (2004), the meaning of the term visualization is more likely a graphical representation of data or concept. How the data can be represented and visualized to facilitate the decision-making process is essential. There is where information visualization takes place. Information visualization uses data and represents them in a meaningful, visual way that users can easily interpret and comprehend, turning data into information and knowledge. It provides an effective way to share insights in a digestible format for non-experts and allows decision-makers to draw conclusions more efficiently and act in an informed way (TIBCO Software, 2022). After all, We observe more information through vision than through all of the other senses combined (Ware, 2004).

Moreover, visualization brings a series of advantages: (1) It provides an ability to comprehend a massive volume of data. (2) It allows the perception of emergent proprieties that were not anticipated. (3) It enables problems with the data to become immediately apparent. (4) It also facilitates understanding both large-scale and small-scale features of the data and (5) hypothesis formation. Furthermore, the process of data visualization includes four basic stages: (1) The collection and storage of the data. (2) A data preprocessing stage aims to transform the data prepared to further manipulation. (3) Mapping the data to the visual representation. (4) The human perceptual and cognitive system (Ware, 2004). Figure 2.1 demonstrates the process described.

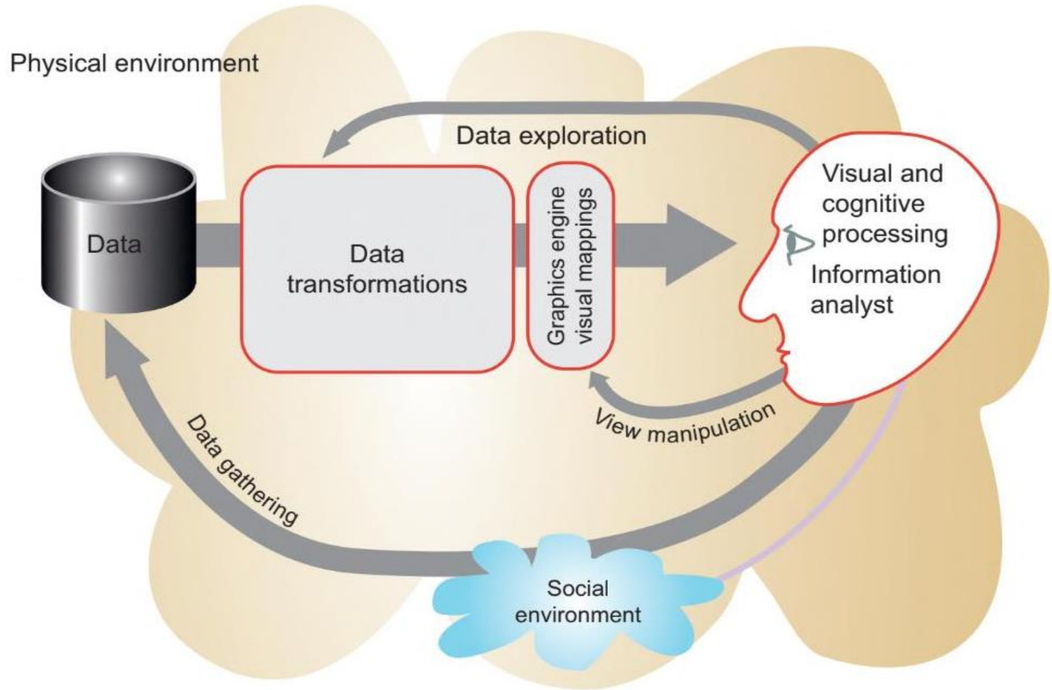

Figure 2.1 – The visualization process (Ware, 2004)

Interactive Data Visualization (IDV), also known as dashboards, refers to BI technology and the technique of gathering, processing, and analysis of large and complex amounts of data for a specific application. It simplifies the process of organizing and visually displaying data in an easy-to-use interface and increases the level of understanding of data-driven performance measurement (Janvrin, Raschke, & Dilla, 2014). Thus, Dashboards play an important role in organizations. For Taras Bakusevych (2018), it “is an at-a-glance preview of the most crucial information for the user at the moment he is looking at it, and an easy way to navigate directly to various areas of the application that require users attention.” Figure 2.2 demonstrates how the IDV works.

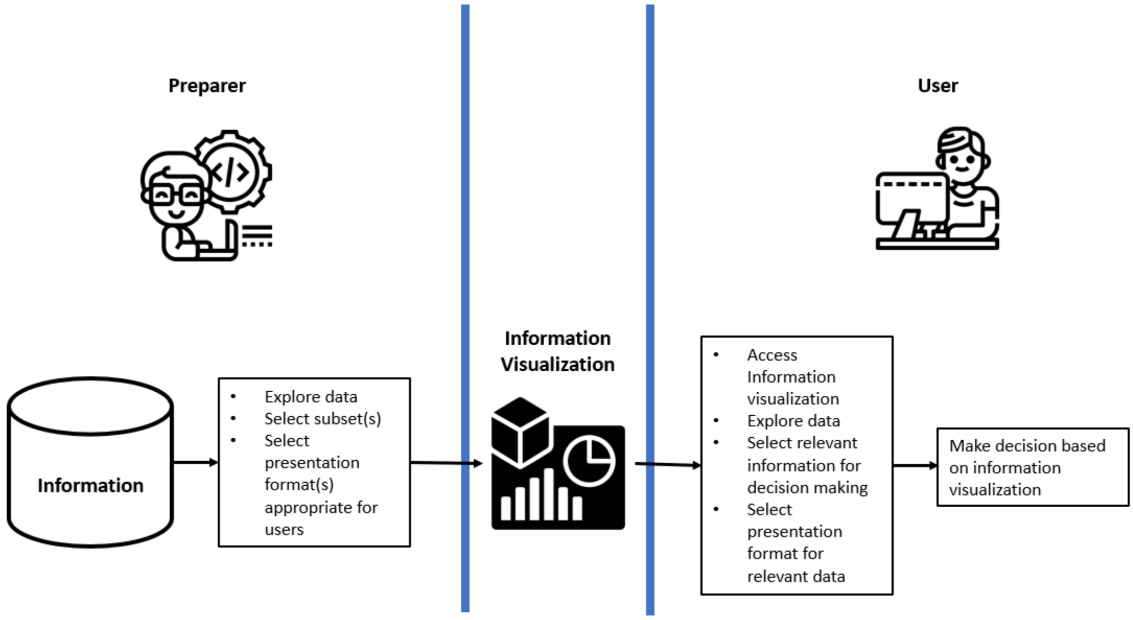

Figure 2.2 – Overview of IDV process. Adapted from (Janvrin et al., 2014)

Taras Bakusevych (2018) provided a practical guide for a better dashboard design. First, define what the purpose of the dashboard is. Depending on the business need, a specific goal needs to be achieved and served. After that, the proper representation of data needs to be selected. Data representation is a complex task, especially when multiple types of information are needed. Choosing the wrong representation can lead to confusion and misunderstanding of the data. Figure 2.3 demonstrates a guideline to help it out.

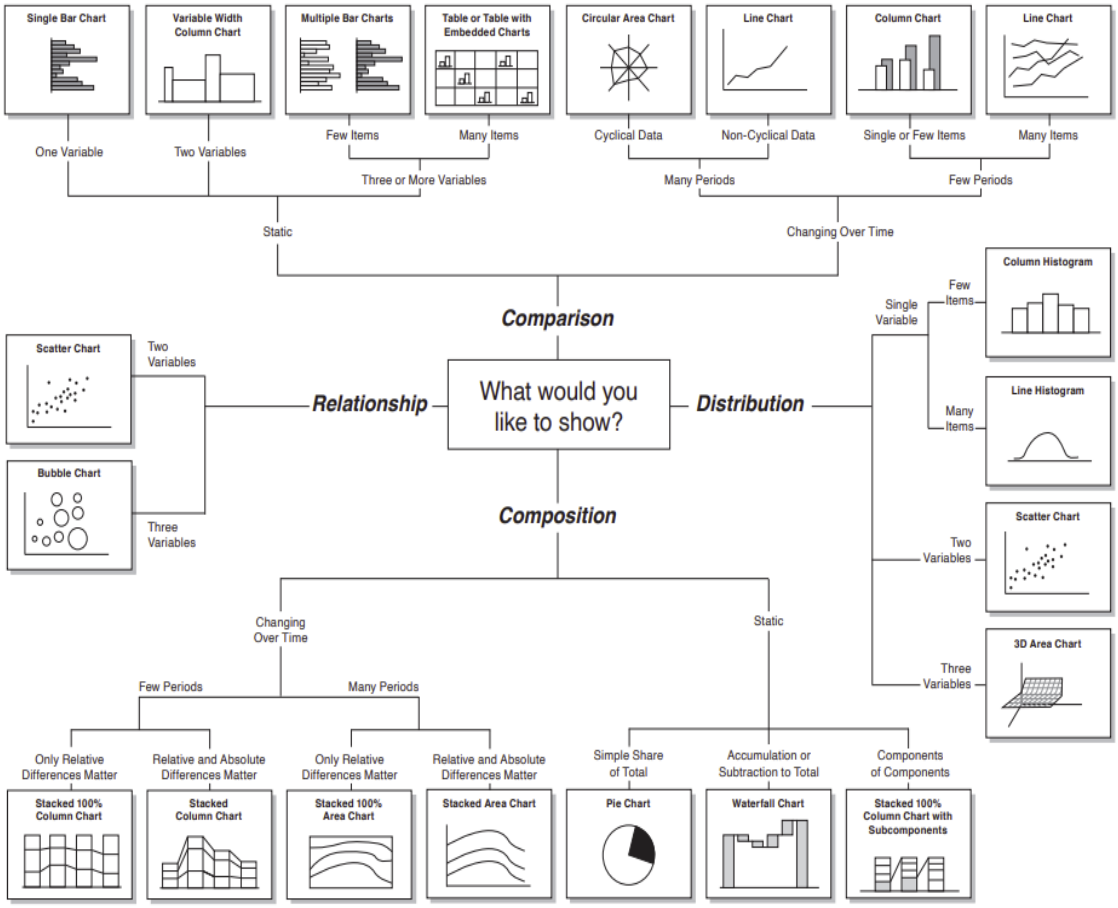

Figure 2.3 – Chart Chooser (Abela, 2006)

Furthermore, clear and consistent naming conventions and consistent date formatting should be followed. Truncate large values are also necessary to deliver precise information. Defining the layout flow and prioritizing the represented information is also important. Depending on the habit of people while reading, the top left corner normally gains more attention from users. Thus, the key information should lift from left to right accordingly (Taras Bakusevych, 2018). Figure 2.4 demonstrates how the users usually scan through dashboard content.

