POLITECNICO DI TORINO

UNIVERSIDAD CARLOS III DE MADRID

Department of Electronics and Telecommunications

Collegio di Ingegneria Elettronica, delle Telecomunicazioni e Fisica

Master's degree programme

ICT FOR SMART SOCIETIES LM-27(DM270)

Service platform to calculate and analyse access distance to public services

Analysis of the case of the Lazio region

Supervisors:

Professor Javier Garcia Guzman

Professor Cristina Pronello

Candidate:

Kevin Gumina S320054

July 2025

Academic year 2024-2025

V1 - 9th of June 2025: First official version of the thesis submitted to University Carlos III of Madrid.

V2 - 16th of June 2025: Preliminary version of the thesis submitted to Politecnico di Torino University

V3 - 26th of June 2025: Official document of the thesis available on the Politecnico di Torino University library

All contents, results, and materials produced in the scope of this work are available at the GitHub repository: https://github.com/keving-ing/Distance-project. If the repository is temporarily or permanently inaccessible, please contact the author at the following institutional addresses: 100547035@alumnos.uc3m.es, s320054@studenti.polito.it or at the personal email address: kevingumina.kg@gmail.com.

Dedicated to my land,

May this project also be one of the tools
that contribute to shaping a more developed,
accessible and inclusive future for the
places that have helped shape my identity.

Questo lavoro rappresenta la fine di un percorso universitario iniziato nel 2019, un cammino che anno dopo anno mi ha arricchito non solo a livello formativo, grazie alle conoscenze acquisite e alle sfide affrontate nei diversi corsi sostenuti, ma anche sul piano personale. In questi sei anni, tra il Politecnico di Torino e l'Universidad Carlos III de Madrid, ho imparato a crescere, a superare ostacoli e a trarre insegnamenti anche dai momenti più difficili.

Tutto questo anche grazie all'aiuto speciale della mia famiglia, che è stata la mia colonna portante in tutto questo percorso. In particolare, grazie ai miei genitori, a Katia, Mattia e Zia Santina: il vostro sostegno, non solo economico ma soprattutto umano, è stato fondamentale. Mi avete sempre incoraggiato, senza pressioni né giudizi, lasciandomi libero di affrontare il mio cammino, con i suoi tempi e i suoi inevitabili momenti di difficoltà, trasformandoli in tappe di crescita.

Un pensiero affettuoso va anche alle Sapone's e ai Tre Moschettieri, amici di sempre che, nonostante la distanza di oltre 2000 km, non mi hanno mai fatto sentire solo tutte le volte che sono tornato a casa. Con la loro presenza, anche solo virtuale, hanno saputo alleggerire giornate pesanti e regalarmi sorrisi con aggiornamenti e "curtigghi" sul nostro gruppo.

Un grazie speciale a Peppe Mangano, che ha sopportato pazientemente le mie lamentele, i miei sfoghi e qualche volta anche la mia rabbia, talvolta ingiustamente riversata su di lui. Nonostante tutto, c'è sempre stato, pronto a chiamarmi, distrarmi e a farmi ridere parlando delle partite dell'Orlandina o delle storie leggendarie di Torrenova e degli arbitri dei "tempi d'oro".

Non posso non menzionare tutte le persone incontrate in questi anni fuori "casa" a Torino e a Madrid, che hanno contribuito a rendere questo percorso unico e indimenticabile, costruendo la nuova "casa" a cui mi sento di appartenere oggi. Un grazie a Mario, Mo, ai Gestionali, a Tonino, a San Millán, a Salmoreno e Ernesto La Torre con Nik e Tony: con voi ho condiviso, risate, esperienze, viaggi e momenti preziosi fuori dallo studio, indispensabili per mantenere l'equilibrio e l'energia necessari per arrivare fino in fondo a questo traguardo.

Infine, un sentito ringraziamento al Professor Javier per avermi accolto e guidato con disponibilità nella realizzazione di questo progetto, a Maria e Aitana per avermi fatto sentire sin da subito parte della squadra del laboratorio 2.2.B.10, e alla Professoressa Pronello, che anche a distanza ha creduto in me e mi ha seguito con preziosi consigli e suggerimenti, fondamentali per la definizione e l'organizzazione di questo lavoro.

¡A tutti voi, grazie davvero!

Abstract

This thesis investigates the accessibility of essential public services in the Lazio region (Italy), with a particular focus on the most vulnerable territorial units: municipalities with fewer than 40,000 inhabitants. The motivation stems from the observation that larger urban centres generally host a full range of services and thus present fewer accessibility issues, whereas smaller municipalities often suffer from infrastructural disparities. The analysis adopts the urban cores as the fundamental spatial unit, as these represent the actual population concentrations within municipal boundaries.

The study focuses on key sectors such as healthcare and education, for which accessibility is assessed by computing travel distances and times from each urban cores to the nearest service facility. These measurements are performed under two distinct scenarios: travel by private vehicle, accounting for traffic conditions, and travel by public transport. Due to data availability limitations, the public transport analysis is limited to the province of Rome. The accessibility measures are synthesized into sector-specific indices that can be analysed both at the urban cores level and in aggregated form at the municipal level.

To support dissemination and usability of the data infrastructure created, an interactive web-based application was developed. The platform includes a dynamic map viewer that allows stakeholders, planners, and citizens to explore the computed accessibility indices, inspect individual urban centres, and access the underlying data infrastructure in a transparent and intuitive manner.

The results reveal spatial disparities in service accessibility, particularly when comparing private vehicle use with public transport. While car accessibility remains relatively consistent across services, with only a few areas showing critical issues, public transport accessibility presents substantial gaps, especially in peripheral zones. These findings underscore the importance of evaluating accessibility across multiple transport modes depending on the type of service and, when combined with other key socio-territorial data, highlight priority areas for infrastructure planning and policy action.

Key words

Data Analysis, Accessibility, Public Services, Lazio Region, Linked Open Data
Web Application, Map Viewer

SUSTAINABLE DEVELOPMENT GOALS

Contents

List of Figures

List of Tables

1. Introduction

The spatial distribution of essential public services and the population's ability to access them efficiently are central themes in regional planning and territorial equity. Accessibility, defined as the ease with which inhabitants can reach key services such as healthcare facilities or educational institutions, is increasingly used as an analytical lens to assess the adequacy of service provision and to identify disparities across different areas. (Geurs and van Wee, 2004) It is commonly conceptualised as comprising three key dimensions: distance, travel time, and impedance, which reflects the relationship between demand and supply or other barriers to access. This study primarily focuses on the first two dimensions, distance and travel time, as they are more readily quantifiable using available data and spatial analysis techniques. The impedance dimension, although not directly addressed here, is acknowledged as a valuable component for future methodological development. In this framework, accessibility studies are crucial for informing policies that aim to reduce territorial inequalities and foster more inclusive development strategies. (Páez et al., 2012)

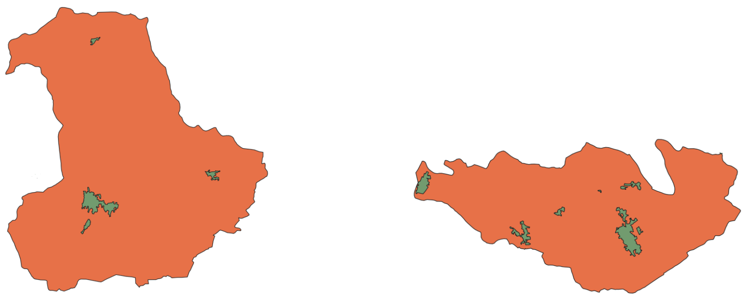
The analysis is applied to the Lazio region in Italy, a territory characterised by strong socio-spatial heterogeneity, where the contrast between urban centres and peripheral or rural areas poses significant challenges in terms of service accessibility. (ISTAT, 2022) To capture this complexity, the study adopts urban centres, defined by the Italian National Institute of Statistics (ISTAT), as the unit of analysis. These centres offer a more precise spatial resolution than municipal boundaries alone, especially when assessing intra-municipal variability. By focusing on urban centres, the research aims to provide a fine-grained and realistic picture of accessibility conditions, with particular attention to municipalities with fewer than 40,000 inhabitants, a threshold often associated with more limited infrastructure and lower service provision. (OECD, 2020)

Building on this territorial framework, the study also incorporates a comparative dimension. Accessibility data from the Community of Madrid are used to benchmark and contrast findings, offering both methodological robustness and policy relevance. This international comparison not only enriches the scientific scope of the research, but also provides practical insights for regional and national decision-making.

To enable such analysis, a robust and interoperable data infrastructure is essential. This infrastructure must support open access, facilitate the integration of diverse datasets from multiple sources, and allow for regular updates. Moreover, it should be designed to be usable even by non-technical stakeholders, such as local administrators and urban planners. By reducing technical barriers, this approach promotes broader engagement with accessibility data and fosters more informed, inclusive governance.

Despite the growing interest in accessibility analysis, several practical challenges remain. One major issue is that disparities are particularly acute in small and medium-sized municipalities, where residents often face difficulties in reaching essential services. These challenges are compounded by the uneven distribution of services and variations in transport infrastructure, especially in geographically large municipalities with scattered population clusters. In such contexts, some zones may have excellent access, while others are severely underserved, yet aggregate municipal-level indicators often fail to capture this internal variability. (OECD, 2022)

To address these gaps, detailed data are needed on travel distances and times, by car and public transport, from each urban centre to key services. For many residents, especially those in peripheral areas, this translates into longer journeys, higher costs, and reduced access to essential opportunities in health, education, and employment. These barriers contribute to social exclusion and degrade quality of life, particularly for those without access to private transport. (ESPON, 2020) Figure 1 illustrates the internal dispersion of urban centres in two municipalities of the Lazio region, showing how population clusters within the same administrative boundary can face drastically different accessibility conditions. (ISTAT, 2024a, 2021)



(a) *Urban centres of Acquapendente* (b) *Urban centres of Palomara Sabina*

Figure 1: *Spatial distribution of urban centres within a single municipality in the Lazio region*

Another methodological hurdle involves integrating diverse and often incompatible datasets, including demographic data, service locations, road networks, public transport schedules, and administrative boundaries, into a coherent, unified database. The lack of standardisation and synchronisation across sources poses significant obstacles to the construction of a consistent and updatable accessibility indicator.

In addition, current policy frameworks frequently rely on municipal-level aggregates, which obscure sub-municipal differences and prevent a precise identification of underserved areas. Although some accessibility indicators are included in institutional reports, these are rarely produced with regularity or integrated into interactive platforms. This limits their usefulness for non-technical users and decision-makers, who often lack the tools to explore and understand spatial inequalities in an intuitive way.

Addressing these limitations requires both technical and institutional innovation. On the one hand, developing user-friendly digital platforms can democratise access to spatial information and make the data actionable for policy design. On the other hand, bridging the gap between data producers and end-users remains a key challenge: researchers, planners, and policymakers need accessible, interoperable tools to interpret complex spatial data and to design effective interventions for territorial cohesion.

Finally, the value of comparative analysis cannot be overstated. Benchmarking results against other regions helps validate the findings, improves methodological transparency, and encourages the exchange of best practices. In this study, comparing the Lazio region with the Community of Madrid offers a concrete example of how differences in spatial organisation, data infrastructures, and planning approaches influence accessibility outcomes, thus informing more equitable and effective regional policies.

To guide the reader through the research process, the structure of this document has been designed to follow the logical development of the project, illustrating how its objectives were pursued and achieved. Each chapter details a specific phase of the work, from methodological foundations to practical implementation and analysis of results.

The second chapter presents the state of the art. It contains a review of existing studies and projects relevant to the domain and geographic context, providing a technical overview of the tools and technologies considered. Socio-economic implications and the relevant legal and policy frameworks are also discussed. Particular emphasis is placed on identifying the gaps in current approaches, thereby establishing what changes and improvements are necessary to achieve the stated objectives.

Chapter three focuses on the objectives and presents the methodology. First, the chapter highlights the importance of accessibility analysis within the selected region. A general objective is then defined, followed by a set of specific objectives aimed at addressing the main research question. Alongside the theoretical references, a detailed account is provided of the practical processes used to meet the objectives, from data collection and preprocessing (including cleaning and ETL procedures), to distance calculation, data integration, the generation of accessibility indicators, and how these indicators are published and presented to stakeholders.

The last chapter presents the results of the analysis and reflects on how well the objectives have been met within the context of the Lazio region. The document concludes with a summary of the findings and a discussion of future work and potential improvements.

Finally, the document contains Appendix A, which describes the organisational structure of the research project. It details the operational phases from start to finish and includes a budget estimate quantifying the economic value of the work undertaken.

2. State of the art

This chapter provides the conceptual, technical, and methodological foundations necessary to understand the project, offering a comprehensive overview of the fields of knowledge involved. Given the interdisciplinary nature of the work, multiple domains have been explored and analysed to support the achievement of the stated objectives.

The section begins by introducing the concept of accessibility, a central element in regional planning and policy evaluation. The analysis will examine how accessibility is defined in the literature, the fundamental components involved in its measurement, and the most commonly used indicators. Particular attention is devoted to the various notions of distance: geometric, topological, temporal, and perceived, which play a critical role in accessibility analysis.

The chapter will then provide an overview of the technologies used for processing spatial data, including open-source platforms and GIS tools. Given the nature of the data collection process adopted in this project, a dedicated section is included on web scraping techniques, describing both the rationale behind their use and the practical considerations necessary to extract and standardise data from online sources.

The functioning of web technologies will be examined, with a focus on front-end development frameworks, mapping libraries, and tools available for creating interactive visualisations. The discussion will provide justification for the design choices made during the implementation of the map-based interface used to present accessibility indicators to stakeholders and citizens.

A core part of this chapter is dedicated to a review of existing literature and applied projects related to accessibility analysis. This includes both European level initiatives and, more specifically, regional projects focusing on the Lazio region. Each will be discussed critically, highlighting their strengths and limitations, and identifying the gaps this project intends to address.

Finally, the chapter contextualises the socio-economic relevance of the work, underscoring its alignment with current challenges in spatial justice and regional cohesion. The policy and legal framework underpinning this type of analysis will also be outlined, with references to national strategies, European directives, and international standards.