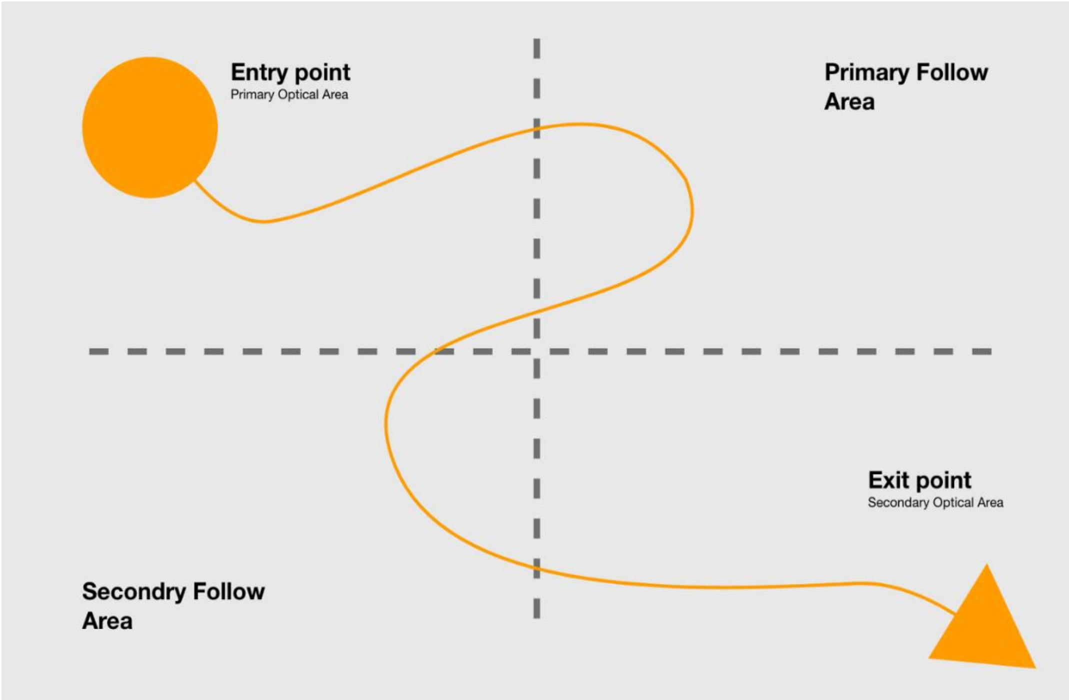

Figure 2.4 – Reading gravity (Soham Sinha, 2021)

3. CONCEPTUAL MODEL PROPOSAL

This section will design a conceptual model proposal to answer the business needs. This step consists of the “Design and Development” step proposed in the DSR methodology. First, a study of the selected walkability evaluation process will be conducted. Then, the overall requirements will be defined. Finally, a proposal for the data model architecture and dashboards conceptual model is made.

3.1. WALKABILITY EVALUATION

As mentioned previously, an evaluation process created by Manzolli et al. (2021) was adopted to proceed with the walkability simulator. This project developed a walkability score by applying an MCDA, which is wildly used to measure and classify routes, paths, and streets in terms of walkability levels (Manzolli et al., 2021; Pelegrina, Duarte, & Romano, 2019). It deals with problems characterized as a choice among alternatives to solve multiple criteria problems. Then, eighteen criteria are selected based on the multidimensional 5Cs Layout developed by the Greater London Authority (GLD), encompassing route conditions, route safety, and route characteristics. In addition, a new criterion concerning street lighting was added to the evaluation process. The well-lit street offers good visibility over a long distance at night. Improving lighting is a direct means of cost-effectively creating a sense of public safety, improving the quality of the built environment, and increasing the number of people on the streets after dark (Painter, 1996). Thus, it is considered an important criterion for route safety. However, all the criteria assume the same weight. Figure 3.1 presents the nineteen selected criteria and hierarchically structured fundamental objectives.

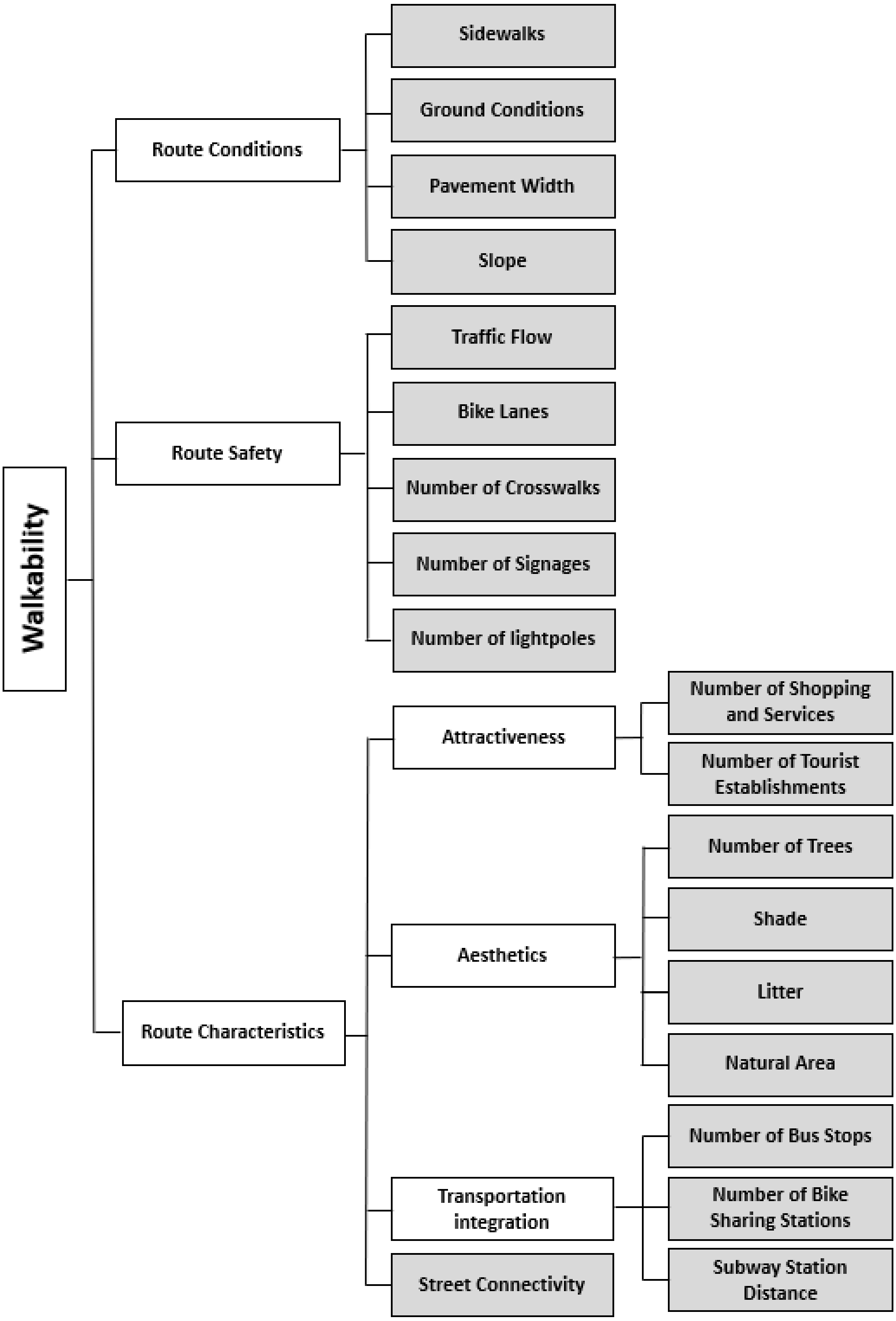

Figure 3.1 – Tree of fundamental objectives. Adapted from (Manzolli et al., 2021)

Furthermore, the Preference Ranking Organization Method for Enrichment of Evaluations (PROMETHEE) method was selected as a support tool to implement the MCDA developed, mainly because it is an outranking model that allows the use of quantitative and qualitative scales. For this purpose, the Visual PROMETHEE software was used, which enables the ranking capability (Manzolli et al., 2021). Then, a case study was conducted in Lisbon, Portugal. These authors selected ten routes located in different zones of the Parque das Nações. During this phase, data were collected from the most varied sources, such as official sources from the Lisbon municipality, companies, and open source applications. Moreover, QGIS and ArcGIS were used to access and manipulate most of the data collected.