All the material presented in this section is based on peer reviewed academic publications, institutional reports, and official standards. The goal is twofold: to capture the perspectives and methodologies adopted by researchers and practitioners in the field, and to provide a solid foundation for readers who may be unfamiliar with the technical and conceptual aspects of the subject.

2.1 Accessibility: definitions and key components

In urban and transport planning, accessibility is a foundational concept, broadly defined as the potential for individuals to reach spatially dispersed opportunities such as employment, healthcare, and education. (Páez et al., 2012) It is increasingly recognised not only as a key performance indicator for transport systems but also as a critical dimension of social equity and territorial cohesion. According to Páez et al. (2012), accessibility can be conceptualised through both positive (descriptive of actual travel behaviour) and normative (prescriptive of desired standards) frameworks. This distinction is essential for informing planning strategies that are not only empirically grounded but also aligned with broader policy objectives. In line with this, accessibility measures typically combine two fundamental components: the cost of travel, which may refer to distance, time, or effort, and the availability of opportunities at various destinations. (Páez et al., 2012) These components can be operationalised through different methodologies, ranging from cumulative opportunity indicators to gravity based or utility based models, each with its own strengths and limitations depending on the use case. (Páez et al., 2012), (Geurs and van Wee, 2004)

For instance, cumulative indicators assess the number of opportunities reachable within a threshold, while gravity based measures weigh opportunities by distance decay functions. From a policy perspective, accessibility is gaining traction for its ability to integrate transportation and land, use planning in a coherent framework, making it a versatile tool for sustainability, oriented governance. (Páez et al., 2012) Moreover, the relevance of accessibility extends to addressing spatial inequalities. Research has shown that individuals in peripheral or rural areas often face considerable barriers to accessing basic services, highlighting the need for disaggregated, fine grained assessments. (Geurs and van Wee, 2004) Accessibility, therefore, offers not only a diagnostic lens but also a strategic lever for enhancing territorial equity and social inclusion.

2.1.1 Types of accessibility measures

Accessibility measures can be broadly categorised into three main classes: cumulative opportunity indicators, gravity-based models, and utility-based measures. (Páez et al., 2012) These approaches differ in how they operationalise the core components of accessibility, namely, the cost of travel and the availability of opportunities at various destinations. Cumulative opportunity measures are among the most intuitive and widely used. They quantify the number of opportunities (e.g., schools, hospitals) that can be reached within a predefined threshold of distance or time from a given origin. For instance, a typical application might measure the number of hospitals reachable within 30 minutes by car. These indicators are relatively simple to compute and easy to interpret, making them useful for policy communication. However, they are sensitive to the choice of threshold and may not account for variations in service quality or individual behaviour. Gravity-based measures, inspired by Newtonian physics, introduce a weighting mechanism that reflects the "attractiveness" of each destination and the "resistance" imposed by travel cost. They consider all possible destinations but weigh them inversely according to travel distance or time, thus, closer opportunities contribute more to the accessibility score than distant ones. These models are more robust and flexible than cumulative measures, but they require careful calibration and are more complex to implement. A third category is formed by utility-based accessibility measures, derived from discrete choice modelling, particularly logistic models. These estimate the probability of individuals choosing specific destinations based on factors such as travel time, mode, and socio-economic attributes. While theoretically appealing and behaviourally grounded, these models can be computationally demanding and often suffer from data limitations and aggregation issues.

Accessibility can also be analysed through the lens of positive and normative frameworks. Positive measures reflect actual travel behaviour, derived from observed data such as travel surveys or GPS traces. They are useful for describing current conditions and identifying patterns of spatial inequality. Normative measures, on the other hand, define how far people should have to travel to access services, often based on planning standards or average acceptable distances. While positive indicators are empirically grounded, normative ones serve as benchmarks for equity and policy evaluation. An important contribution to the literature is the integration of these two perspectives. For example, Páez et al. (2012) apply both normative and positive cumulative indicators in a case study in Montreal to compare actual access to day care services with policy-based thresholds. Their findings show significant mismatches, especially for vulnerable groups like single mothers without private transport, who often fall short of normative accessibility levels.

Finally, the choice of distance metric used in accessibility calculations plays a critical role. Euclidean distance is easy to compute but rarely reflects real-world travel conditions. Network distance and travel time offer more realistic measures, accounting for infrastructure, traffic, and mode of transport. In some cases, perceived distance, a subjective measure influenced by personal constraints or experiences, can be relevant, especially in studies of accessibility for vulnerable populations or in contexts involving psychological barriers.

Distance metrics in accessibility analysis

An essential aspect of accessibility measurement concerns the definition of distance, which may seem intuitive but in fact encompasses a variety of approaches with substantial analytical implications. The type of distance used can alter both the quantitative outcomes and the policy interpretation of accessibility indicators.

Euclidean distance, or straight-line distance, is the simplest and most commonly used metric due to its computational ease. However, it often fails to capture real travel conditions, especially in regions with complex terrain or fragmented infrastructure. (Páez et al., 2012) For instance, a school may be 2 km away "as the crow flies", but 5 km by road due to natural or infrastructural barriers. The map 2 of the Lazio region highlights exactly the diversity of the territory, characterised by mountainous, hilly and coastal areas. These geographical features can significantly affect accessibility to services, underlining the importance of considering the appropriate distances based on territorial context in accessibility analyses.

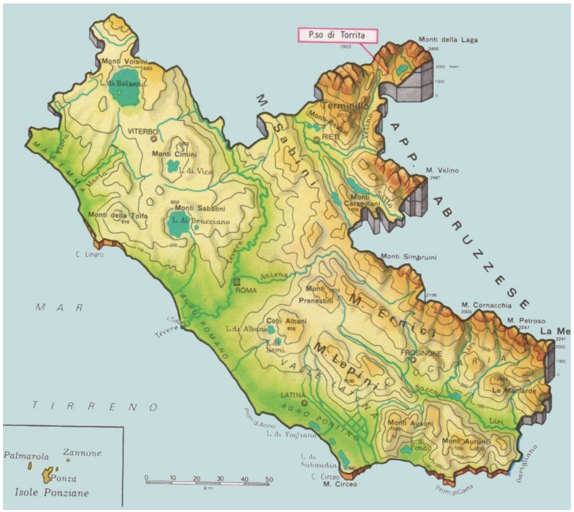


Figure 2: *Physical map of the Lazio region*

Network distance represents a more realistic option by computing the shortest or fastest path along the actual road or transport network. It accounts for factors like road hierarchy, walkability, or transport mode availability, making it suitable for analyses that require operational precision. (Geurs and van Wee, 2004) Travel time is often regarded as the most behaviourally relevant measure, particularly when it includes dynamic elements such as congestion, speed limits, or timetable constraints in the case of public transport. It allows for a direct translation of accessibility into user experience, making it more suitable for demand-side analysis and policy scenarios. Perceived distance introduces a subjective component. It accounts for the way individuals psychologically experience distance, which may vary according to age, income, physical ability, or trip purpose. For example, 800 metres may feel very different to a healthy adult compared to an elderly person or someone with limited mobility. As Páez et al. (2012) point out, this subjectivity makes perceived distance crucial in equity-focused accessibility studies, particularly when analysing access for vulnerable groups. Ultimately, the choice of distance metric should align with the goal of the analysis. While Euclidean distance may be acceptable for preliminary screenings, more sophisticated metrics like travel time or network-based distance are required for operational planning and policy interventions. Moreover, recent research encourages multi-metric approaches, where different distance types are used in parallel to triangulate and validate results.

In summary, the choice of distance metric plays a crucial role in accessibility studies, as it directly influences how accessibility scores are computed, interpreted, and applied in planning contexts. While distance is often perceived as a simple, objective quantity, its measurement can vary significantly depending on the metric selected. Euclidean distance is the most basic and computationally efficient metric, representing a straight line between two points. However, it fails to account for the actual structure of the transport network, resulting in systematic underestimation of real travel distances, particularly in urban environments with complex street layouts. For example, Mizen et al. (2015) found that Euclidean-based estimates can diverge considerably from network-based ones, with differences exacerbated in areas with fragmented or indirect road systems. (Mizen et al., 2015) Network distance, calculated along the actual transport infrastructure, offers a more realistic alternative. It considers the layout of roads or paths and can incorporate route restrictions. This makes it especially useful when evaluating pedestrian or vehicular access to services in both urban and rural settings. (Geurs and van Wee, 2004) Travel time is even more behaviourally relevant, especially when it includes dynamic factors such as traffic congestion. Accessibility scores based on free-flow travel times can significantly overestimate real-world accessibility, particularly during peak hours or in dense urban areas. Incorporating real-time or average congestion data allows for a more accurate representation of temporal barriers to access. (Geurs and van Wee, 2004) In policy contexts, this distinction can be decisive, what may seem accessible in kilometres might be functionally inaccessible in minutes. Finally, perceived distance accounts for subjective factors such as safety, comfort, or familiarity. Though harder to quantify, it is critical when analysing vulnerable populations such as elderly residents or people with mobility limitations, who may perceive even short distances as significant barriers. (Páez et al., 2012) In light of these considerations, researchers increasingly recommend using multi-metric approaches, combining several types of distance to triangulate insights and avoid bias. Moreover, aligning the metric with the goal of the study, whether technical planning, equity analysis, or behavioural research, is essential for ensuring the validity of conclusions.

The image 3 shows how different distance metrics influence neighbourhood configuration and, consequently, the analysis of accessibility to healthcare services. In particular, it highlights the discrepancies between Euclidean, Manhattan and Minkowski distances in determining service catchment areas. This underlines the importance of choosing the most appropriate distance metric to ensure accurate accessibility analysis.

2.2 Methodological approaches to study accessibility

Over the years, numerous studies have addressed the analysis and measurement of accessibility to public services, developing a range of methodological frameworks. In the literature, two main perspectives are commonly identified: positive approaches, which are based on observed behaviours such as actual travel patterns, and normative approaches, which are grounded in desired standards or policy objectives, reflecting a prescriptive view of accessibility.

A notable example of a normative approach is found in a study on public service accessibility in the Netherlands. This work examines accessibility through the lens of government interventions, financial instruments (such as subsidies and grants), and specific sectoral challenges in areas like healthcare, education, postal services, and transportation. The analysis focuses on the interplay between institutional planning tools and territorial accessibility, showing how public policies can significantly influence equity in access to essential services. (Kompil et al., 2019) In parallel, a case study set in Montréal illustrates a more empirical approach, comparing revealed accessibility, based on observed mobility data, with desired accessibility, defined by normative benchmarks. This study employs cumulative opportunity measures, which quantify the number of reachable destinations within defined spatial or temporal thresholds. The findings highlight significant discrepancies between perceived and actual accessibility, depending on individual socio-demographic profiles and urban geographic characteristics. Both studies emphasize the complexity of the accessibility concept and the need for robust, multidimensional analytical tools to properly evaluate and improve it. A common thread across these approaches is the central role played by the spatial distribution of destinations and the characteristics of the transport network: accessibility measurements can vary substantially based on territorial configurations and the transportation systems considered.

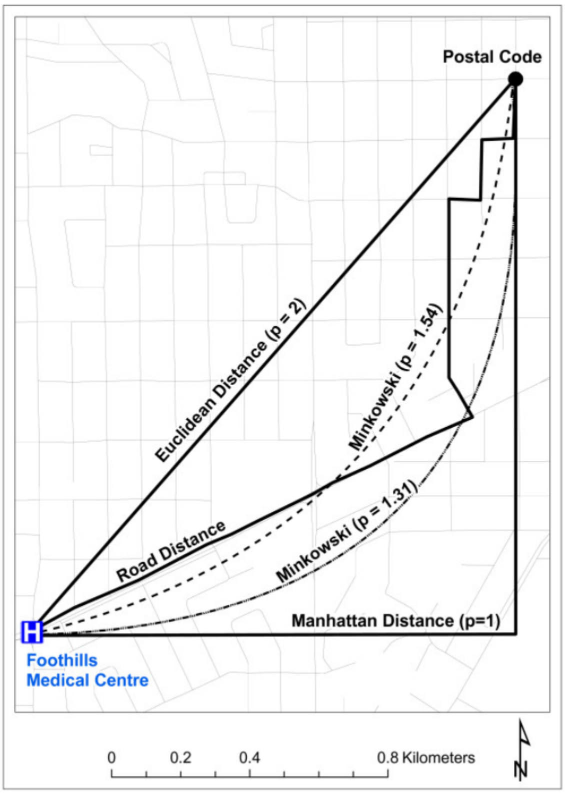


Figure 3: *Neighbourhood configurations determined by different distance metrics (Shahid et al., 2009)*

Building upon these methodological distinctions, it is important to note that accessibility is not solely interpreted as a metric tied to transport networks and spatial distances. Beyond the assessment of individual mobility and transport efficiency, several studies have emphasized the role of accessibility as a strategic tool for shaping territorial cohesion policies. (Kompil et al., 2019) In this broader perspective, accessibility becomes a key parameter in promoting balanced territorial development, guiding decisions on the optimal location of public services and ensuring equitable access for all communities, particularly in peripheral or underserved areas. This expanded understanding of accessibility integrates spatial justice concerns and aligns with policy objectives aimed at reducing regional disparities. By evaluating accessibility in relation to the distribution and planning of public infrastructure, these studies provide critical insights into how access to services can influence social inclusion, economic opportunity, and long-term regional sustainability. (Verachtert et al., 2023)

Continuing this broader policy-oriented perspective, recent research has increasingly recognized accessibility as a driver of territorial cohesion and population distribution, especially within the European context. The European Union has placed growing emphasis on projects aimed at evaluating accessibility not only as a static spatial indicator but as a dynamic component influenced by, and influencing, demographic trends.

2.2.1 Accessibility and population dynamics: The role of EU spatial planning

In particular, the EU has supported studies investigating how changes in population distribution can affect and be affected by accessibility levels. (Jacobs-Crisioni et al., 2016) (Spiekermann et al., 2015) One such initiative involved the use of the LUISA (Land-Use-based Integrated Sustainability Assessment) model, which simulates future land use and population changes based on biophysical and socio-economic drivers.