3.2. Overall Requirements

Manzolli et al. (2021) suggested two objectives for improving their research: (1) To apply different weights to the selected criteria to create different analysis scenarios. (2) a more significant number of alternatives can be analyzed in the future, perhaps carrying out a complete mapping of the "Parque das Nações" area in walkability rankings.

A discussion was carried out based on the author's needs, and several requirements were defined.

- The data model must be capable of processing and storing a more significant number of alternatives.
- The tool must be capable of ranking alternatives.
- Simulation capability must be delivered.
- The tool should enable the users to change the weight of the criteria.
- Preset scenarios should be established according to the users' interests.
- The process should be automatic.

3.3. Data Model Architecture Proposal

In this section, a data model architecture will be established. The dimensional modeling method was selected, and the four-step dimensional design process proposed by Kimball & Ross (2011) was adopted.

The first step is to select a business process. Considering the requirements defined in the previous section, the focus should be on the walkability simulator and its visualization. After that, the grain of the business should be declared. Using other words, define the granularity of the dimensional model. However, it is pretty straightforward. All the attributes in the fact table will be at the segment level. Thus, the grain of the proposed model would be one row per segment.

After defining the business process and its grain, the next step is to define the dimensions and the fact tables. The model contains two dimension tables and only one fact table.

Location dimension (DIM_Location): This dimension presents all the selected alternatives on the segment level. A location hierarchy is created in this table, starting with the city until it reaches the segment level (City -> District -> Municipality -> Parish -> Street -> Segment).

Time Dimension (DIM_Date): This dimension presents the date when the attributes are generated and stored in the data model. It enables the possibility of storing and analyzing the walkability score along the time, as well as the value of each criterion. In the end, a comparison over time can be established. A date hierarchy is created in this table, starting with the year number until it reaches the month level (Year -> Quarter -> Month).

Walkability fact (Fact_Walkability_Features): This table is a typical fact table that only consists of keys and attributes. As discussed previously, each record represents a segment, and it is only connected with the location dimension via segment ID.

An auxiliary table will be defined to support the tool to achieve the goals defined. Auxiliary tables do not have any connection with other tables and are not considered a dimension.

Scenario auxiliary table (Scenario): As mentioned previously, one of the authors' suggestions is to have the possibility to apply different weights to the criteria. Thus, a table consisting of the weight of each criterion by preset scenario was defined. It allows users to select the preset scenario and calculate the walkability score based on its weight.

Figure 3.2 demonstrates the conceptual model designed.

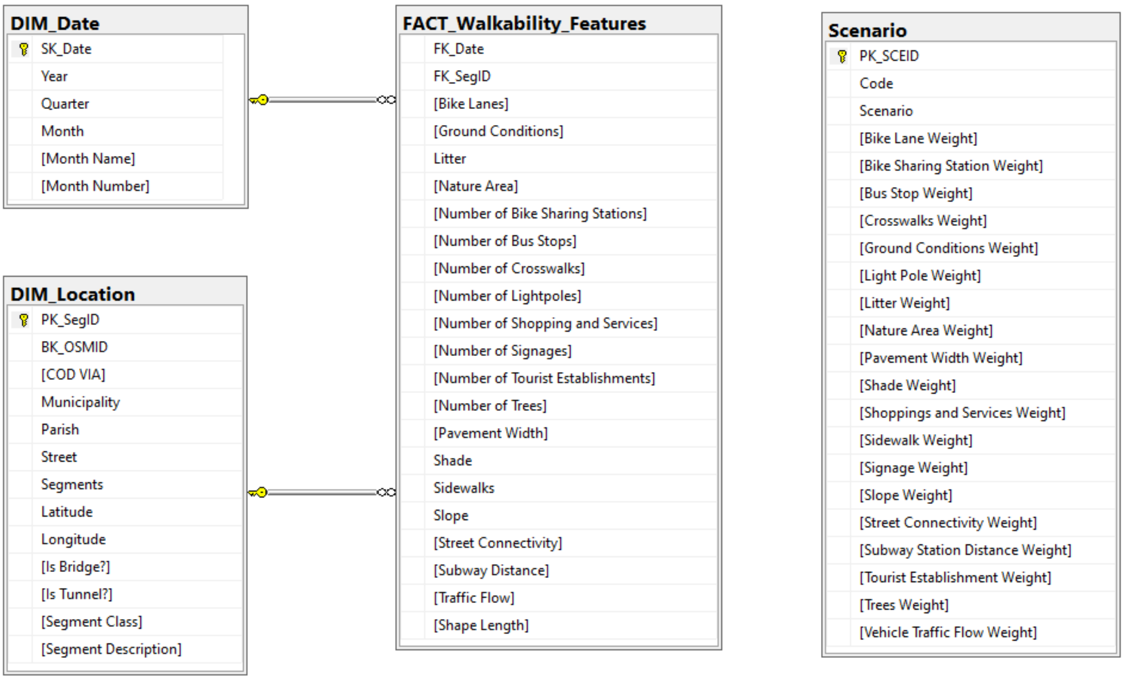

Figure 3.2 – Conceptual data model diagram

3.4. Dashboard Conceptual Model Proposal

This section will establish a conceptual model to proceed with the final dashboard creation. Considering the business needs previously defined, three pages will be created: (1) Walkability Overview, (2) Walkability Weight Simulator, and (3) Walkability Calculator. The Walkability Overview provides a clear view of the walkability score without any simulation applied. Then, in the Walkability Weight Simulator, weight simulation will be enabled, allowing users to calculate the score using the weight defined by themselves. In the end, the Walkability Calculator provides the possibility of simulating random scenarios. It calculates the score of a particular segment based on the attribute values and their weight defined by users.

The dashboard models established were designed based on the best practices studied in the literature review section and presented below (Figures 3.3 and 3.4).

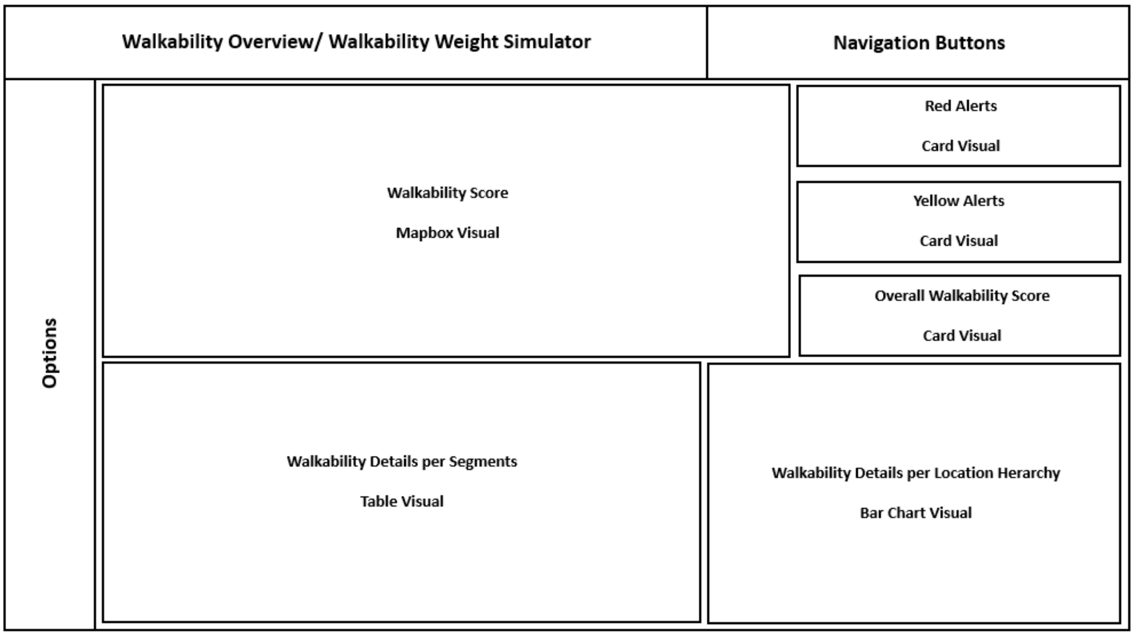

Figure 3.3 – Walkability Overview and Walkability Weight Simulator dashboard conceptual model