In this model, accessibility is treated not only as an outcome but also as a determinant of land use allocation, thus capturing feedback loops between territorial development and infrastructure. The study compares different scenarios:

1. The current baseline situation;
2. A scenario considering only transport infrastructure investments;
3. Two forward-looking scenarios incorporating both transport investments and future population distributions simulated by LUISA under different urbanisation patterns (Compact vs. Business-As-Usual).

To assess accessibility, a variety of indicators are employed, including location accessibility, relative network efficiency, potential accessibility, and daily accessibility, alongside inequality measures such as the coefficient of variation, Gini, Atkinson, and Theil indices.

Key findings and policy implications

The results reveal that while average accessibility levels are only marginally affected by variations in local population, the degree of territorial inequality in accessibility can change significantly. In particular:

- Compact urban development scenarios tend to reduce disparities in potential and daily accessibility, but can increase inequalities in positional accessibility, reflecting a trade-off in spatial equity outcomes;
- When future population trends are accounted for, the cohesion-enhancing effect of infrastructure investment becomes more limited, suggesting that infrastructure alone may not be sufficient to counterbalance spatial disparities without coordinated land use planning.

Toward integrated policy approaches

This study highlights the necessity of integrating demographic forecasts and spatial planning into infrastructure investment strategies. There is no universally "optimal" urban growth model for territorial cohesion: the impacts depend on which type of accessibility, e.g., potential, daily, or positional, is prioritized. Consequently, policymakers must adopt a nuanced approach, avoiding excessive urban concentration (e.g., in capital cities like Rome) and instead fostering distributed accessibility across regions to support balanced population distribution and sustainable territorial development.

2.2.2 Types of public services considered

To meaningfully assess accessibility, it is essential to define which types of public services are being considered, as their spatial distribution and functional roles vary significantly. In recent studies (UNE, 2016), especially those involving participatory planning with stakeholders, public services have been systematically classified based on two key dimensions: the frequency of use and the geographic extent of their service area. This classification typically distinguishes three categories: basic, regional, and metropolitan services.

- Basic services refer to essential, frequently used amenities required for everyday life and full participation in society. These include primary schools, childcare facilities, general practitioners, and pharmacies. Their accessibility is typically evaluated at a local level, as they are expected to be within reach of all residents on a daily basis.
- Regional services serve broader catchment areas, often catering to multiple municipalities within a region. They include secondary schools, general hospitals, assisted-living complexes, shopping centres, cultural venues, and administrative offices. A well-served region is one that ensures comprehensive and balanced access to this full set of services, thus enhancing regional functionality and cohesion.

- Metropolitan services possess a supra-regional, and often international, relevance. They include universities, major cultural institutions, and tourist attractions, which are typically accessed by a wide range of users, from local residents to international visitors, students, and entrepreneurs. These services play a critical role in positioning metropolitan areas such as Brussels or Antwerp as global hubs within the Flanders region and beyond.

This typology is crucial in accessibility analysis, as it allows planners and researchers to calibrate their evaluation frameworks according to the hierarchical importance and expected spatial reach of each service type. It also supports the formulation of tailored policy interventions, ensuring that essential services are equitably distributed while reinforcing the strategic role of higher-level functions in urban and regional development.

2.2.3 Data Sources

Following the classification of public services, a fundamental aspect in accessibility studies lies in the selection and integration of reliable data sources. The quality and granularity of input data directly influence the validity of accessibility assessments, especially in spatial planning and territorial policy contexts.

Most recent studies rely on geospatial data from platforms such as ArcGIS Information Services, including OpenStreetMap (Tahmasbi et al., 2019) and APIs like Amap (Xia et al., 2022), in combination with official datasets provided by regional public transport operators. (Verachtert et al., 2023) These sources offer essential geographic and infrastructural information, enabling researchers to map networks, estimate travel distances, and simulate mobility patterns.

However, a major research challenge lies in improving both the accuracy and the cost-efficiency of distance and time calculations to various types of public services. Recent advances have introduced new data providers and APIs capable of incorporating real-time traffic data and multimodal public transport information, leading to more realistic estimates of accessibility. (Xia et al., 2022)

From distance to travel time: A shift in methodological focus

Traditionally, accessibility has been measured through Euclidean or network-based distances between origins and destinations. Yet, several studies argue that travel time, rather than raw distance, provides a more realistic measure of accessibility, particularly in urban contexts with variable traffic conditions. (Xia et al., 2022), (Tahmasbi et al., 2019), (Verachtert et al., 2023), (Kompil et al., 2019), (Jacobs-Crisioni et al., 2023), (Mohri et al., 2021) This shift is particularly important in multi-modal analyses, where researchers assess access to transport nodes (e.g., metro stations, bus stops, train terminals) and evaluate alternative modes such as private car use versus public transit. (Tahmasbi et al., 2019), (Xia et al., 2022), (Verachtert et al., 2023)

Accounting for daily congestion patterns, service frequency, and modal interchange times is therefore essential to reflect the real-world experience of users and to avoid overestimating effective accessibility. (Xia et al., 2022)

Consistency of temporal and spatial data

A further requirement in these analyses is the temporal consistency of datasets. The geolocation of public services must align temporally with the transport network data and journey cost information, ensuring that all sources refer to the same reference period. Discrepancies in data currency may undermine the reliability of the findings, especially in rapidly evolving urban contexts. (Xia et al., 2022), (Tahmasbi et al., 2019), (Verachtert et al., 2023)

Additionally, contextual data, such as population distribution, employment density, and socioeconomic indicators, must be included to provide a comprehensive analysis of accessibility outcomes and their equity implications. (Xia et al., 2022), (Tahmasbi et al., 2019), (Verachtert et al., 2023)

Data quality, standardisation, and longitudinal analysis

Due to the high dependency on external data providers, researchers are often confronted with the uncertainty and variability of data quality. Many studies rely on datasets collected several years before the time of analysis, which may no longer accurately represent current accessibility conditions. (Tahmasbi et al., 2019), (Verachtert et al., 2023), (Kompil et al., 2019), (Jacobs-Crisioni et al., 2023) This temporal lag can limit the policy relevance of results unless regularly updated datasets or time series are used.

Consequently, it is crucial to adopt standardised access mechanisms and flexible computational frameworks that allow for the adaptation of accessibility calculations to the most cost-effective and reliable data sources available for each region. Furthermore, researchers must define evaluation protocols to assess the quality and representativeness of the data provided by each source in the specific study context.

Looking ahead, the use of time series data, ideally at an annual frequency, could significantly enhance our understanding of how accessibility evolves over time. Such temporal insights are key for assessing the impacts of infrastructure investments, land use changes, and policy interventions, as well as for identifying emerging disparities in service provision.

2.2.4 Methodological approaches to accessibility indicators

Accessibility is a multidimensional concept that integrates the spatial distribution of services with the performance of transport systems. It reflects the extent to which land use and transport networks enable individuals or goods to reach desired destinations, considering variables such as time, cost, and effort. (Geurs and Ritsema van Eck, 2001) Over time, a variety of accessibility indicators and methodological approaches have been developed to operationalize this concept, focusing on different components including infrastructure, spatial location, and user experience. (Van Wee et al., 2013)

Classification of accessibility measures

A comprehensive classification proposed in the paper "Planning for sustainable accessibility: Developing tools to aid discussion and decision-making" (Curtis and Scheurer, 2010) identifies seven major categories of accessibility indicators, each reflecting distinct methodological perspectives:

1. Spatial separation measures: These capture the degree of resistance or impedance between origins and destinations based on physical or network distance, travel time, travel cost, or transport service quality. While easy to compute using GIS and cartographic data, they do not account for land use patterns or the distribution of opportunities.

2. Contour (or cumulative opportunity) measures: Define catchment areas by travel time thresholds and count the number of accessible opportunities (e.g., jobs, schools) within each contour. While simple and intuitive, they can be arbitrarily defined and fail to capture variations within the same contour.

3. Gravity-based measures: Use continuous impedance functions (e.g., negative exponential or inverse power) to weigh opportunities by their distance or cost. They offer finer sensitivity than contour methods but are harder to interpret and do not usually distinguish between travel purposes or individual preferences.

4. Spatio-temporal measures: Consider time windows and sequences of activities, suitable for trip chaining and spatial clustering analysis. These require detailed, project-specific data and are therefore more limited in scale and transferability.

5. Utility-based measures: Evaluate the perceived or real benefit derived from accessing services. While insightful in terms of behavioural modelling, they often rely on complex assumptions and struggle to predict feedback effects between land use and transport behaviour.

6. Competence (or capacity-based) measures: Incorporate constraints in both supply and demand, enabling more realistic estimation of congestion and saturation in public services.

7. Network-based measures: Apply graph theory metrics (e.g., centrality, proximity, impedance) to assess the connectivity and efficiency of the transport network. They can capture spatial legibility and simulate structural impacts of interventions.

Travel impedance and distance decay functions

A common thread in most methods is the use of impedance functions, which quantify the effort required to reach a destination and simulate trip distribution. Two of the most widely used functions are the negative exponential and the inverse power function, whose form and parameters depend on the mode of transport and the trip purpose. When measuring accessibility by walking, cycling, driving, or public transit, it is necessary to define a separate impedance function for each mode.

The distance decay function complements this by expressing how accessibility scores decrease with increasing travel time. Stakeholders often define thresholds for what constitutes "close" or "far" for different types of services. For example, for basic services, a travel time of 7–17 minutes (1.75–4.25 km by bike at 15 km/h) is considered acceptable, aligning with standards used in European studies (e.g. Kompil et al., 2019 (Kompil et al., 2019)).

Similarly, regional services are typically located within 10–20 minutes (2.5–5 km), and metropolitan services within 14–25 minutes (3.5–6.25 km). These thresholds reflect stakeholder intentions to encourage active mobility for local access while reserving public transport for longer-range travel.

To differentiate between spatial proximity and network connectivity, some models use a dual-dimension approach, where "place" dimension measures proximity by walking or cycling, while "node" dimension reflects access to public transport. This separation allows analysts to distinguish between areas where services are physically close and those where access relies on transport infrastructure.

Marginal utility functions and service saturation

In some studies, accessibility is further refined using marginal utility functions, derived from behavioural economics. These functions recognize diminishing returns in accessibility gains: the first nearby service (e.g., a bakery or pharmacy) adds substantial value, while subsequent services add progressively less. One model assigns the following weights:

- Second service: +0.5
- Third service: +0.3
- Fourth service: +0.2

After the fourth, additional services do not increase the accessibility score. This adjustment is applied as a correction factor to the initial distance decay score and has been uniformly applied across 50 service types in certain studies.

Potential-based and floating catchment methods

Other widely used approaches include potential accessibility measures, such as the Population-to-Provider Ratio (PPR). This is particularly common in healthcare studies due to its simplicity and ease of communication. However, PPR is sensitive to scale and zoning, which can distort results. (Xia et al., 2022)

To address this, methods like GraBAM (Gravity-Based Accessibility Measure) and 2SFCA (Two-Step Floating Catchment Area) have been developed:

- GraBAM weighs supply-demand ratios using distance-decay functions.
- 2SFCA improves on GraBAM by iteratively calculating accessibility on both demand and supply sides, reducing overestimation in underserved areas.

The WATT (Weighted Average Travel Time) is another travel-cost-based metric used for comparing accessibility across modes, locations, or time periods.

Additionally, market potential indicators are often computed by aggregating population values (Pi) weighted by distance (Dij), typically excluding grid cells beyond a defined threshold.

Measuring equality of access

As accessibility plays a central role in reducing spatial and social inequality (Kompil et al., 2019), (Tahmasbi et al., 2019), several scholars have proposed methodologies to evaluate equity in service distribution:

1. Relative distribution methods, such as the Gini coefficient, Lorenz curve, and concentration indices;
2. Composite index systems, designed to capture disparities using multidimensional indicators.
3. Service equalisation frameworks, focusing on supply modes, demand patterns, user rights, and institutional responsibility.

These frameworks are instrumental for planners and policymakers aiming to ensure fair and inclusive access to essential services such as healthcare, education, and transport infrastructure.

2.2.5 Accessibility on spatial planning support systems

Over the past few decades, research on spatial planning support systems (PSS) has expanded significantly, evolving from the development and testing of individual tools to comprehensive reviews and the construction of generalized knowledge frameworks for planning support. However, despite these advancements, the actual adoption of such systems by planning practitioners remains limited. Often, these professionals are unaware of the tools available, particularly accessibility measures, or lack the experience and trust required to incorporate them into decision-making processes. (Pelzer, 2015; te Brömmelstroet, 2017) On the other side, tool developers tend to emphasize scientific rigour and technical sophistication, sometimes at the expense of usability and contextual relevance. This imbalance has contributed to a growing "black box effect", which increases the gap between what is offered and what is actually needed in practical planning contexts. (Pelzer, 2015)

Pelzer (Pelzer, 2015) identified 10 recurring usability characteristics of PSS: transparency, ease of use, interactivity, flexibility, computational time, data quality, level of detail, comprehensiveness, reliability, and communicative value. Nonetheless, both Pelzer and te Brömmelstroet argue that improving these usability attributes alone may not be sufficient to close the implementation gap. (te Brömmelstroet, 2017) More recent studies have shifted focus toward the concept of usefulness, evaluating whether a PSS can meaningfully support planning tasks, rather than merely being technically sound.

Empirical assessments of PSS usefulness are typically conducted through workshops involving planners or planning students, who engage with the tool in practical or semi-structured scenarios. These sessions are often followed by surveys or Likert-scale questionnaires that assess participants' perceptions. (Arciniegas et al., 2013; te Brömmelstroet, 2015) Emerging evaluation frameworks have introduced multidimensional criteria such as emotional reaction, problem perception, participant engagement, behavioural influence, communication effectiveness, shared language development, consensus-building, and group cohesion. (te Brömmelstroet, 2017)

Such frameworks reveal the broader socio-cognitive dimensions of PSS interaction, going beyond usability metrics.

Silva et al. (Silva et al., 2017) further explored how accessibility measures are perceived in planning practice, examining dimensions such as their role in informing debate, problem generation, problem analysis, strategy selection, implementation support, and understanding land-use/transport interactions. Their findings highlight a generally positive reception of accessibility measures, especially in fostering mutual understanding, shared language, and consensus on planning problems. However, difficulties still persist in reaching consensus over planning goals.

Moreover, attitudes toward accessibility tools are shaped by personal and organizational dimensions, including perceived professional relevance, familiarity, cultural fit within the organization, and inter-agency cooperation. (Silva et al., 2017) These results suggest that accessibility-based indicators are valuable not only for technical analysis but also for stimulating dialogue, building capacity, and aligning perspectives among stakeholders.

In conclusion, while accessibility metrics offer valuable insights for territorial analysis and planning, their effective application within this work hinges on aligning methodological robustness with practical usability. Achieving the objectives of this study, namely, to explore, compare, and operationalize accessibility indicators for supporting equitable and evidence-based spatial planning, requires not only technically sound tools, but also an awareness of the institutional and organizational contexts in which they are applied. Bridging the persistent gap between scientific rigour and the practical needs of policymakers remains essential for translating accessibility analysis into meaningful planning outcomes.

2.3 Analyses and case studies related to accessibility

Over the last few years, loads of studies and projects have looked at how easy it is to get to essential services from different places and using different methods. This work has helped to define indicators, methodologies and tools that are useful for analysing and monitoring spatial inequalities in access to services such as education, health and mobility. In this section, several significant examples of such studies will be presented and discussed, with particular focus on European and national projects that propose innovative solutions and operational tools. These include an analysis of the European ESPON PROFECY project, a survey on accessibility in the Valle Sabbia region and an analysis of accessibility conditions from suburbs to schools, with the aim of highlighting strengths, limitations and opportunities for improvement that can inform and support this work.

2.3.1 European accessibility study: EPSON PROFECY project

The ESPON PROFECY project (Processes, Features, and Cycles of Inner Peripheries in Europe) (ESPON, 2017a), (ESPON, 2017b), conducted between 2016 and 2018, aimed to identify and analyse European regions facing challenges in accessing basic services of general interest (SGIs). These regions, termed "inner peripheries", are characterized by their relative remoteness from economic and demographic centres, leading to diminished accessibility to essential services such as healthcare, education, banking, cultural amenities, and transportation hubs.

The project developed a comprehensive methodological framework to delineate inner peripheries across Europe, utilizing indicators related to accessibility, economic potential, and demographic trends. By mapping these indicators at the NUTS-3 level, PROFECY provided a nuanced understanding of the spatial distribution of inner peripheries and their specific challenges. In EU nomenclature of territorial units for statistics, NUTS-3 level indicate the smallest units in an hierarchical system and are used for specific diagnoses and socio-economic analyses of regions.

One of the key contributions of PROFECY was the identification of areas at risk of becoming inner peripheries (RIPs), enabling policymakers to implement preventive measures.

The project also offered strategic recommendations to enhance territorial cohesion, emphasizing the importance of tailored, place-based policies that address the unique needs of these regions.

By highlighting the multifaceted nature of peripheral, not solely defined by geographic distance but also by socio-economic factors, PROFECY has informed subsequent research and policy initiatives aimed at fostering inclusive and balanced regional development across Europe.

Following efforts like the PROFECY project to identify accessibility gaps and territorial disparities, a parallel line of inquiry concerns the functional roles of settlements in the provision of services. The most intuitive method to infer these roles is through the analysis of observed spatial interactions, such as commuting flows, or by evaluating actual levels of service provision across different settlements.

This approach has been explored using commuting data in several studies (e.g., ESPON, 2014; Sýkora & Mulíček, 2009), which offer insights into employment-related spatial dynamics. However, relying solely on commuting patterns may lead to biased interpretations of settlement functionality. Commuting behaviour varies considerably across labour market sizes, levels of educational attainment, and is significantly influenced by age and gender-specific participation rates (Groot et al., 2012). As such, employment-based interactions capture only a partial and uneven aspect of a settlement's role in the regional service network.

Ideally, settlement functionality should be assessed through harmonized, fine-grained, and pan-European data on the acquisition of goods and services, encompassing not only employment but also access to healthcare, education, retail, and cultural services. However, such data remains largely unavailable for academic research due to fragmentation, lack of standardization, and insufficient spatial resolution.

Even datasets that do exist, such as those published by ESPON (2017b) (ESPON, 2017b), offer only a partial picture, often constrained by:

- Limited scope of publicly available service data;
- Crowdsourced sources, which exhibit uneven spatial coverage depending on local participation levels;
- Systematic biases, including classification errors, omissions, and false positives (Johnson et al., 2016; Yeow et al., 2021)

These limitations pose a significant challenge for scholars attempting to map functional hierarchies or understand service centrality in an evidence-based way. Therefore, while policy frameworks increasingly call for place-based strategies tailored to local service dynamics, methodological innovations and data infrastructure improvements remain a necessary prerequisite for realizing such goals.

Methodological framework of the ESPON PROFECY project

The ESPON PROFECY project (Processes, Features, and Cycles of Inner Peripheries in Europe) was initiated to identify and analyse regions within Europe that face challenges in accessing essential services, despite not being geographically remote. These "inner peripheries", are characterized by their relative disconnection from economic and service centres, leading to potential socio-economic disadvantages.

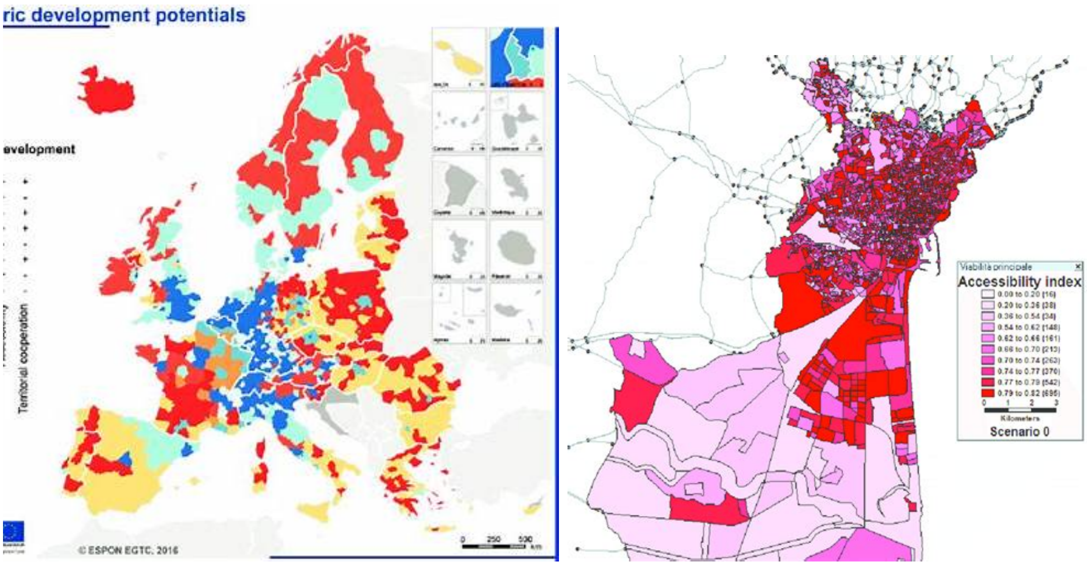
Identification of inner peripheries

The project employed a multi-dimensional approach to delineate inner peripheries, considering factors such as:

- Accessibility to regional centre's: Evaluating the ease with which populations can reach regional hubs that offer a range of services and employment opportunities;

- Availability of Services of General Interest (SGIs): Assessing the presence and accessibility of essential services, including healthcare, education, banking, and cultural amenities:
- Socio-economic performance: Analysing demographic trends, economic indicators, and employment rates to understand the vitality of regions.

By integrating these dimensions, the project identified areas that, while not geographically isolated, experience functional disconnection from broader socio-economic networks.



(a) *Overview EPSON PROFECY project report* (b) *Highlighted functionally disconnected areas*

Figure 4: *Examples of accessibility visualization in ESPON PROFECY final report*

Methodological tools and data sources

To achieve its objectives, PROFECY utilized a combination of quantitative and qualitative methods:

- Geospatial analysis: Employing Geographic Information Systems (GIS) to map and visualize accessibility patterns and service distributions across regions;
- Statistical modeling: Applying statistical techniques to assess correlations between accessibility, service availability, and socio-economic indicators;
- Case studies: Conduct in-depth analyses of selected regions to understand the unique challenges and potentials of inner peripheries.

Data sources included European statistical databases, national and regional datasets, and input from local stakeholders.

Policy Implications and Recommendations

The findings of the PROFECY project have significant implications for regional development policies.

- Targeted interventions: Identifying inner peripheries allows for the design of tailored policies aimed at improving accessibility and service provision in these areas;
- Integrated planning: Encouraging coordination between transport, service delivery, and economic development strategies to address the multifaceted challenges of inner peripheries;

- Monitoring and evaluation: Establish frameworks to continuously assess the effectiveness of policies and adapt strategies as needed of the European Commission.

By highlighting the existence and characteristics of the inner peripheries, the PROFECY project provides a foundation for more equitable and effective regional development initiatives in Europe.

In Italy, several studies and projects have been conducted to assess the accessibility of public services, such as schools and hospitals, in terms of travel time for residents. However, there are currently no specific initiatives of this kind in the Lazio region. In this section, an analysis is conducted of similar work developed at the national level and of other accessibility studies undertaken in different regions of Italy.

2.3.2 POLIS - Accessibility to educational services in Italy's internal areas

A comprehensive study conducted by Openpolis in collaboration with "Con i Bambini" highlights significant disparities in access to educational services across Italy's internal and less accessible municipalities. (Openpolis and i Bambini, 2023) The research emphasizes that transportation barriers to schools are a critical issue affecting equitable educational opportunities, particularly for children residing in remote areas.

The study reveals that approximately one in four Italian municipalities lack proximity to railway stations or highway networks, rendering them less accessible. Notably, 61% of these municipalities are classified as peripheral or ultra-peripheral, situated over 40 minutes away from the nearest service centre. These areas are home to about 389,211 minors, underscoring the scale of the challenge in ensuring equitable access to education. In terms of school accessibility, the research indicates that 88.1% of schools in these territories are connected via alternative means to private vehicles, such as public transportation. While this figure appears substantial, it falls below the national average of 89.3%, highlighting a gap in accessibility that could impact students' educational experiences.

The methodology employed in the study involves analysing data on school locations, transportation infrastructure, and demographic information to assess the accessibility of educational institutions. By examining the availability and reach of public transportation options, the study provides insights into the logistical challenges faced by students in remote areas. These findings underscore the need for targeted policy interventions to enhance transportation infrastructure and services in Italy's internal regions. Improving accessibility to educational facilities is crucial for promoting equal opportunities and addressing educational disparities among students in different geographic locations.

Visualizing spatial inequality: visualization tool

A powerful visual output from the Openpolis and "Con i Bambini" study is the spatial map that highlights accessibility disparities across Italian municipalities (see figure 5). This map uses a choropleth-style visualization to indicate the proportion of buildings connected to educational services via public or alternative transport. In red, the borders of municipalities located more than 40 minutes from the nearest service hub and not close to highways or train stations. Darker shades of blue correspond to municipalities with higher levels of connectivity, while lighter tones reveal critical gaps in service access.

The map also outlines in red the borders of municipalities that are both inaccessible, lacking nearby highways or railway stations, and located over 40 minutes from the nearest service hub, emphasizing their peripheral or ultra-peripheral classification under the National Strategy for Inner Areas (SNAI). These territories, primarily in mountainous or rural regions, experience systemic logistical challenges that compromise the right to education, especially for children without access to private transport.

This visual representation adds a spatial layer to the quantitative findings, underlining the geographic concentration of educational disadvantage and the need for territoriality policy interventions.

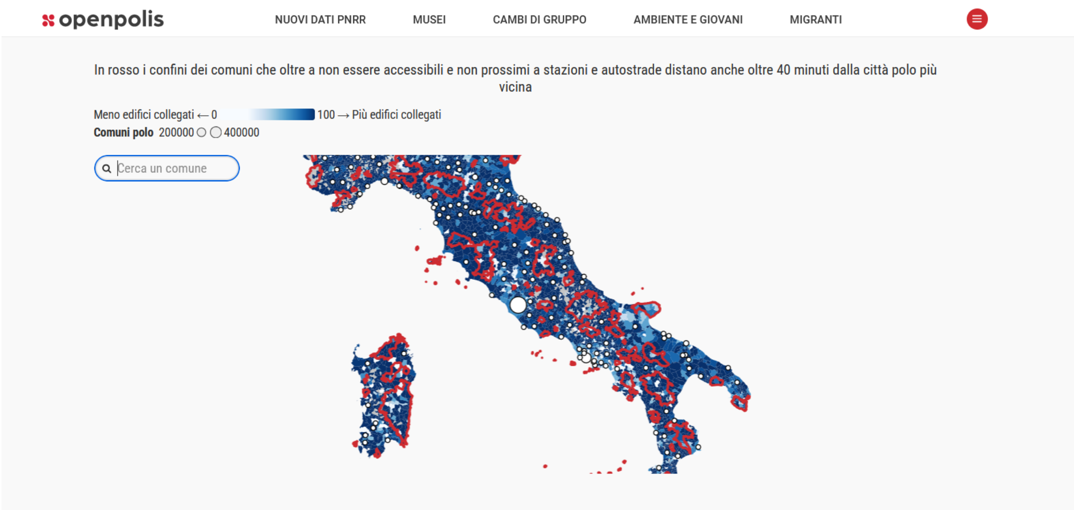


Figure 5: *Accessibility to educational buildings across Italian municipalities*

.

2.3.3 Case study: Accessibility and service provision in Valle Sabbia

A notable example of spatial accessibility analysis in Italy is the comprehensive study conducted in 2003 on the Valle Sabbia region (Brescia, Italia), entitled "Provision of public and private services and accessibility in Valle Sabbia". (Comunità Montana di Valle Sabbia, 2003) This study represents one of the earliest and most methodologically robust attempts in Italy to evaluate the distribution and accessibility of both public and private services across a mountainous and geographically fragmented area. The research underscores how topographical features, such as elevation, surface area, and road network complexity, influence human activity and access to essential services in sparsely populated communities.