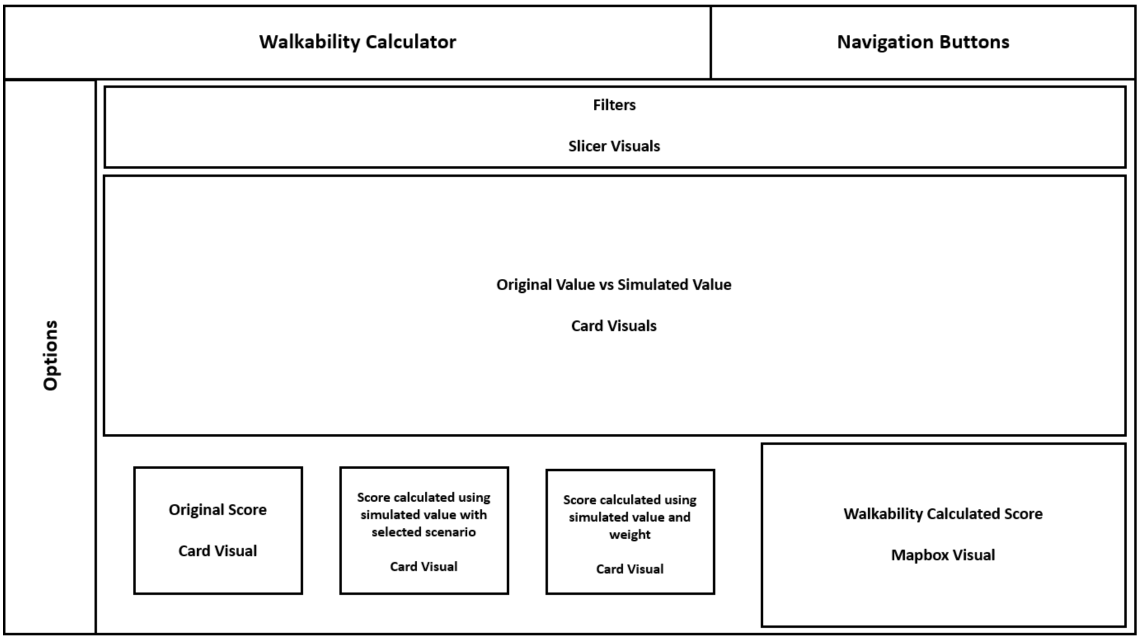

Figure 3.4 – Walkability Calculator dashboard conceptual model

4. DEVELOPMENT

This section demonstrates the dashboard creation process, which consists of the "Demonstration" step of the DSR methodology. Firstly, a study that intends to identify the most appropriate tool to proceed with the proposed BI system was conducted. Then, as mentioned in previous sections, a study area was defined. Next, the data collection and preparation will be done. In the end, the dashboards proposed will be established.

4.1. STUDY AREA

As previously mentioned, this work is supported by the C-TECH project, which is co-financed by the European Regional Development Fund (ERDF) through the Operational Program for Competitiveness and Internationalisation (COMPETE 2020), the North Portugal Regional Operational Program (NORTE 2020), and the Portuguese Foundation for Science and Technology (FCT under Massachusetts Institute of Technology Portugal [MIT]). Based on a three-dimensional representation of the city and its combination with multiple data from different data sources, for instance, domains such as weather, energy and water consumption, mobility, and, most of all, user's behavior determined by their mobile phone use, the C-TECH project aims at researching, developing, and pilot-scale a digital smart city platform for urban modeling and planning. It will allow simulating scenarios regarding energy-efficiency of buildings, green structures, creation, and urban mobility, empowering local authorities to identify and tackle specific environmental issues, overcoming the global challenge of decreasing the urban carbon footprint, and fostering the transition to a net-zero ecosystem (C-Tech, 2020).

Based on that, a study area from Lisbon city was selected to proceed with the evaluation. The study area includes three parishes in the eastern area of Lisbon along the Tagus River, namely Marvila, Beato, and Parque das Nações. The Parque das Nações parish has a more recent urban fabric than the other parishes, as it was renovated to receive the Universal Exhibition of Lisbon in 1998 (Expo'98). This pilot area has approximately 14 km2 and is reasonably diverse, e.g., incorporating residential, commercial, business areas, and cultural and leisure areas along the Tagus River. In order to avoid the effect of abrupt border changes, in other words, "artificial borders," a 100 meters buffer was applied. Figure 4.1 demonstrates the selected area.

Figure 4.1 – Study Area with 100m buffer (Lisbon City)

4.2. Software Selection

Selecting the proper software is important for the success of every BI project. It helps developers to achieve the objectives faster and more efficiently. Besides the skillset of BI developers, it also depends on the availability of the software to complete the project.

In this project, Power BI from Microsoft was adopted. According to Gartner´s research, Power BI is recognized as a leader in BI&A platforms (Microsoft Power BI, 2022c) and demonstrated in Figure 4.2. It is a powerful business intelligence tool that easily connects to, models, visualizes the data, and creates personalized reports with KPIs and brands (Microsoft Power BI, 2022a). “It is capable of turning the unrelated source of data into coherent, visually immersive, and interactive insights” (Microsoft Power BI, 2022b).

Figure 4.2 – Magic Quadrant for Analytics and BI Platforms (Gartner, 2022)

4.3. Data Collection and Preparation

After selecting the study area and the software to use, the next step is to prepare the data to create the dashboards. Firstly, an extensive search was conducted to obtain the most reliable, accurate, and up-to-date data. Raw data was collected from various sources and stored in excel files in the first instance. QGIS and AGIS software were used to visualize, analyze and validate the collected data. In the end, the collected data was loaded into Power BI to be prepared for dashboard development.

Based on the four fundamental steps for the data preparation process studied in the literature review section, the data loaded into Power BI was manipulated using Power Query Editor. There were several non-validated records in the fact table that could contain incorrect data. Thus, a data cleaning process was conducted to remove those records. Then, excel files containing information regarding the selected criterion were merged to obtain a fact table with all the measurements available. After the data integration, the next step is to transform the data into the proper format. In the end, some irrelevant columns and content were eliminated to improve the dashboard's performance.

Moreover, a Max-Min normalization was applied to the selected criteria to ensure these are comparable and facilitate the calculation of the walkability score.

Selected criteria are described below and also in Table 4.1.

- Sidewalks: This binary qualitative criterion describes each segment's presence or absence of pedestrian sidewalks.
- Natural Areas: This binary qualitative criterion evaluates the presence of natural areas (mainly parks) in the selected segment.
- Ground Conditions: This quantitative criterion demonstrates the quality of the project, construction, and pavement maintenance of each segment. The number of occurrences on the Minha Rua LX app concerning pavement conditions and maintenance in a timeframe of two years was used to measure it.
- Pavement Width: This criterion is quantitative. It provides the width of each segment in meters and evaluates how effective pavement width is for walking, considering the presence of obstacles.
- Slope: This quantitative criterion evaluates the inclination of each segment in percentage. QGIS 3. 14 software was used as a support-aid tool to access the data and calculate the slope on the selected segment. It is a free and open-source desktop Geographic Information System (GIS) application that supports viewing, editing, and analyzing geospatial data (QGIS, 2022).
- Traffic Flow: This quantitative criterion evaluates the pedestrians' sense of safety pertaining to the presence of high vehicle traffic flow. Data concerning the level of traffic jams during rush hours (8-11 a.m. and 6-9 p.m.) were collected. The data provider classifies the traffic severity from one to five, where one is the lowest and five is the highest. Then, levels concerning high traffic (four and five) are counted to obtain the desired result.
- Number of Signages: This quantitative criterion evaluates the number of signage per segment.

- Street Connectivity: This quantitative criterion describes how the footpath network connects to other potentially available segments. It contains the number of street intersections for each segment.
- Shade: This quantitative criterion evaluates the percentage of the segment covered in the shade throughout the day. A 3D model of the area studied was developed. The days with the most sunlight were selected to obtain the percentage of shade during the day. After setting the time period, the simulation was conducted to provide the shade incident during the day and its location. The total length of segments covered falling into shade was calculated through this approach. ArcGIS Pro software was used as a support tool to access and acquire the data. ArcGIS is an application maintained by the Environmental Systems Research Institute (Esri), which works with maps and geographic information (Manzolli et al., 2021).
- Subway Station Distance: This quantitative criterion evaluates the nearest metro station of the selected segment. ArcGIS software was adopted to obtain the centroid of each segment, link subway stations nearby, and pick up the shortest walking distance of each reference. Then, the subway station with the minimum walking distance of each segment was selected to evaluate this criterion.
- Bike Lanes: This binary qualitative criterion evaluates the existence of bike lanes on the segment level. QGIS 3.14 software was used as a support aid to access the data.
- Number of Crosswalks: This quantitative criterion evaluates the number of crosswalks per segment. QGIS 3.14 software was used as a support aid to access the data.
- Number of Shopping and Services: This quantitative criterion evaluates the number of shops and services along the segment (e.g., bakeries, markets, pharmacies, restaurants, bars).
- Number of Tourist Establishments: This quantitative criterion evaluates the number of tourist establishments per segment (e.g., museums, hotels, sightseeing spots, points of interest).
- Number of Trees: This quantitative criterion evaluates the presence of trees on the segment level. QGIS 3.14 software was used as a support aid tool to access and manipulate the data.
- Litter: This quantitative criterion evaluates the presence of litter per segment. The number of occurrences on the Minha Rua LX app concerning urban cleaning in a timeframe of two years was considered to determine it.

- Number of Bus Stops: This quantitative criterion evaluates the presence of bus stops on the segment level.
- Number of Bike Sharing Stations: This quantitative criterion evaluates the presence of bike-sharing stations along the segment. QGIS 3.14 software was used as a support aid to access the data.