The methodology integrates demographic, socio-economic, and territorial data, primarily derived from the XIV General Census of 2001 and elaborated using GIS software (Census 2000 GIS). The analysis was performed at the sub-municipal level (neighbourhood), allowing for a highly granular perspective on service provision. Key datasets included the population registry, road atlases, thematic maps, and business census records. A core feature of the methodology was the construction of Origin/Destination (O/D) matrices, which quantified distances and travel times between neighbourhood using different average speeds (2, 10, and 40 km/h). These matrices provided insight into mobility patterns and accessibility deficits, especially in areas where walking was unfeasible due to terrain or distance.

The study focuses part of the research to choose the services that are considered for the analysis, mainly them are classified into two categories: metropolitan (supra-local services like hospitals and public safety) and local (local services like pharmacies, post offices, and primary schools). It calculated theoretical catchment areas (theoretical catchment areas) based on proximity, regardless of administrative boundaries. Importantly, the analysis included the impact of opening or closing service locations, allowing for dynamic assessment of service reach and population exposure to service gaps. Specifically, a specific set of public and private services considered relevant to local communities were analysed. These services were classified into thematic groups. The services analysed in the study are sowed in table 1

Table 1: Categorization of public and private services in Valle Sabbia case study

Category	Services
Education	• Nursery schools • Kindergartens • Primary schools • Lower secondary schools • Upper secondary schools
Health and social services	• Hospitals and clinics • Medical clinics affiliated and not affiliated with the national health service • Social assistance • Pharmacies
Vehicle maintenance, trade and public facilities	• Vehicle maintenance • Food trade • Trade in other products • Public facilities (e.g. bars, restaurants)
Recreational and sporting activities	• Entertainment activities • Libraries • Sports facilities
Credit and communications	• Banks • Post offices and telecommunications
Public administration	• Public administration • Public safety

The presence of these services is indicated in tables, by municipality or district, simply indicating whether or not the service is available in that location. The study does not focus on the exact number of service points within a district, but on the existence of a “proximity service” (a nearby service).

Population groups were disaggregated by age, with particular attention to vulnerable cohorts such as the elderly (65+) and children (6–10 years old), evaluating their proximity to critical services. Moreover, the study introduced a unique socioeconomic dimension by assessing family autonomy through two indicators: the Absolute Autonomy Index (IAA) and the Relative Autonomy Index (IRA). These metrics measured the balance between income producers and consumers within households, offering a proxy for economic resilience and capacity to access services. Families with low IRA values were flagged as particularly at risk, often correlating with isolation or poverty, especially in the case of single elderly individuals or large households with young children.

Results were presented through extensive tables, graphs and static maps, detailing population density, dependency ratios, service presence by location, and accessibility metrics. Although the study lacked interactive visualization tools, it laid the groundwork for future, more advanced accessibility assessments and provided a replicable model for other mountainous or rural regions in Italy.

This case study demonstrates the relevance and feasibility of applying spatial analysis techniques to assess service accessibility in complex territorial contexts. It also highlights the necessity of integrating physical geography, demographic structure, and socioeconomic variables into accessibility evaluations to ensure equitable public service provision. As such, it serves as a valuable precedent for similar analyses in other Italian regions, including the Lazio region, where comparable accessibility assessments are still limited or fragmented.

2.4 Digital tools and technologies supporting the platform

In the previous sections, a review was conducted of studies and research efforts exploring various methodologies for assessing accessibility in spatial and service planning. As clearly emerged, data plays a pivotal role in enabling these analyses, serving as the foundation for both qualitative evaluations and quantitative measurements. However, in many practical contexts, the required data is either not publicly available, not accessible in open formats, or not updated in real time. This limitation has necessitated the use of alternative methods for data acquisition.

Among these, web scraping has gained prominence in recent years as a valuable technique for extracting information directly from web pages, particularly when datasets are not readily available through official APIs or open data portals. This section provides a general introduction to the concept of web scraping, discusses its relevance and applications in the context of accessibility analysis, and describes the tools and technologies adopted in this project. Further details on the implementation and purposes of this technique are provided in the methodology chapter. Legal considerations related to this aspect are also addressed following this paragraph.

A fundamental development that intersects with web scraping and data extraction is the concept of the Semantic Web, a transformative structure that enables the extraction and interpretation of vast amounts of data by computers in a meaningful way. Codina (2009) examines the evolution of the Web from its early iterations (Web 2.0) to the more sophisticated Web 3.0 (Codina, 2009), also known as the Semantic Web. This new paradigm is based on an innovative and complex idea: allowing computers not just to access but to "understand" and reason about web content. Central technological components developed for the Semantic Web include Uniform Resource Identifiers (URI/IRI), XML, and the Resource Description Framework (RDF). Among these, RDF is notable for its complexity and sophistication in encoding metadata, as it describes relationships between objects on the web in a way that is interpretable both by humans and machines.

The integration of Semantic Web technologies with web scraping offers promising avenues for more efficient, automated, and semantically rich data collection. This is particularly important in accessibility studies, where data heterogeneity poses challenges. By leveraging semantic structures and standards, data interoperability is improved, enabling advanced querying and reasoning capabilities that enhance the quality and usability of accessibility datasets.

2.4.1 Linked Open Data (LOD)

A core technology enabling the Semantic Web vision is the concept of Linked Open Data (LOD), which provides a framework for publishing and interlinking structured data on the web to enhance interoperability and reuse across domains (Chiarcos et al., 2013; Ullah et al., 2018). (Chiarcos et al., 2013), (Ullah et al., 2018) LOD operates on four fundamental principles:

1. Entities must be identified using globally unique Uniform Resource Identifiers (URIs);
2. These URIs should be dereference-able via standard HTTP protocols to retrieve meaningful descriptions;
3. Data should be expressed using W3C standards such as the Resource Description Framework (RDF);
4. Resources should include links to other related resources to facilitate integration and discovery.

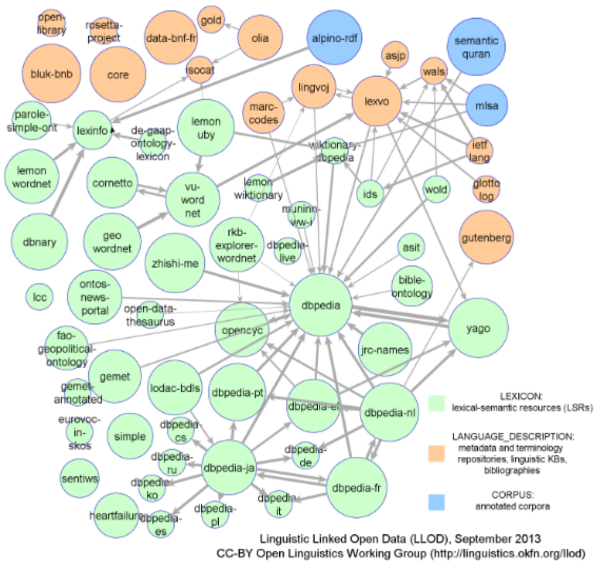


Figure 6: *RDF graph model representing Linked Open Data (Chiarcos et al., 2013)*

This framework enables the creation of rich, interconnected datasets where information is represented as triples, subject, predicate, and object, forming a directed labelled graph structure. RDF nodes, identified by URIs, allow disparate datasets hosted at different locations to reference each other, generating a dense network of linked knowledge. The open nature of LOD is further emphasized by the "5-star" rating system introduced in 2010, where the highest rating is awarded to datasets that are linked to other people's data, providing contextual richness and improving usability.

Publishing data as Linked Open Data allows resources to be globally and uniquely identifiable, accessible via standard web protocols, and uniformly interlinked. This structural interoperability is essential for complex domains like territorial accessibility, where data originates from heterogeneous sources and formats. Adopting LOD principles facilitates automated data integration, semantic querying, and enhanced decision-support capabilities, laying the groundwork for dynamic, transparent, and scalable spatial data infrastructures.

2.4.2 Data collection techniques: Web Scraping

To effectively implement web scraping techniques, it is essential to understand the core technologies behind the World Wide Web, namely, the Hypertext Transfer Protocol (HTTP), the Hyper Text Markup Language (HTML), and the Document Object Model (DOM). These components form the technical foundation upon which websites are built and through which automated tools interact with web content.

HTTP is the protocol used for transmitting hypermedia documents. It is the foundation of any data exchange on the web and a protocol used for fetching resources. HTTP follows a client-server model where through HTTP requests, resources such as HTML documents, CSS style sheets, JavaScript files, and images are retrieved. HTTP methods like GET, POST, or PUT define how clients interact with the server, and mastering these is essential for correctly acquiring web pages during a scraping process. (Ferrara et al., 2012)

HTML is the standard markup language for creating web pages. It structures the content on the web, using a series of elements represented by tags. These elements form a hierarchical structure that can be parsed and manipulated, which is essential for extracting data during web scraping. (Ferrara et al., 2012)

HTML, also, provides the content structure of a web page. It is composed of elements defined by tags (e.g., `<div>`, `<a>`, `<table>`, etc.) that can contain text, links, images, or other nested elements. These HTML elements, once loaded by a browser, are rendered into a DOM tree, which becomes the interactive, programmable structure of the page. The elements in the DOM also contains the essential data that the web scraping are interested to take off.

The Document Object Model (DOM) is a hierarchical tree representation of an HTML (or XML) document. Every element, attribute, or piece of text is represented as a node in the tree. The root node is typically the `<html>` element, which branches into `<head>` and `<body>`, and from there into deeper levels of nested elements like `<div>`, `<span>`, or `<ul>`. This structure allows programs to navigate through the document using parent-child relationships and select or modify nodes. For example, a scraper can identify all table rows within a specific table by locating a `<table>` node and then iterating over its child `<tr>` nodes. Similarly, the attributes (like class, id, or href) of each element can be inspected or filtered to extract relevant data.(López et al., 2012) Figure 7 illustrates a simplified DOM tree for a basic HTML document.

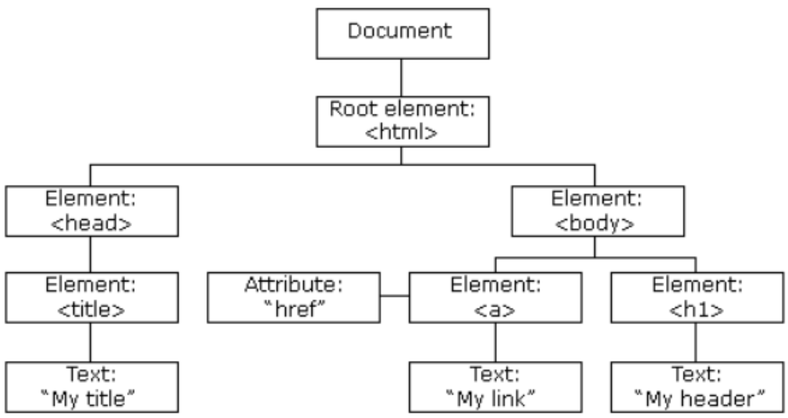


Figure 7: *Simplified structure of a DOM tree representing a basic HTML document*

In the context of web scraping, these technologies work together: HTTP is used to request web pages, HTML provides the structure of the content, and the DOM allows for the manipulation and extraction of specific data points. A comprehensive understanding of these components is essential for developing effective and efficient web scraping tools. (López et al., 2012)

Web scraping is an automated method used to extract information from websites. It has become an essential tool in data-driven research and applications, especially in contexts where structured datasets are not available via official APIs or public open data portals. Web scraping involves sending requests to web servers, retrieving the underlying HTML content, parsing the document structure, and extracting relevant data from the rendered Document Object Model (DOM). (Mitchell, 2021)

In general, the main goal of this technique, is employ to collect accessibility-related data from institutional websites that do not provide machine-readable formats. The process relies heavily on understanding and navigating the DOM structure of a page, as detailed earlier, and is implemented using a combination of powerful Python libraries: `Selenium`, `BeautifulSoup`, and `webdriver-manager`. Before implementing the logic explained below, an important analysis of the web page must be carried out to understand the exact names of the elements and create code specific to the page in question.

Selenium is a browser automation framework originally designed for testing web applications. In the context of scraping, it enables the simulation of real browser behaviour, including JavaScript rendering, drop-down interaction, and dynamic page loading, by launching a browser driver (such as Chrome Driver). This makes Selenium particularly useful for scraping JavaScript-heavy or dynamic websites that cannot be parsed through static HTML alone. (Kasliwal and Singh, 2018)

Through the use of `webdriver.Chrome()` in combination with the `webdriver_manager` module (specifically `ChromeDriverManager()`), the correct version of Chrome Driver is automatically downloaded and configured. `WebDriverWait` and `expected_conditions` allow fine-grained control over the timing of interactions, ensuring that content is fully loaded before parsing begins.

Once the page content is rendered and fully loaded, the source HTML is passed to BeautifulSoup, a lightweight Python library designed for parsing HTML and XML documents. BeautifulSoup provides a simple interface for navigating, searching, and modifying the parse tree of the document, making it ideal for extracting elements using tags, classes, or IDs. (Richardson, 2007)

While Selenium provides access to the live browser and facilitates interaction with page elements, BeautifulSoup excels in post-processing and content extraction. For example, once an HTML table is rendered in the DOM, BeautifulSoup can be used to iterate over rows (`<tr>`), extract data from each cell (`<td>`), and store the values in structured formats like lists or CSV files.

The typical scraping workflow consists of the following steps:

1. Launch ChromeDriver via Selenium and navigate to the target URL.
2. Wait until a specific element appears on the page using `WebDriverWait`.
3. Extract the full HTML content (`driver.page_source`).
4. Pass the content to BeautifulSoup for DOM parsing and element selection.
5. Extract and clean the desired data, then store it in a structured output format.
6. Repeat steps 2-5 iteratively until all data of interest has been retrieved.

This approach offers robustness against dynamic and complex layouts, and allows fine-tuned targeting of content sections relevant to accessibility data.

2.4.3 Web-based platforms for the representation of accessibility

An understanding of web page structure from an information retrieval perspective has been established. The subsection proceeds with a general overview of the technologies employed to develop web pages suitable for the publication of accessibility results.

Modern web applications are structured into distinct layers, primarily the frontend and back-end, to enhance modularity, scalability, and maintainability. The frontend represents the client-side interface that users interact with, typically developed using technologies like HTML, CSS, and JavaScript. The back-end, on the other hand, encompasses the server-side logic, handling data processing, business rules, and database interactions. This separation allows for independent development and scaling of each layer, facilitating a more organized and efficient development process (ScopicSoftware, 2023)

To further improve scalability and flexibility, many applications adopt a micro services based architecture. In this approach, the application is broken down into smaller, independent services, each responsible for a specific business functionality. These services communicate via lightweight protocols, such as HTTP/REST or message queues, allowing for independent deployment and scaling. This modular structure is in line with domain-driven design principles and supports continuous delivery and integration practices. (Dragoni et al., 2017) This micro services-based structure allows us to separate the various layers of development, from data to programming logic to the user interface.

Additionally, the Back-end for Frontend (BFF) pattern is employed to tailor back-end services to the specific needs of different client interfaces, such as web or mobile applications. By creating separate back-end services for each frontend, this pattern optimizes performance and user experience, ensuring that each client receives only the data it requires in the appropriate format. (Microsoft, 2023)

JavaScript Architecture with Vite, NPM, and OpenLayers

The implementation of the web application presented, studied, and considered in this thesis adopts a modular and modern front-end architecture based on JavaScript, leveraging a set of contemporary tools to ensure performance, scalability, and maintainability. The core technologies include NPM (Node Package Manager) for dependency management, Vite as a build tool and development server, and OpenLayers as the primary library for interactive map rendering.

This type of applications are developed in vanilla JavaScript (ES6+), allowing direct control over the DOM and application logic. This decision favours lightweight execution and a clear understanding of the interactions between components, especially useful in the context of a customized geographic user interface. JavaScript asynchronous capabilities (through Promises and async / await) are crucial for handling data fetch operations and ensuring responsive user interactions, particularly when loading dynamic content such as geographic data layers or accessibility indicators. (Flanagan, 2020)

To manage external libraries and maintain a clean project structure, the application uses NPM, the standard package manager for JavaScript ecosystems. NPM facilitates the installation, versioning, and integration of external modules such as OpenLayers (for map rendering) and utility libraries for DOM manipulation or UI control. This modularity supports a more maintainable and extensible architecture.

For development and bundling, the project uses Vite, a modern frontend tooling solution that offers faster startup and hot-module replacement during development. Unlike traditional bundlers, Vite serves source files over native ES modules and performs on-demand compilation, dramatically improving the development experience. During production builds, Vite leverages Rollup to generate optimized bundles, improving load times and performance. (You and Contributors, 2023)

The mapping functionality is powered by OpenLayers, an open-source JavaScript library designed for displaying dynamic map content in web applications. OpenLayers provides extensive tools for working with raster and vector data, tile services (e.g., WMTS, WMS), custom layers, feature overlays, and coordinate transformation. The library supports interaction with various geospatial data formats (GeoJSON, KML, GML), making it ideal for visualizing accessibility-related spatial data. The application initializes and manages map views, dynamically adds layers based on user input, and supports interactivity such as feature selection, tooltips, and filtering. (Contributors, 2023a)

This combination of technologies results in a component-based architecture where:

- JavaScript controls business logic and user interaction.
- OpenLayers handles geospatial visualization and spatial queries.
- Vite accelerates development and builds with fast refresh and modular bundling.
- NPM manages external libraries and dependencies cleanly and reproducibly.

Altogether, this architecture allows for the development of a responsive, interactive web interface capable of integrating complex geospatial data with accessibility metrics.

2.4.4 Useful tools for accessibility mapping

The implementation of modern digital platforms for geospatial data analysis and visualization often relies on a combination of tools and libraries designed for data processing, spatial computation, mapping, and interface development. This section outlines some of the most widely adopted technologies in the domain, providing a foundational understanding of their purpose and capabilities.

Visual Studio Code and Python for data processing

Python is a high-level programming language particularly suited for data science, automation, and back-end development due to its readability, extensibility, and strong ecosystem of libraries. For data cleaning and manipulation, libraries like pandas and NumPy offer efficient handling of structured data and numerical operations. Spatial extensions such as GeoPandas and Shapely further enhance Python's capabilities by enabling operations on geographic objects, including buffering, intersection, and coordinate transformations.

Visual Studio Code (VS Code) serves as a widely adopted integrated development environment (IDE) for Python and other languages. Its support for extensions, real-time linting, and debugging tools make it a convenient platform for developing, testing, and managing code in data-intensive workflows. (Flanagan, 2020)

Distance calculation: Google Maps APIs, OpenStreetMaps, OpenTripPlanner

When assessing accessibility or mobility, estimating distances and travel times is crucial. Services such as the Google Maps Distance Matrix API provide programmatic access to distance and duration calculations between multiple origins and destinations, taking into account different travel modes (e.g., driving, walking, public transport). The Geocoding API, on the other hand, converts human-readable addresses into geographic coordinates and vice versa. (Developers, 2023)

For open-source or multimodal transport scenarios, OpenTripPlanner (OTP) is often used. OTP is a routing engine that integrates OpenStreetMap (OSM) for road and pedestrian network data and GTFS (General Transit Feed Specification) datasets for public transport schedules. GTFS is a standardized format developed to describe transit agency data, including stops, routes, timetables, and frequencies, enabling realistic routing and accessibility computations in multimodal contexts. (Google and Developers, 2023; Contributors, 2023b)

Geographic Information Systems: QGIS and Shape files

QGIS is a free and open-source desktop Geographic Information System used to view, edit, and analyse spatial data. It supports a wide range of vector and raster formats and provides tools for spatial analysis, map styling, coordinate transformation, and data merging. QGIS is frequently employed in academic, governmental, and environmental sectors for decision support and territorial analysis. (QGIS, 2023)

A common format handled in GIS environments is the shapefile, a vector data storage format originally developed by ESRI. Shape files consist of several interrelated files that store geometry (points, lines, polygons) and associated attribute data in tabular form. They are widely used for representing administrative boundaries, road networks, and spatial distributions of services, and remain a standard in many spatial data repositories.

2.5 Socio-economic and regulatory context

This section discusses the socio-economic, regulatory, and legal implications in accordance with the study carried out and its potential contribution to both citizens and institutions.

From a social perspective, the project offers significant value by providing an intuitive, user-friendly tool that enables individuals to make more informed decisions regarding their daily lives. For example, users can evaluate different residential options based on estimated travel times to essential public services, not just at the municipal level, but also across different sub-zones within a municipality. This localized granularity promotes greater equity in spatial awareness and empowers citizens with accessible, actionable geographic information. On the institutional side, the platform offers strategic insights for public administrations, particularly regional planning bodies such as those of the Lazio Region. By visualizing accessibility gaps and spatial disparities in service provision, policymakers are better equipped to identify critical issues within the territory and design targeted interventions. This capacity for data-driven territorial planning enhances the ability of institutions to allocate resources more efficiently and develop equitable urban policies, aligned with the real needs of underserved communities.

Furthermore, the project contributes to the broader data economy, especially through the generation and dissemination of high-value open data. The underlying infrastructure integrates heterogeneous datasets from multiple providers and transforms them into a harmonized, reusable format, accessible to both public and private actors. This aligns with European strategies promoting open data as a driver of innovation, transparency, and economic development. A recent study by the European Commission on the economic impact of open data in Europe highlights the growing value of the open data market in terms of employment, efficiency gains, and cross-sectoral innovation. The report emphasizes how open data can enable time savings, reduce environmental impact, enhance public service delivery, and support informed decision-making across sectors. (Capgemini Invent, 2020) Notably, it includes concrete examples of European institutions leveraging open data to improve service quality, increase operational efficiency, and foster citizen engagement.

Beyond its social and data-related dimensions, the project also has potential for considerable economic impact. One of the most direct benefits is the potential for greater efficiency in the allocation of public resources. By identifying underserved areas with limited access to essential services, regional and municipal governments can optimize infrastructure investments, reduce duplication of efforts, and better target interventions where they are most needed.

Furthermore, enhanced accessibility data can contribute to the real estate and urban development sectors. Investors, planners, and developers can use spatial indicators to identify high-potential zones for residential or commercial development, especially in the context of smart city initiatives and sustainable mobility planning.

On an individual level, improved accessibility insights may lead to time savings for citizens, which in turn can increase overall productivity. Reduced commute times not only improve quality of life but also generate positive externalities for the local economy through increased economic participation and labour market flexibility.

Finally, by releasing data as open and interoperable resources, the project contributes to the economic value of open data ecosystems. According to the European Data Portal, the open data market in the EU was valued at €184 billion in 2020, with projections of further growth through the development of new products and services, cost savings in public administration, and increased market efficiency. (Capgemini Invent, 2020) Table 2 summarizes the main economic indicators associated with open data usage in the European Union, based on projections from official studies. (Capgemini Invent, 2020)

Table 2: Estimated socio-economic impact of open data in the European Union

Indicator	Value (2019)	Projected (2025)
Estimated open data market size	€184 billion	€199.5–334.2 billion
Jobs enabled by open data	884,000	1.97 million
Potential time saved on EU roads	–	629 million hours
Potential energy consumption reduction	–	16%
Sectors with high impact	Transport, Environment, Health, Public Services	Expanded use across all sectors

Legal framework for public data reuse

Turning to the legal perspective, it is essential to frame the technological development of the platform within the appropriate regulatory context, particularly given the nature of the data involved.

In an increasingly data-driven society, where digital platforms interact with huge amounts of publicly available and georeferenced data, it is essential to place any technological development within an appropriate regulatory and legal framework. Respecting legal principles is not only a matter of compliance, but a fundamental requirement for ensuring transparency, accountability, anonymization and the responsible use of information, particularly when the data involved originates from public institutions or could impact individuals indirectly.

This section outlines the main regulatory dimensions relevant to data-centric applications, especially those dealing with territorial and accessibility information. While the platform described in this project does not handle personal data directly, the reuse of public data, the potential for user interaction, and the inclusion of geospatial datasets raise important legal considerations that must be acknowledged.

The legal overview begins with the regulation of public sector information (PSI). In recent years, the European Union and national governments have promoted policies to increase the availability and reusability of public data for social and economic benefit. However, access to such data is still subject to a series of legal obligations concerning licensing, attribution, and data integrity. Understanding these requirements is key for developing open-access platforms that comply with legislation while fostering innovation.

Legal frameworks are not static; they evolve alongside technology. For this reason, it is vital that projects involving open data, public services, or spatial analytics remain adaptable and anticipate legal implications even when not immediately apparent. This approach ensures ethical consistency, prepares for future developments, and reinforces the legitimacy of the technological solutions proposed.

In the European Union, the reuse of public sector information (PSI) is governed by Directive (EU) 2019/1024 (European Union, 2019), commonly referred to as the Open Data Directive. Adopted on 20 June 2019 and entering into force on 16 July 2019, this directive replaced the previous PSI Directive (Directive 2003/98/EC) and its amendment (Directive 2013/37/EU), establishing a comprehensive legal framework for the availability and reuse of public sector data across Member States.

The Open Data Directive aims to promote the use of public sector data to foster innovation, economic growth, and transparency. It sets out rules to facilitate the availability of public sector information in a manner that is open, accessible, and reusable, thereby contributing to the development of a data-driven economy within the EU. The directive applies to documents held by public sector bodies at national, regional, and local levels, including ministries, state agencies, and municipalities. It also extends to public undertakings operating in sectors such as transport, energy, and postal services, as well as to research data resulting from public funding.

The directive mandates that public sector bodies make their documents available for reuse by default, unless specific restrictions apply. This principle ensures that data is proactively published in formats that are machine-readable and accessible, facilitating its reuse.

A significant innovation introduced by the directive is the identification of High-Value Datasets (HVDs). These datasets are characterized by their potential to generate significant socioeconomic benefits and are required to be made available free of charge, in machine-readable formats, and via Application Programming Interfaces (APIs). The thematic categories for HVDs include:

- Geospatial
- Earth observation and environment
- Meteorological
- Statistics
- Companies and company ownership
- Mobility

The European Commission is tasked with adopting implementing acts to define the specific datasets within these categories.

To remove barriers to data reuse, the directive stipulates that the reuse of documents should be free of charge. However, exceptions are allowed where charges are necessary to cover the marginal costs of reproduction, provision, and dissemination. For certain public undertakings and cultural institutions, charges may also include a reasonable return on investment.

The directive encourages the use of standard open licenses that are easy to understand and facilitate data reuse. It also requires public sector bodies to be transparent about the conditions applicable to data reuse, including any charges and licensing terms.

To ensure fair competition, the directive prohibits exclusive arrangements for the reuse of public sector information, except where such arrangements are necessary for the provision of a service of general interest. Any such exceptions must be transparent and subject to periodic review.

Member States were required to transpose the Open Data Directive into national law by 17 July 2021. The directive is expected to have a significant impact on the availability of public sector data, promoting its reuse in various sectors, including transportation, environmental monitoring, and urban planning. By facilitating access to high-value datasets, the directive aims to stimulate the development of new products and services, enhance transparency, and support policymaking.

In summary, the Open Data Directive establishes a robust legal framework for the reuse of public sector information in the EU. Its provisions are designed to unlock the potential of public data, drive innovation, and contribute to the growth of the digital economy. Compliance with this directive is essential for any project or initiative that involves the use of public sector data within the European Union including the platform presented in this study, which relies on open and reusable public data to support equitable territorial planning.