Ref.	Variable	Type	Aggregation	UoM	Impact	Source
v1	Sidewalks	Binary	Average	Yes/No	Positive	OSM
v2	Natural Areas	Binary	Maximum	Yes/No	Positive	Lisboa Aberta
v3	Ground Conditions	Quantitative	Sum	Number	Negative	Minha Rua LX
v4	Pavement Width	Quantitative	Average	Meter	Positive	OSM
v5	Slope	Quantitative	Maximum	Percentage	Negative	Lisboa Aberta
v6	Traffic Flow	Quantitative	Sum	Number	Negative	Waze Google Street Maps
v7	Number of Signages	Quantitative	Sum	Number	Positive	OSM Lisboa Aberta
v8	Street Connectivity	Quantitative	Sum	Number	Positive	OSM
v9	Shade	Quantitative	Maximum	Percentage	Positive	DTM aster2 or DTM/DGT
v10	Subway Station Distance	Quantitative	Minimum	Km	Negative	OSM Lisboa Aberta
v11	Bike lanes	Binary	Average	Yes/No	Positive	OSM Lisboa Aberta
v12	Number of Crosswalks	Quantitative	Sum	Number	Positive	OSM
v13	Number of Shopping and Services	Quantitative	Sum	Number	Positive	Google Cloud Turismo Portugal Lisboa Aberta
v14	Number of Tourist Establishments	Quantitative	Sum	Number	Positive	Google Cloud Turismo Portugal Lisboa Aberta
v15	Number of Trees	Quantitative	Sum	Number	Positive	Lisboa Aberta
v16	Litter	Quantitative	Sum	Number	Negative	Minha Rua LX
v17	Number of Bus Stops	Quantitative	Sum	Number	Positive	PGIL Data Lake
v18	Number of Bike Sharing Stations	Quantitative	Sum	Number	Positive	Lisboa Aberta
v19	Number of Lightpoles	Quantitative	Sum	Number	Positive	Lisboa Aberta

Table 4.1 – Selected criteria

According to the data collected, three criteria were defined differently from the selected evaluation process: ground conditions, street connectivity, and traffic flow. This project was capable of collecting quantitative data for those criteria instead of qualitative as previously defined and demonstrating more detailed information and, consequently, obtaining a more precise walkability score.

4.4. DASHBOARD

With objectives defined and the data prepared, the final stage is dashboard development. As mentioned in the "Dashboard Conceptual Model Proposal" section, three reports will be created during this phase, namely Walkability Overview, Walkability Weight Simulator, and Walkability Calculator.

4.4.1. Walkability Overview

This report provides a simple view of the walkability score based on the data prepared and illustrated in Figure 4.3. Its focuses on the actual state of walkability level of each segment in the study.

First, a Mapbox visual was created that provided location information for each segment and its respective walkability score. Then, based on the target defined, the card visual was adopted to illustrate the general walkability score and the distribution of the segments. Afterward, a table was created to demonstrate detailed information and the ranking of each segment. In the end, a bar chart that shows the walkability score per location hierarchy and its comparison to the target was created.

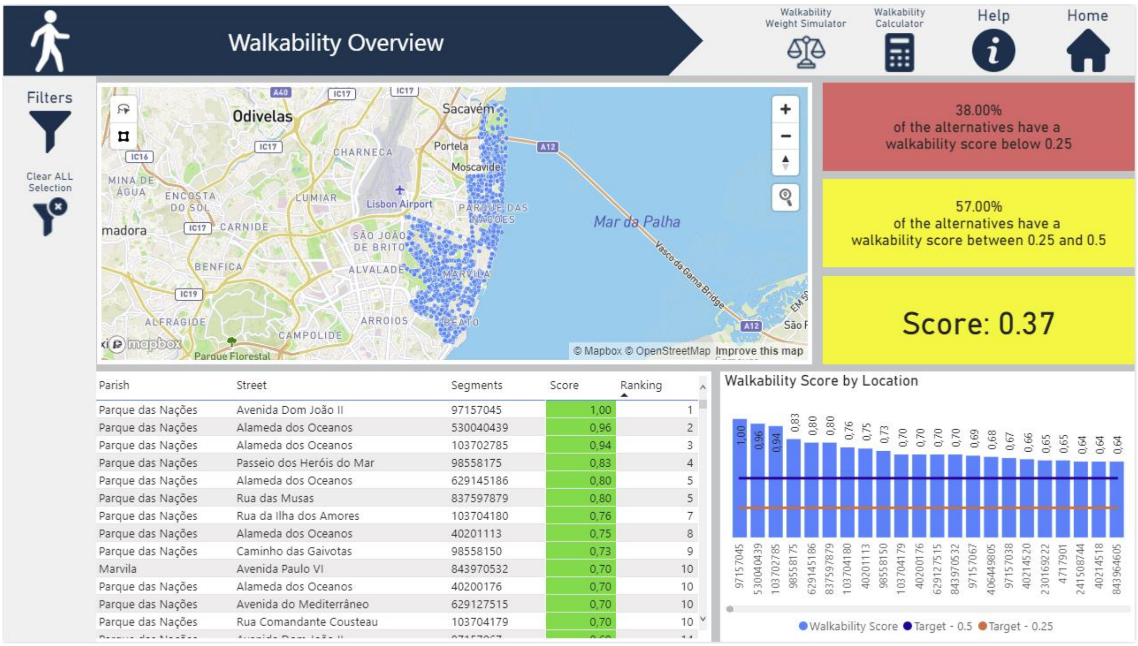

Figure 4.3 – Walkability Overview

4.4.2. Walkability Weight Simulator

The theme of this report is the same as the Walkability Overview report and is represented in Figure 4.4.

This report enables the weight simulation. Users are capable of calculating the walkability score based on their strategy. To do so, parameters for each criterion were created. Based on that, users are capable of rating the importance of each criterion between 0 to 5, where 0 represents the lowest importance and 5 represents the highest importance. Then, a weight between 0 and 1 for each criterion was calculated considering the rate defined. Figure 4.5 illustrates an example of weight simulation.

Moreover, the table visual also enables the comparison between the original and simulated scores. A Key Performance Indicator (KPI) was created using the tabular editor to illustrate the effect of the simulation. The tabular editor is an alternative to SQL Server Data Tools (SSDT) for authoring Tabular models for Analysis Services even without a workspace server. It is an open-source project capable of manipulating and managing measures, calculated columns, display folders, perspectives, and translations (SQLBI, 2022).