2.6 Remarks from the literature and how to go beyond

Today, regions face major environmental, social, and economic challenges, including traffic congestion, insufficient public services and infrastructure, and limited mobility and access to essential services. To meet these challenges, regions implement sustainable territorial cohesion policies, focus to maintain economic productivity, improve quality of life and address other environmental and social problems. The spatial distribution of activities, the efficient use of resources, and accessibility to different services and facilities are crucial to promote more sustainable regions. One of the ultimate goals of urban planning is to make cities more sustainable and thus liveable and attractive. This makes accessibility to essential services a key factor in determining a city's liveability and attractiveness. Batty et al. (2012p. 481–482) identify several challenges, including *"ensuring greater and more effective mobility and access to opportunities for urban populations"*. Access to opportunity is intrinsically linked to the accessibility of services such as education, health care, and other social and public services. Among the main benefits of smart and sustainable cities are the improvement of participation, equity, justice, security, mobility, and accessibility.. This again emphasises that accessibility is considered a key benefit of integrating smart and sustainable approaches in urban development. Finally, among the main future urban planning practices and emerging scientific and technological trends, urban areas associated with sustainability dimensions, including mobility and accessibility, are mentioned. (Bibri and Krogstie, 2017)

In this context, the debate on how to effectively prioritise accessibility over mere mobility in transport planning becomes crucial. Ferreira, Beukers, and te Brömmelstroet (2012) critically examine this issue, highlighting the limitations of conventional cost-benefit analysis (CBA) in transport policy in Dutch transport planning, arguing that, the prevailing approach disproportionately prioritises mobility over accessibility. They contend that this emphasis on mobility stems from a narrow interpretation of its benefits, which ultimately skews decision-making in favour of projects that enhance movement rather than improving access to key destinations. The authors propose a paradigm shift towards accessibility, highlighting its greater societal value in fostering economic competitiveness and sustainable development. Through a hypothetical case study, they illustrate how reorienting transport policy towards accessibility can lead to more effective and equitable outcomes. Their insights contribute to the ongoing debate on how transport evaluation frameworks can better reflect broader social and economic objectives. This perspective aligns with broader trends in European transport and urban planning policies, which increasingly focus on improving accessibility rather than simply increasing mobility. Common approaches across regions include integrating transport and land-use planning, promoting mixed-use development near public transport hubs, and encouraging active mobility through walking and cycling initiatives. These strategies reflect a growing recognition that sustainable and efficient transport systems should prioritise accessibility to services and opportunities rather than merely facilitating greater traffic flows. (Ferreira et al., 2012)

Building upon the growing recognition that accessibility plays a fundamental role in sustainable urban and regional development, various international and supranational frameworks have placed increasing emphasis on accessibility-driven policies. In parallel with the UN Sustainable Development Goals (SDGs), the European Union, within the framework of the Urban Agenda for the EU (Parra-Domínguez et al., 2022), collaborates with cities and regions to develop comprehensive sustainable mobility strategies. These policies extend beyond the urban scale to encompass regional and county-level transport planning, ensuring efficient public transport systems, promoting active mobility solutions, such as walking and cycling, and enhancing accessibility to essential services for both residents and commuters. In this context, the integration of sustainability information in urban planning and smart city initiatives has gained increasing attention. Parra-Domínguez, López-Blanco, and Pinto-Santos (2022) highlight the importance of advanced information systems in monitoring sustainability indicators, particularly in relation to the SDGs. Their study underscores how smart cities leverage technological solutions to collect, process, and present sustainability data, enabling more effective decision-making in mobility and accessibility planning. By incorporating such data-driven approaches, policymakers can develop more inclusive and sustainable urban environments, ensuring that accessibility remains a central pillar of future transport policies.

Accessibility can be broadly defined as the extent to which relevant destinations can be reached using available means of transport. It is increasingly regarded as a key policy objective in land-use planning and has been promoted as one of the most relevant criteria in policy assessments. Improving access to essential services is a fundamental aspect of spatial planning in Europe, as it can help reduce social and territorial disparities while enhancing citizens' quality of life. (Kompil et al., 2019) A market-potential-based approach to accessibility assessment, as proposed by Kompil et al. (2019), provides valuable insights into how accessibility to services can varies across different regions. In general, their study highlights that disparities in access to healthcare, education, and other essential services persist across Europe, particularly between urban and rural areas. However, one of the major challenges in assessing accessibility remains the lack of comprehensive and up-to-date data on the location and availability of services and the possibility to keep these accessibility evaluations updated during the years. This data gap complicates the evaluation of policy impacts on service distribution and accessibility, underscoring the need for improved data collection and monitoring mechanisms to support evidence-based decision-making in urban and regional planning. (Kompil et al., 2019)

However, addressing these accessibility disparities and effectively assessing the impact of spatial policies requires reliable and harmonised data sources. For the evaluation of territorial cohesion policies, it is essential to integrate and standardise data from various providers. Furthermore, to ensure an accurate periodic assessment of the effectiveness of implemented policies, these sources should be updated annually, at a frequency higher than that provided by some existing research frameworks, such as PROFECY. (ESPON, 2017a,b) The absence of such mechanisms complicates the work of policymakers and researchers responsible for formulating and evaluating territorial cohesion policies, limiting their ability to make informed, evidence-based decisions.

A critical factor in promoting evidence-based territorial cohesion policies is the availability of integrated and accessible datasets that consolidate information from multiple, often heterogeneous sources. While many institutions publish spatial and statistical data in open formats such as CSV, Shape files, or Web Map Services, these datasets are frequently fragmented, non-interoperable, and built around locally defined structures and identifiers. As a result, combining them to generate actionable knowledge often requires extensive manual preprocessing, which hinders timely decision-making and the ability to perform continuous monitoring.

As highlighted by Tim Berners-Lee (Berners-Lee, 2006) in his five-star open data model, the transition from simply publishing data to creating semantically linked and machine-readable datasets is essential for unlocking their full value. According to this framework, datasets that are openly licensed, available in structured formats, and linked to other datasets using standard identifiers provide the highest level of accessibility and utility. In the context of territorial analysis, such standards enable automated data integration and reasoning across domains, facilitating deeper insights into complex spatial phenomena such as accessibility to essential services.

In conclusion, it can be highlighted that an effective and accurate understanding of accessibility must be integrated into a broader perspective, one that goes beyond the concept of distance and transport network evaluation, and also incorporates aspects such as environmental and territorial sustainability, territorial cohesion, and regional development. In this context, it should serve as a strategic tool to help cities move from car-centric mobility and land use systems to more inclusive and multimodal transport strategies, or even to assess with awareness whether it is necessary to improve road or transport infrastructure or expand the range of public services. This involves not only identifying and promoting the strengths of non-motorised and public transport systems, but also highlighting their spatial and political weaknesses, enabling targeted interventions. Accessibility analysis must therefore be accurate and meaningful, based on verifiable and freely accessible data, and representative of the multiple dimensions that define accessibility in the real world.

As demonstrated by the literature, research and case studies examined in this chapter, accessibility is inherently multidimensional. It cannot be understood solely through geometric distance or travel time, but must be interpreted through economic, demographic, infrastructural and institutional lenses. A meaningful metric of accessibility must therefore integrate different perspectives and data sources to reflect the complex reality faced by citizens, planners and policymakers.

To ensure that platforms meet these requirements, usability, transparency and policy relevance, including administrative constraints are key aspects. Particular attention should been paid to the readability and interpretability of the tool for institutional users, in particular land-use and transport planners, through iterative validation and feedback collection, which can enable to improve the visualisation of results in terms of accessibility. The interactive platform, rather than a static report, should serve not only as a visualisation interface, but also as a decision support system, capable of highlighting gaps in services, providing information on territorial cohesion strategies and identifying the potential impacts of infrastructure or demographic changes. So far, platform do not include the above characteristics that will be the focus of this research, to go beyond the current limitations, as detailed on the next chapters.

3. Objectives and Methodology

The Lazio region, home to Rome, the capital and most populous city of Italy, presents a unique context for accessibility analysis. Rome's significant attractiveness, driven by a high concentration of public and private services, cultural amenities, and economic opportunities, draws a large influx of residents and commuters. This centralisation creates intense demand pressures that manifest in challenges such as rising housing costs and congestion. Consequently, understanding and addressing accessibility deficits in the wider Lazio region is crucial. An effective accessibility analysis can reveal spatial disparities and identify underserved areas, providing critical insights to promote a more balanced population distribution across the region and to ensure equitable access to essential services for all residents.

To this end, this thesis aims to contribute to the broader debate on accessibility by applying a fine-grained methodology to the Lazio region, with a particular focus on small and medium-sized municipalities. Building on the conceptual and policy frameworks outlined below, the work provides an operational approach to assessing disparities in access to essential services, while also exploring the potential for comparative analysis with other European regions such as the Community of Madrid.

Therefore, in response to the challenges and needs outlined above, this study sets out to develop a comprehensive and operational framework for assessing accessibility to public services at the sub-municipal level. The following objectives guide the research and structure the implementation of the proposed methodology. To achieve this overall objective, the project defines the following specific sub-objectives:

1. **Service classification and zoning framework**
   - Catalogue public services according to the standards UNI ISO 37123 (ISO - UNI Ente Italiano di Normazione, a) and UNI ISO 37155 (ISO - UNI Ente Italiano di Normazione, b), based on their frequency of use (frequent/local, occasional/sub-regional, and punctual/regional).
   - Define a zoning scheme (sub-regional, municipal, urban) for the assessment of service accessibility, aligned with regional decision-making needs and with standardised European frameworks to ensure comparability.
2. **Distance computation and multimodal comparison**
   - Accurately measure travel distances and times from each urban centres to essential services:
     - By private vehicle (considering real-world traffic conditions);
     - By public transport.
   - Using and evaluate different tools and technologies for calculating distances and travel times to services:
     - Google Maps API
     - OpenStreetMap
     - ArcGIS Network Analyst
     - OpenTripPlanner with OSM and GTFS data
   - Quantify and compare differences between the two transport modes, highlighting areas with critical public transport gaps.
3. **Accessibility indicator development**
   - Develop an indicator that summarises the accessibility conditions of each urban centre, which is:
     - Easy to calculate, update and interpret;
     - Sensitive to changes in population distribution, transport and infrastructure;

    - Able to handle public/private ownership and different service scales (local to regional).
  - The indicator provides a concrete measure of accessibility, expressed in terms of both distance (kilometres) and travel time (minutes), calculated at the level of each urban centre and subsequently aggregated at the municipal level.

4. **Data integration and interoperability**
   - Collect and harmonise heterogeneous data sources (demographics, service locations, roads, public transport, territorial units and accessibility data) into a coherent and updatable spatial database.
   - Construct a structured and interoperable dataset by integrating heterogeneous sources: population data, service locations, road networks, public transport and driving distances, and territorial units.

5. **Spatial analysis**
   - Identify intra-municipal disparities in accessibility, comparing urban centres within the same municipality.
     - *Which urban centres exhibit the best accessibility for each analysed service, and is this also reflected at the municipality level?*
     - *Which urban centres exhibit the worst accessibility for each analysed service, and is this condition also observable at the municipality level?*
     - *Which municipalities show the greatest intra-municipal disparities in accessibility, both in terms of time and distance, between their urban centres?*
   - Analyse accessibility patterns across the different provinces of the Lazio region, identifying territorial inequalities and structural gaps.
     - *How do the provinces rank, in terms of accessibility, for each analysed service?*
   - Quantify and compare differences between the two transport modes (private vehicle and public transport), highlighting areas with critical public transport gaps.
     - *What are the differences in accessibility, both in terms of distance and time, in the province of Rome?*
     - *Are the differences between car access (considering traffic) and public transport access to a given service consistent across the region, or are there significant outliers?*
   - Comparative analysis and benchmarking between the results obtained by the Community of Madrid and those of the Lazio region
     - *Is there a correlation in accessibility between municipalities with similar population densities?*
     - *What is the relationship between the most and least accessible services in the two case studies (Community of Madrid and Lazio)?*

6. **Stakeholder-oriented design and data visualisation**
   - Define an evaluation scheme that considers the roles and needs of all stakeholders involved in data production, publication, and policy use.
   - Develop an user-friendly web application with an interactive map that enables both citizens and policymakers to visualise accessibility indicators and compare different areas intuitively.

The chapter continue to describe how these objectives have been pursued through the design of the data infrastructure, the selection and implementation of the methodology, and the development of the interactive platform, describing in detail the methodological approach adopted for analysing accessibility to public services in the study area.

The entire process was divided into several phases, each of which played a fundamental role in building a solid data infrastructure and meaningful, interpretable indicators.

The first step was data collection, which involved integrating heterogeneous sources, such as geographic datasets, GTFS public transport data, and public services location, carefully selected to ensure spatial and temporal consistency. Next, the process of data preparation and transformation is described, which is necessary to obtain origin-destination matrices (OD matrices) consistent with the spatial structure adopted. In addition, the method used to retrieve data from other sources when not obtained from Google Maps will be shown, along with the conclusions reached.

A specific section is dedicated to the concrete creation of accessibility indicators, which required the aggregation and synthesis of information on distances and travel times between areas and points of interest. This section also explains the method used to impute missing distance data that could not be obtained in the previous phase. The indicators were then subjected to an interpretative analysis to assess their meaning, distribution and territorial implications.

Finally, the last parts of the methodology explains the methods to build a tool to visualize and publish the results, with the aim of making the data accessible and easily understandable even to public decision-makers and non-technical stakeholders, while ensuring the transparency and applicability of the entire process. The data publication process is developed in accordance with the European open data standards. Figure 8 aims to clarify the stages of the methodology, which are detailed below.

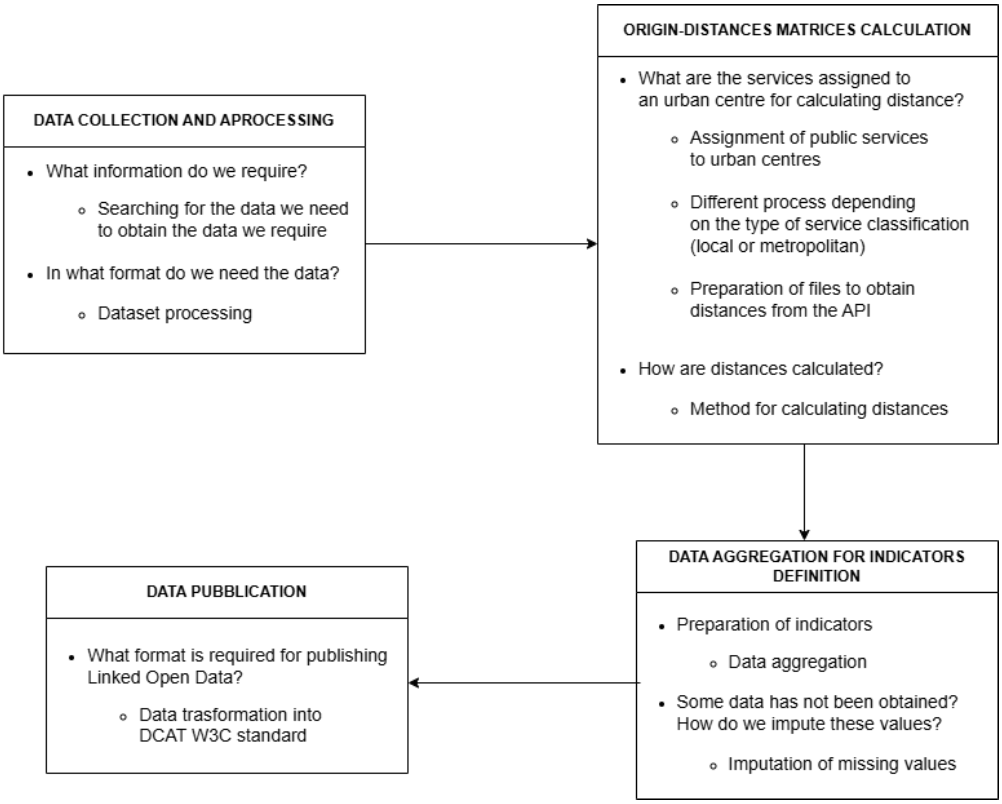


Figure 8: *Diagram explain the methodology phases*

Before delving into the details of the methodology adopted, it is important to clearly reiterate the final objectives to be achieved through the application of the methods that are about to be described. The main objective of this work is to construct an accessibility indicator that quantifies the degree of accessibility of each urban centre in the Lazio region to public services of interest. In particular, the accessibility indicator is understood as the ease of reaching, in terms of both time and distance, the public services considered, which in this project are currently limited to the education and health sectors. The classification of services was carried out according to the criteria proposed by Kompil (Kompil et al., 2019) and Verrachert (Verachtert et al., 2023), which distinguish between local and metropolitan services.