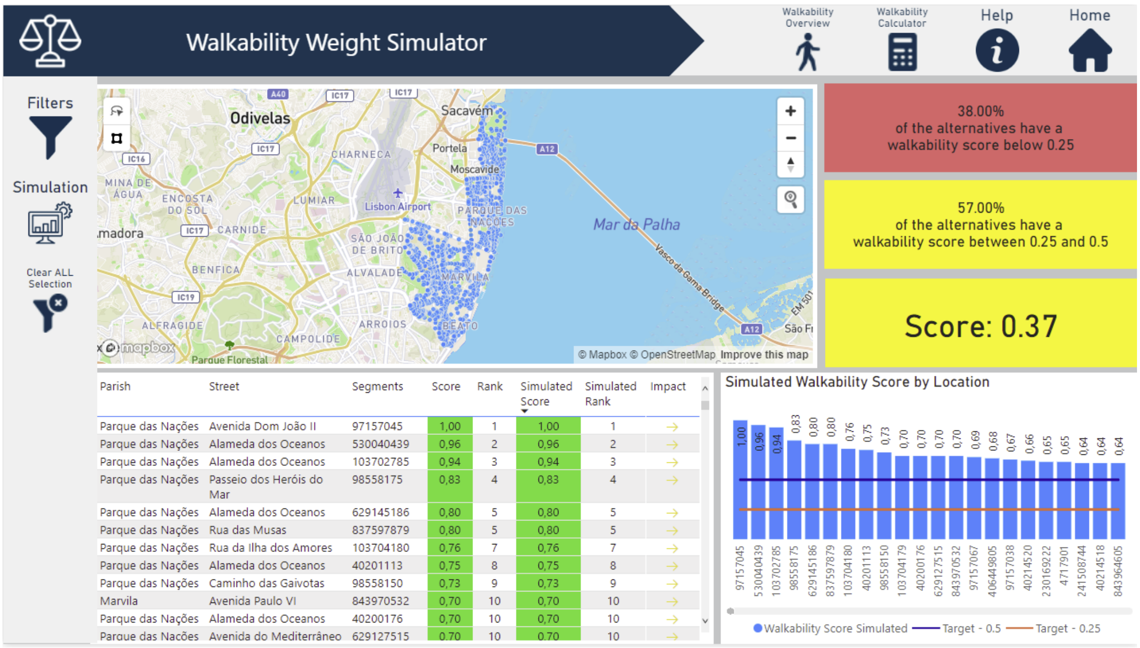

Figure 4.4 – Walkability Weight Simulator

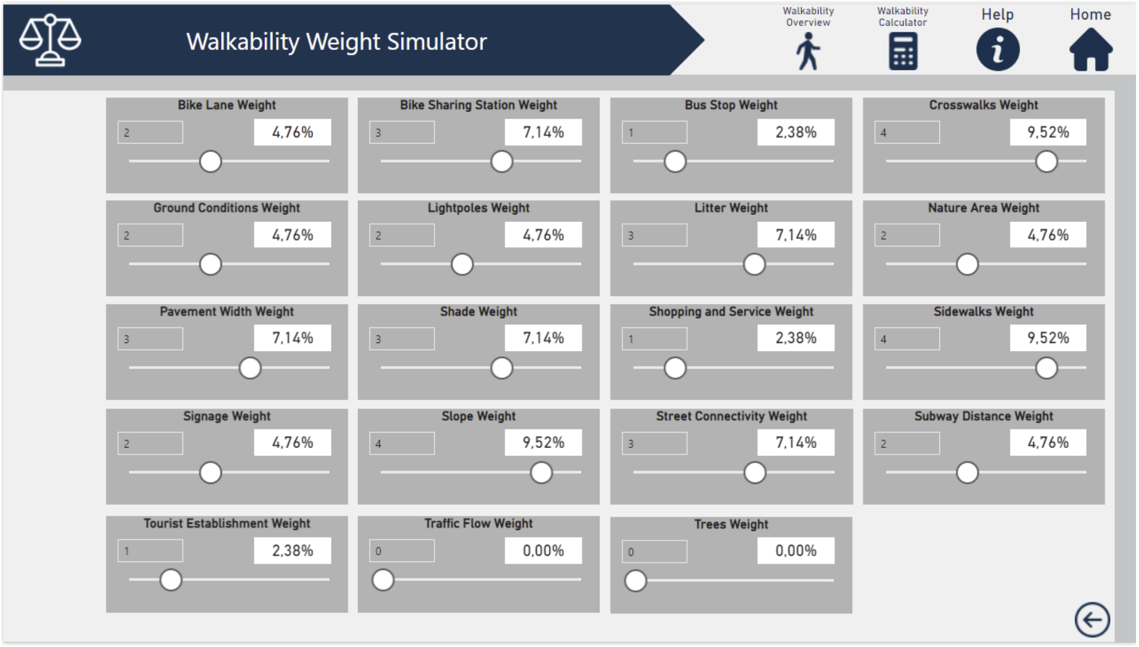

Figure 4.5 – Criterion Weight Simulation page

4.4.3. Walkability Calculator

This report intends to provide a way to calculate the walkability score where users themselves can define the value and the weight of each criterion on the segment level. Moreover, calculate the three possible results based on the segment selected: (1) Orignal Score. (2) Score calculated based on the value simulation and pre-defined weights. (3) Score calculated based on the value and weight simulated. Moreover, it also provides a comparison between the original and the simulated value of each criterion on the segment level. The output of this report is illustrated in Figure 4.6.

For the simulation purpose, the methodology described in the previous section was applied and represented in Figures 4.7 and 4.8.

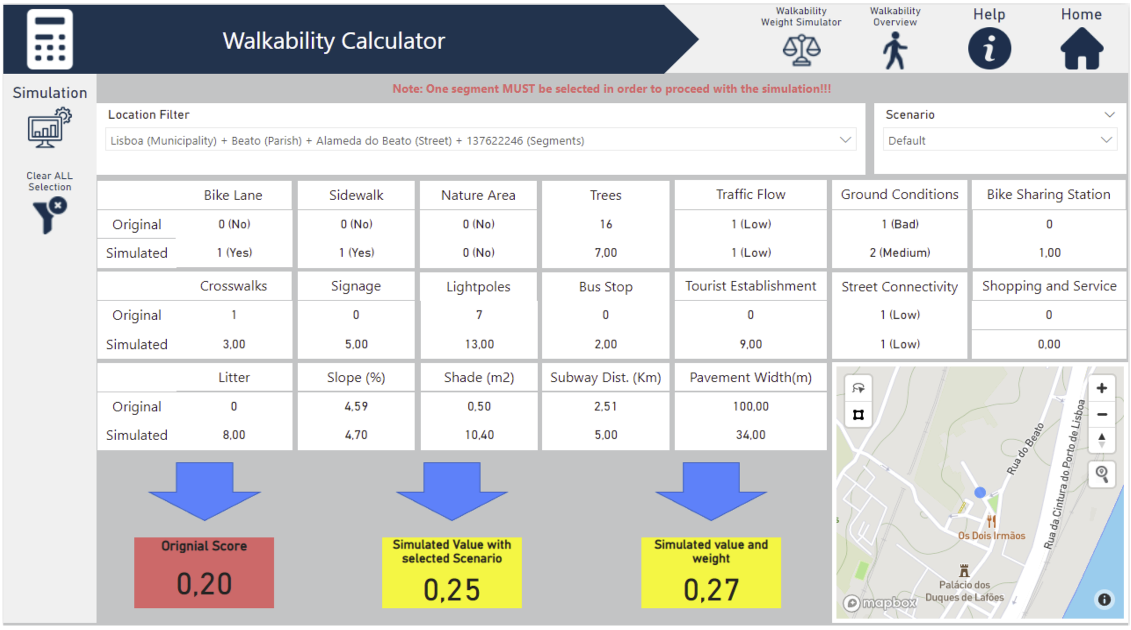

Figure 4.6 Walkability Calculator

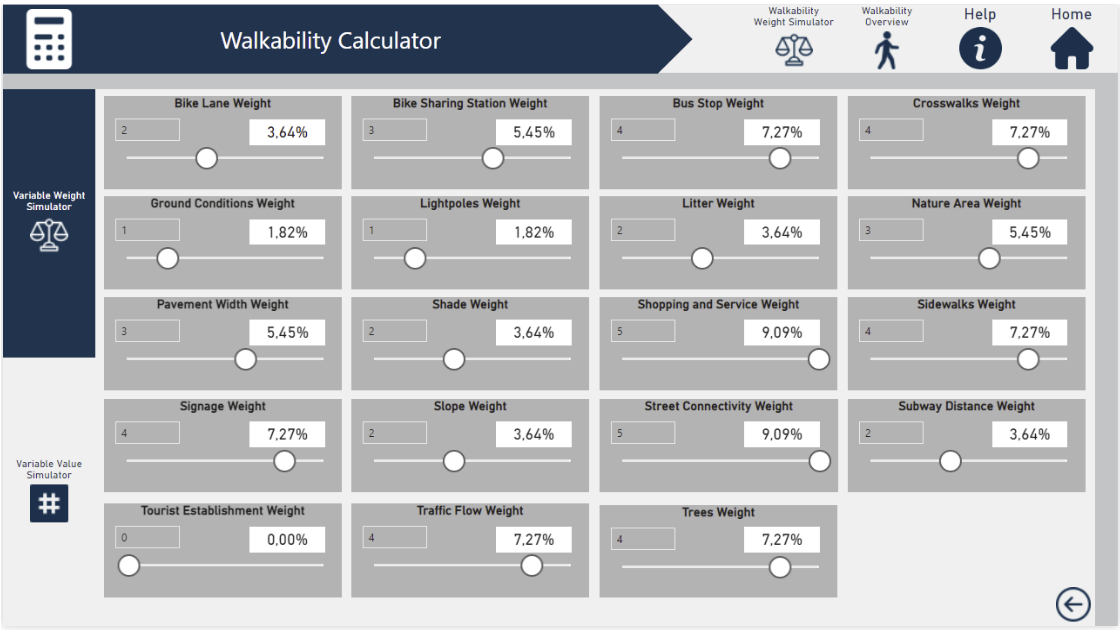

Figure 4.7 Criterion Weight Simulation page

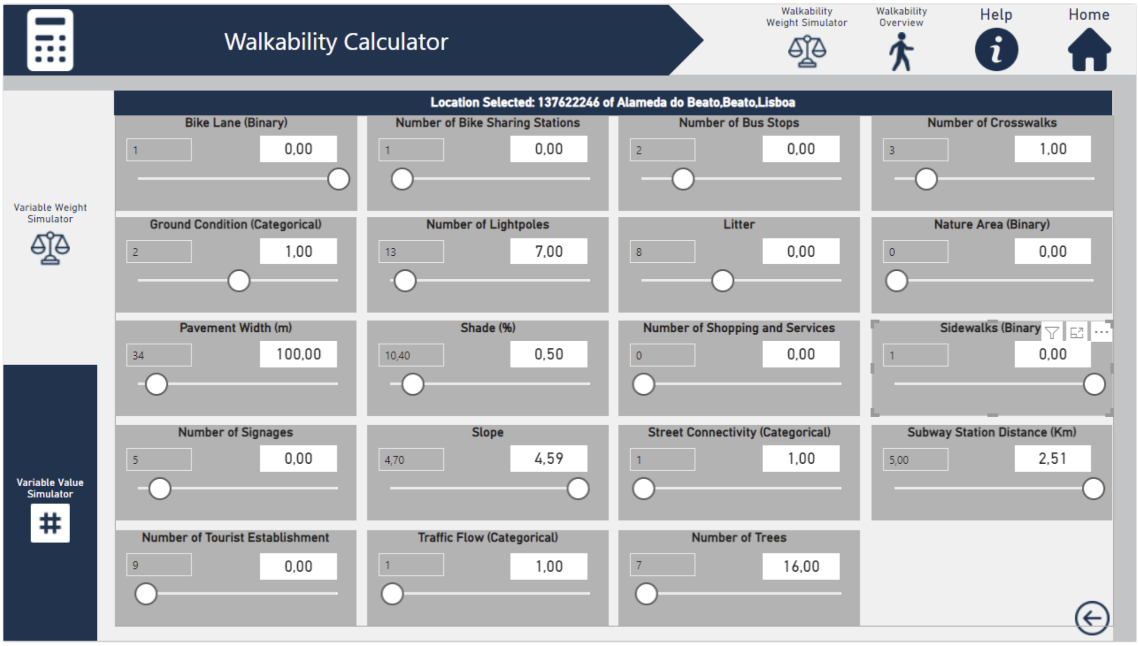

Figure 4.8 Criterion Value Simulation page

5. RESULTS AND DISCUSSIONS

This section represents the final stage of the DSR methodology, the "Evaluation." This step focuses on evaluating whether the result obtained met the previously defined objectives.

This project aims to develop a walkability simulator and visualization platform based on the evaluation proposed by (Manzolli et al., 2021). Thus, it is essential to identify the problems with the current process of the walkability measurement and its visualization. Then, a list of objectives is firstly defined based on the authors' interests and business needs. Afterward, research meant to figure out the best practice and previous works in literature that dealt with similar problems.

Based on the selected evaluation process and the suggestions of the authors, one more criterion was added to the walkability calculation, the number of light poles. It is considered an essential attribute for route safety, especially at night. By adding it, the model may obtain a more precise result regarding walkability and provide another possibility for improvements. Moreover, three previously qualitative criteria (ground condition, street connectivity, and traffic flow) were obtained in a different manner and are currently represented on a quantitative scale. This allows for a more precise representation of these criteria and a more accurate result.

Considering the aspects mentioned above, a conceptual model was created. Firstly, it is important to observe the conceptual data model created. The dimensional model was adopted in the project, which is wildly used in the BI area to deal with data. Based on the four-step dimensional design process proposed by Kimball & Ross (2011), one fact and two dimensions were identified. Moreover, an auxiliary table was created to deal with preset scenarios, which constituted one of the main goals defined. Then, the dashboard conceptual model and dashboard design was created based on the best practices studied in the literature review section.

During the project's implementation stage, Power BI was adopted to proceed with the development. It is a robust BI software that is appropriate for the problem. Based on the conceptual model defined, three reports were created and delivered. Firstly, it was necessary to create a report that demonstrates the original state of the evaluation. It provides a clear view for an intuitive reading and interpretation of the walkability score calculated based on the data collected. The ranking capability is enabled, which is desired by the authors. From another perspective, this project deals with locations. Therefore, it is interesting to have the alternatives presented on the map. Thus, Mapbox visual was adopted for this purpose.

This project enables two types of simulations. In the first place, based on the visualization created, users are capable of simulating the weight of each criterion. It allows users to choose their strategy for calculating the walkability score. Its layout is similar to the first one described above. However, it is important to understand the impact of the simulation compared to the original one. Thus a comparison between the original and the weight simulated score was created and illustrated using a KPI.

In the end, a walkability calculator was established, which allows users to simulate the weight of each criterion and the values. This functionality was restricted on the segment level, meaning users must select a segment to proceed with the simulation. The idea is to support the users, especially the DMs, in understanding the impact of an action that intends to improve the walkability score of a specific segment and how efficient and effective it can be. Moreover, a comparison between the three scores obtained in this project: the original score, the score calculated using the weight simulator, and the score calculated using the walkability calculator, as well as the original value and the simulated value of each criterion.

By evaluating the reports created and the result obtained, it is possible to mention that the proposed and implemented dashboard satisfies the defined requirements. Nevertheless, there were still some challenges during the simulator creation. Some desired capabilities cannot be created due to the limitation of the tool. Firstly, one of the desired functionality is to color the points in the Mapbox visual according to their score and the targets defined. However, the selected visual cannot customize the color of each entity until the delivery of this project. Another limitation faced is not capable of defining dynamic ranges for the parameters created. Therefore, it was hard to define the maximum and minimum values for the simulation. Especially for the walkability calculator, this functionality is only available for segment level due to this limitation. Moreover, the data flow is not automatic due to the lack of resources, which is a desirable capability that this project intends to achieve. There were still human interventions on the data updates.

It is also important to mention that the tool requires a training section for the end-users to ensure that all target users fully understand the information provided. Not only the navigability of the dashboard but also the evaluation process should be transparent to all kinds of users.

6. CONCLUSION

Improving urban pedestrian environment quality is critical while dealing with sustainable city planning. Walking, which is the most common way of transportation mode, is directly linked to improving the quality of life, taking into account health, social, and environmental aspects. Supported by findings, this project explored the concept of walkability and presented as its main goal the development of a walkability simulator based on an evaluation process previously created. Based on that, a BI solution was created using Power BI.

Firstly, this project explored the evaluation process adopted. It was created using an MCDA approach, and then eighteen criteria were defined based on the 5Cs layout. The selected criteria can be divided into three categories: route condition, route safety, and route characteristics. Then, based on the studies, a new criterion was added to the evaluation process, the number of light poles, and classified as route safety. A conceptual model was created where we defined the overall requirements, the conceptual data model, and the dashboard design.

The solution delivered is a Lisbon case study encompassing three selected parishes: Marvila, Beato, and Parque das Nações. An intensive data collection was conducted to obtain the most reliable data for the selected study area. Then, a data preparation was done to transform the collected data into an appropriate format. The currently collected data made it possible to obtain an improved evaluation process, when compared to the selected base methodology.

In the end, the dashboard was created and delivered. It provides a clear view of the walkability score for the selected alternatives. An easy read and interpret visualization platform was established. From another hand, a walkability simulator was developed. The main idea is to provide a tool capable of simulating an action that intends to improve the walkability score, analyzing its cost-efficiency and effectiveness, and helping the city planners to build a more pedestrian-friendly environment.

Moreover, this tool is built based on a solid framework for the walkability evaluation, which is replicable in other urban contexts. Thus, this tool was also designed to receive more data and be ready to be used in such scenarios.

Analyzing the result makes it possible to conclude that the dashboard delivered meets almost all the objectives defined. All the listed targets were considered during the development, and best practices mentioned in the literature review were applied.

However, many challenges and limitations were faced during the project development. Firstly, due to the lack of resources, data were read directly from the excel files. Thus, some human interventions were still needed to get the most up-to-date data. Moreover, parameters created using power BI do not allow dynamic rages, which is desirable while simulating values of each criterion. Furthermore, the dashboard used Mapbox visual to display the alternatives and their walkability scores on the map. This visual is not capable of customizing the color of each entity. Considering the targets defined, it is highly desired that each data point shows colors based on its score, which was impossible until the delivery of this project.

It is suggested that future works include subjective criteria in the evaluation process, such as pedestrian groups (adults, children, seniors, and impaired mobility pedestrians) and trip purposes (utilitarian and leisure). It is interesting to understand the impact of those factors on the walkability score and how to improve it. Depending on the pedestrian groups or trip purposes, how the walkability will be evaluated may change. Regarding the tool developed, it is interesting to have the comment capability. It enables interactions between users. Users can use this functionality to justify why a particular street or area is unfavorable to pedestrians and what can be done to improve it. In the end, more details regarding the street or area are also desirable.

7. REFERENCE

Abass, Z. I., & Tucker, R. (2018). Residential satisfaction in low-density Australian suburbs: The impact of social and physical context on neighbourhood contentment. *Journal of Environmental Psychology*, *56*, 36–45. https://doi.org/10.1016/J.JENVP.2018.02.005

Abela. (2006). Choosing a good chart - The Extreme Presentation(tm) Method. Retrieved May 11, 2022, from https://extremepresentation.typepad.com/blog/2006/09/choosing_a_good.html

Adler, P., Florida, R., King, K., & Mellander, C. (2019). The city and high-tech startups: The spatial organization of Schumpeterian entrepreneurship. *Cities*, *87*, 121–130. https://doi.org/10.1016/J.CITIES.2018.12.013

Boongaling, C. G. K., Luna, D. A., & Samantela, S. S. (2021). Developing a street level walkability index in the Philippines using 3D photogrammetry modeling from drone surveys. *GeoJournal*, 1–24. https://doi.org/10.1007/s10708-021-10441-2

Campisi, T., Basbas, S., Skoufas, A., Akgün, N., Ticali, D., & Tesoriere, G. (2020). The impact of covid-19 pandemic on the resilience of sustainable mobility in sicily. *Sustainability (Switzerland)*, *12*(21), 1–25. https://doi.org/10.3390/su12218829

Chaudhuri, S., Dayal, U., & Narasayya, V. (2011, August). An overview of business intelligence technology. *Communications of the ACM*, Vol. 54, pp. 88–98. https://doi.org/10.1145/1978542.1978562

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

Craig Stedman, Ed Burns, & Mary K. Pratt. (2022). What is Data Preparation? An In-Depth Guide to Data Prep. Retrieved May 11, 2022, from https://www.techtarget.com/searchbusinessanalytics/definition/data-preparation

C-Tech. (2020). Retrieved May 23, 2022, from https://www.mitportugal.org/research/flagship-projects/c-tech-climate-driven-technologies-for-low-carbon-cities

David Stodder. (2016). Improving Data Preparation for Business Analytics. *Redpoint.Net*. Retrieved from http://www.redpoint.net/wp-content/uploads/2016/10/TDWI_BPReport_Q316_RedPoint_F_rev2_code_Final.pdf

Dovey, K., & Pafka, E. (2020). What is walkability? The urban DMA. *Urban Studies*, *57*(1), 93–108. https://doi.org/10.1177/0042098018819727

Ewing, R., & Handy, S. (2009). Measuring the unmeasurable: Urban design qualities related to walkability. *Journal of Urban Design*, *14*(1), 65–84. https://doi.org/10.1080/13574800802451155

Fancello, G., Congiu, T., & Tsoukiàs, A. (2020). Mapping walkability. A subjective value theory approach. *Socio-Economic Planning Sciences*, *72*. https://doi.org/10.1016/j.seps.2020.100923

Florida, R., Mellander, C., & Stolarick, K. (2010). Beautiful Places: The Role of Perceived Aesthetic Beauty in Community Satisfaction. *Https://Doi.Org/10.1080/00343404.2010.486784*, *45*(1), 33–48. https://doi.org/10.1080/00343404.2010.486784

Gartner. (2022). Definition of Analytics and Business Intelligence (ABI) - Gartner Information Technology Glossary. Retrieved May 11, 2022, from https://www.gartner.com/en/information-technology/glossary/business-intelligence-bi

Gilderbloom, J. I., Riggs, W. W., & Meares, W. L. (2015). Does walkability matter? An examination of walkability's impact on housing values, foreclosures and crime. *Cities*, *42*(PA), 13–24. https://doi.org/10.1016/j.cities.2014.08.001

Giles-Corti, B., Macaulay, G., Middleton, N., Boruff, B., Bull, F., Butterworth, I., … Christian, H. (2014). Developing a research and practice tool to measure walkability: A demonstration project. *Health Promotion Journal of Australia*, *25*(3), 160–166. https://doi.org/10.1071/HE14050

Glazier, R. H., Creatore, M. I., Weyman, J. T., Fazli, G., Matheson, F. I., Gozdyra, P., … Booth, G. L. (2014). Density, destinations or both? A comparison of measures of walkability in relation to transportation behaviors, obesity and diabetes in Toronto, Canada. *PLoS ONE*, *9*(1). https://doi.org/10.1371/journal.pone.0085295

Götz, F. M., Yoshino, S., & Oshio, A. (2020). The association between walkability and personality: Evidence from a large socioecological study in Japan. *Journal of Environmental Psychology*, *69*. https://doi.org/10.1016/j.jenvp.2020.101438

Iivari, J., & Venable, J. R. (2009). *Action research and design science research-Seemingly similar but decisively dissimilar*. 2009. Retrieved from https://aisel.aisnet.org/ecis2009/73/

Janvrin, D. J., Raschke, R. L., & Dilla, W. N. (2014). Making sense of complex data using interactive data visualization. *Journal of Accounting Education*, *32*(4), 31–48. https://doi.org/10.1016/J.JACCEDU.2014.09.003

Kimball, R., & Ross, M. (2011). *The data warehouse toolkit: the complete guide to dimensional modeling*. Retrieved from https://books.google.com/books?hl=pt-PT&lr=&id=XoS2oy1IcB4C&oi=fnd&pg=PA1&dq=data+warehouse+toolkit+%22dimensional+modeling%22&ots=1DGepEmPhG&sig=k00eoMlUcsijKjJ23EdYClpksas

Leinberger, C., & Alfonzo, M. (2012). *Walk this way: The economic promise of walkable places in metropolitan Washington, DC*. Retrieved from https://nyuscholars.nyu.edu/en/publications/walk-this-way-the-economic-promise-of-walkable-places-in-metropol

Litman, T. A. (2003a). Economic Value of Walkability. *Transportation Research Record*, (1828), 3–11. National Research Council. https://doi.org/10.3141/1828-01

Litman, T. A. (2003b). Economic Value of Walkability: *Https://Doi.Org/10.3141/1828-01*, (1828), 3–11. https://doi.org/10.3141/1828-01

Lo, R. H. (2009). Walkability: What is it? *Journal of Urbanism*, *2*(2), 145–166. https://doi.org/10.1080/17549170903092867

Manzolli, J. A., Oliveira, A., & Neto, M. de C. (2021). Evaluating Walkability through a Multi-Criteria Decision Analysis Approach: A Lisbon Case Study. *Sustainability*, *13*(3), 1450. https://doi.org/10.3390/su13031450

Mariana Bento. (2021). *HOW DATA VISUALIZATION TOOLS CAN IMPROVE DECISION MAKING PROCESSES BASED ON CUSTOMER SATISFACTION?*

Microsoft Power BI. (2022a). What is Power BI | Microsoft Power BI. Retrieved May 11, 2022, from https://powerbi.microsoft.com/en-us/what-is-power-bi/

Microsoft Power BI. (2022b). What is Power BI? - Power BI | Microsoft Docs. Retrieved May 11, 2022, from https://docs.microsoft.com/en-us/power-bi/fundamentals/power-bi-overview

Microsoft Power BI. (2022c). Why Power BI - Features & Benefits | Microsoft Power BI. Retrieved May 11, 2022, from https://powerbi.microsoft.com/en-us/why-power-bi/

Moura, F., Cambra, P., & Gonçalves, A. B. (2017). Measuring walkability for distinct pedestrian groups with a participatory assessment method: A case study in Lisbon. *Landscape and Urban Planning*, *157*, 282–296. https://doi.org/10.1016/j.landurbplan.2016.07.002

Oussous, A., Benjelloun, F. Z., Ait Lahcen, A., & Belfkih, S. (2018). Big Data technologies: A survey. *Journal of King Saud University - Computer and Information Sciences*, *30*(4), 431–448. https://doi.org/10.1016/J.JKSUCI.2017.06.001

Painter, K. (1996). The influence of street lighting improvements on crime, fear and pedestrian street use, after dark. *Landscape and Urban Planning*, *35*(2–3), 193–201. https://doi.org/10.1016/0169-2046(96)00311-8

Paiva, S., Ahad, M. A., Tripathi, G., Feroz, N., & Casalino, G. (2021, March 2). Enabling technologies for urban smart mobility: Recent trends, opportunities and challenges. *Sensors*, Vol. 21, pp. 1–45. MDPI AG. https://doi.org/10.3390/s21062143

Papageorgiou, G., Maimaris, A., Efstathiadou, T., & Balamou, E. (2018). Evaluating attitudes on the quality of service of pedestrian networks. *WIT Transactions on the Built Environment*, *176*, 35–40. WITPress. https://doi.org/10.2495/UT170041

Patnaik, S., & Jain, V. (2017). *Recent Developments in Intelligent Computing, Communication and Devices: Proceedings of ICCD 2017*. Retrieved from https://link.springer.com/chapter/10.1007/978-981-10-8944-2_25

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, *24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Pelegrina, G. D., Duarte, L. T., & Romano, J. M. T. (2019). Application of independent component analysis and TOPSIS to deal with dependent criteria in multicriteria decision problems. *Expert Systems with Applications*, *122*, 262–280. https://doi.org/10.1016/j.eswa.2019.01.008

Porębska, A., Rizzi, P., Otsuki, S., & Shirotsuki, M. (2019). Walkability and resilience: A qualitative approach to design for risk reduction. *Sustainability (Switzerland)*, *11*(10). https://doi.org/10.3390/su11102878

QGIS. (2022). Welcome to the QGIS project! Retrieved May 11, 2022, from https://qgis.org/en/site/

Rebecchi, A., Buffoli, M., Dettori, M., Appolloni, L., Azara, A., Castiglia, P., … Capolongo, S. (2019). Walkable environments and healthy urban moves: Urban context features assessment framework experienced in Milan. *Sustainability (Switzerland)*, *11*(10). https://doi.org/10.3390/su11102778

Rosenberg, D., Ding, D., Sallis, J. F., Kerr, J., Norman, G. J., Durant, N., … Saelens, B. E. (2009). Neighborhood Environment Walkability Scale for Youth (NEWS-Y): Reliability and relationship with physical activity. *Preventive Medicine*, *49*(2–3), 213–218. https://doi.org/10.1016/J.YPMED.2009.07.011

Ruiz-Padillo, A., Pasqual, F. M., Larranaga Uriarte, A. M., & Cybis, H. B. B. (2018). Application of multi-criteria decision analysis methods for assessing walkability: A case study in Porto Alegre, Brazil. *Transportation Research Part D: Transport and Environment*, *63*, 855–871. https://doi.org/10.1016/j.trd.2018.07.016

Soham Sinha. (2021). Gutenberg & modern-day UX. Retrieved May 11, 2022, from https://uxdesign.cc/gutenberg-modern-day-ux-14ac545e151c

SQLBI. (2022). Tabular Editor. Retrieved May 11, 2022, from https://www.sqlbi.com/tools/tabular-editor/

Talen, E. (2002). Pedestrian access as a measure of urban quality. *Planning Practice and Research*, *17*(3), 257–278. https://doi.org/10.1080/026974502200005634

Taras Bakusevych. (2018). 10 rules for better dashboard design. Retrieved May 11, 2022, from https://uxplanet.org/10-rules-for-better-dashboard-design-ef68189d734c

TIBCO Software. (2022). What is Information Visualization? | TIBCO Software. Retrieved May 11, 2022, from https://www.tibco.com/reference-center/what-is-information-visualization

UNDP. (2022). Goal 11: Sustainable cities and communities. Retrieved May 11, 2022, from https://www1.undp.org/content/seoul_policy_center/en/home/sustainable-development-goals/goal-11-sustainable-cities-and-communities.html

van Aken, J. E. (2004). Management Research Based on the Paradigm of the Design Sciences: The Quest for Field-Tested and Grounded Technological Rules. *Journal of Management Studies*, *41*(2), 219–246. https://doi.org/10.1111/J.1467-6486.2004.00430.X

Vercellis, C. (2009). Business Intelligence: Data Mining and Optimization for Decision Making. *Business Intelligence: Data Mining and Optimization for Decision Making*, 1–417. https://doi.org/10.1002/9780470753866

Ware, C. (2004). Information Visualization: Perception for Design: Second Edition. *Information Visualization: Perception for Design: Second Edition*, 1–486. https://doi.org/10.1016/B978-1-55860-819-1.X5000-6

Watson, H. J. (2009). Tutorial: Business intelligence - Past, present, and future. *Communications of the Association for Information Systems*, *25*(1), 487–510. https://doi.org/10.17705/1CAIS.02539

Zook, J. B., Lu, Y., Glanz, K., & Zimring, C. (2012). Design and pedestrianism in a smart growth development. *Environment and Behavior*, *44*(2), 216–234. https://doi.org/10.1177/0013916511402060