Category	Sector	Type of service
Local services	Education	Kindergarten Primary school Secondary school
	Health	General practitioners
Metropolitan Services	Health	Hospital facilities Emergency services

Table 3: Classification of public services considered for the calculation of accessibility indicators

Table 3 shows the categorisation adopted, highlighting the specific types of public services considered for each category and sector. For each type of service, the aim is to calculate an accessibility indicator expressed in terms of both time (travel time) and space (distance in kilometres), starting from urban centres and, in addition, from municipal boundaries. The accessibility indicator for each public service will be calculated considering two main modes of transport: by car and by public transport. This approach allows accessibility to be assessed on the basis of two of the main means of transport used to reach services, providing a more comprehensive and realistic view of the actual availability of services in the area.

3.1 Data collection

The first phase of the work involved researching and collecting the data needed to calculate accessibility indicators from population centres in Lazio to the main public services, with a particular focus on education and healthcare. The data were obtained from official sources, in accordance with European and Italian legislation on access to public data, and came from institutional websites or recognised national statistical bodies.

The raw data necessary and used to achieve the research objectives are, as follows:

1. Geometry of urban centres in the Lazio region: used to define the origin points for accessibility calculations and routing analyses.
2. Geometry of municipalities in the Lazio region: required for the territorial assignment of services and for spatial aggregations at the municipal level.
3. Road network of the Lazio region: necessary to support routing operations and to reposition centroids that fall outside urban geometries, ensuring realistic travel paths.
4. Population data for each urban centre and municipality: used for population weighted aggregations and to meet project specific criteria, such as evaluating accessibility in municipalities with fewer than 40,000 inhabitants.
5. Location data for schools in the Lazio region, including school category: serves as destination points in the OD matrix for assessing access to local educational services.
6. Location data for family doctors (general medicine) providing outpatient services: included as another category of local service for OD matrix computation.
7. Location data for hospitals in the Lazio region: considered as destination points for metropolitan-level services in the OD analysis.
8. Location data for emergency services in the Lazio region: similarly used to evaluate access to urgent care facilities at the metropolitan scale.

9. GTFS data of public transport providers in Rome: used to validate public transport availability and to complement missing data for transit-based accessibility calculations.

This initial section provides a basic explanation of the data collection process and includes a brief description of its structure and format.

In order to carry out the analysis, it was necessary to gather heterogeneous data from various official sources. The following main data sources were selected: the website of Italian National Institute of Statistics (ISTAT), OpenStreetMaps, the portal of the Italian Ministry of Education and Merit (MIUR), the portal of the regional health system of the Lazio Region (ASL LAZIO), the National Agency for Regional Health Services (PNE) and the official page of the administration of open data of Rome (OPEN DATA ROMA). The sources present datasets in different formats and not all data can be downloaded immediately or are structured in a uniform manner, making pre-processing and normalisation necessary.

Table 4: Sources, formats and update frequencies of datasets used for accessibility analysis

	Dataset	Source	Format	Update_frequency
1	Geometry of urban centres in the Lazio region	ISTAT (ISTAT, 2021)	Shapefile (.shp)	Every 10 years
2	Geometry of municipalities in the Lazio region	ISTAT (ISTAT, 2024a)	Shapefile (.shp)	Yearly
3	Road network of the Lazio region	OpenStreetMaps (Data/Maps Copyright 2018 Geofabrik GmbH and OpenStreetMap Contributors, 2025)	Shapefile (.shp)	Continuously updated
4	Population of each urban centre and municipality	ISTAT (ISTAT, 2024b)	CSV	Yearly
5	Location of schools with category	MIUR (MIUR, 2024)	CSV	Yearly
6	Location of general practitioners	ASL LAZIO (ASL LAZIO, 2025a)	WEB	Varies by ASL
7	Location of hospitals	PNE (PNE - Esplora Strutture, 2025)	CSV	Yearly
8	Location of emergency services	ASL LAZIO (ASL LAZIO, 2025b)	CSV	Varies by ASL
9	GTFS data of public transport companies in Roma	OPEN DATA ROMA (Google and Developers, 2023)	ZIP	Monthly

Table 4 summarises the sources and formats for each dataset needed to achieve the project objectives, numbered according to the list of data types listed above. Some data are not available in structured or easily downloadable formats. In these cases, it was necessary to use web scraping techniques to extract the information directly from the official web pages. The purpose and reason for choosing this data will be explained in detail later on.

The objective of this first phase of the methodology is to build a solid and consistent data infrastructure containing all the information necessary for generating the accessibility indicators listed in table 3. A detailed description is provided below of the characteristics of the data from the various sources, the methods used for their acquisition, and the pre-processing operations carried out to adapt and prepare them for application in the subsequent methodological phases.

3.1.1 Geospatial data

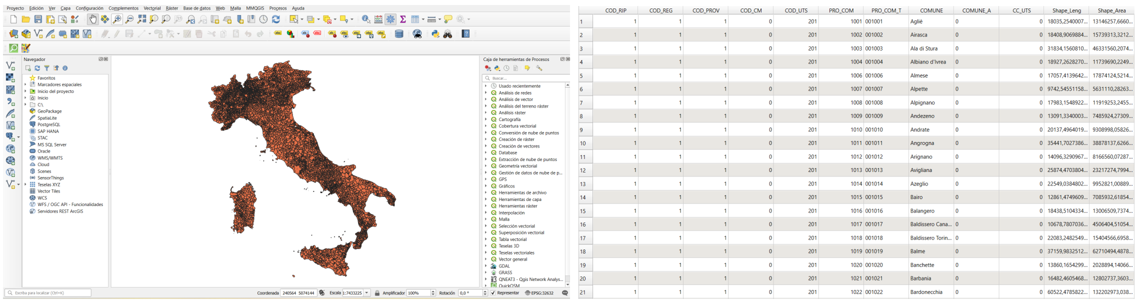
The geographical data was obtained from the official ISTAT website. The data was originally contained in a compressed archive (.zip) which, once extracted, contained vector files in shape file format (.shp). This data was imported into QGIS for initial visual inspection and for filtering, extraction and calculation of the necessary information. The datasets were then exported in GeoJSON format to facilitate their use in subsequent stages of the analysis. In particular, the goal of this step is extract the following key elements from the shape files of the municipalities and urban centres:

- The centroids of each municipality
- The list of neighbouring municipalities for each municipality
- The centroids of each urban centre

This information forms the spatial basis for constructing the origin-destination matrices and calculating the accessibility indicators that will be analysed in the next steps of the methodology.

Municipality data

After importing the vector layer (.shp) containing the raw data of Italian municipal boundaries into QGIS. The dataset covered the entire national territory. Figure 9a shows the initial cartographic representation of the layer, together with an example of the associated attribute table.



(a) *Municipality raw data visualization in QGIS* (b) *Variables of the municipality data*

Figure 9: *Visualization and variables of raw data about municipality in Lazio*

Since the analysis focuses exclusively on the Lazio region, it was necessary to extract a subset of data relating solely to municipalities belonging to that region. This operation was performed using the filtering function integrated in QGIS, accessible via the layer properties panel. In particular, the following query in SQL language (using OGR SQL syntax) was entered in the field dedicated to the filter expression:

```
"COD_REG" = 12
```

where `COD_REG` represents the variable containing the numerical code identifying the Lazio region within the ISTAT dataset. This condition made it possible to select only spatial elements with a region code equal to 12, generating a new vector layer containing only the municipalities of the Lazio region. This initial selection process optimised the dataset, bringing it into line with the geographical area of interest and improving the efficiency of subsequent spatial processing.

Once the vector layer containing the municipalities of the Lazio region had been isolated, two fundamental operations were carried out in QGIS:

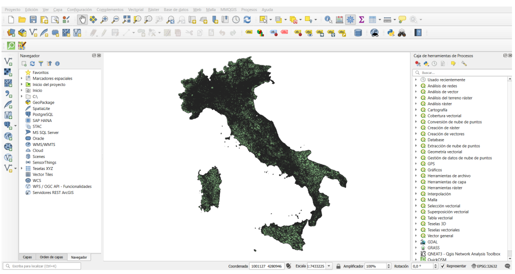
1. **Extraction of centroids**: Using the integrated QGIS tool called Centroids, available in the process tools menu, in the vector geometry section, a new point layer was generated containing the geometric centroids of each municipal polygon. This step allows each municipality to be represented by a single point representing the centre of the town hall, which is essential for subsequent analysis. The output was then exported from QGIS in GeoJSON format.

2. **Extraction of data relating to municipal boundaries**: As regards the extraction of neighbouring municipalities for each municipality, the process was divided into two distinct phases:
   - Attribute union operation by location in QGIS: Using the general vector analysis tool in QGIS called "Join attributes by location", a self-join operation was performed on the municipality layer, setting both the input layer and the comparison layer as the same municipal layer. 'Intersect' was chosen as the spatial predicate in order to identify all municipal polygons that intersect or share boundaries. The result of this operation was a new vector layer that was subsequently exported in GeoJSON format to facilitate further processing.
   - Cleaning and improving the GeoJSON dataset using Python scripts: Since the exported GeoJSON file was not completely clean or easily interpretable, a Python script based on the geopandas library was applied to optimise the information on neighbouring municipalities. Specifically, for each municipality contained in the GeoDataFrame, geometric (topological) adherence to other polygons was verified using the `.touches()` method, which returns adjacent polygons. The names of the neighbouring municipalities were collected in a list and assigned to the dedicated column 'neighbours' in the respective row of the GeoDataFrame. Finally, the dataset was simplified by keeping only the columns "MUNICIPALITY" and 'neighbours' and saved in a well-formatted JSON file, ready to be used in the subsequent stages of spatial analysis and territorial modelling.

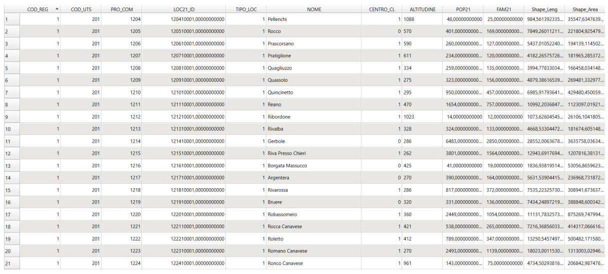
The data relating to municipalities will be used to assign public services classified as local to each municipality in the region.

Urban centre data

As with municipalities, the layer relating to urban centres, downloaded from the ISTAT website, contains the geometric definition of all urban centres in Italy. As shown in the figure 10a, the dataset includes spatial information on each urban centre, but for the purposes of analysis it was necessary to isolate only the urban centres in the Lazio region.



(a) *Urban centre raw data visualization in QGIS*



	COD_REG	COD_UTS	PRO_COM	LOC21_ID	TIPO_LOC	NOME	CENTRO_CL	ALTITUDINE	POP21	FAM21	Shape_Leng	Shape_Area
1	1	201	1204	120410001,00000000000	1	Pellenchi	1	1088	48,00000000000	25,00000000000	984,561392335...	35547,6347639...
2	1	201	1205	120510001,00000000000	1	Rocco	0	570	401,000000000...	169,000000000...	7849,26011211...	221804,925479...
3	1	201	1206	120610001,00000000000	1	Prascorsano	1	590	260,000000000...	127,000000000...	5437,01052240...	194139,114502...
4	1	201	1207	120710001,00000000000	1	Pratiglione	1	611	234,000000000...	120,000000000...	4182,26575726...	181965,285372...
5	1	201	1208	120810001,00000000000	1	Quagliuzzo	1	334	259,000000000...	135,000000000...	3994,77833034...	166458,034148...
6	1	201	1209	120910001,00000000000	1	Quassolo	1	275	323,000000000...	156,000000000...	4879,38616539...	269481,332977...
7	1	201	1210	121010001,00000000000	1	Quincinetto	1	295	950,000000000...	457,000000000...	6985,91793641...	429480,450059...
8	1	201	1211	121110001,00000000000	1	Reano	1	470	1654,00000000...	757,000000000...	10992,2036847...	1123097,01921...
9	1	201	1212	121210001,00000000000	1	Ribordone	1	1023	14,00000000000	12,00000000000	1073,62604545...	26106,1041805...
10	1	201	1213	121310001,00000000000	1	Rivalba	1	328	324,000000000...	133,000000000...	4668,53304472...	181674,605148...
11	1	201	1214	121410001,00000000000	1	Gerbole	0	286	6483,00000000...	2850,00000000...	28552,0063678...	3635758,03634...
12	1	201	1215	121510001,00000000000	1	Riva Presso Chieri	1	262	3801,00000000...	1564,00000000...	12943,6917694...	1207816,38131...
13	1	201	1216	121610001,00000000000	1	Borgata Massucco	0	425	41,00000000000	19,00000000000	1836,93819514...	53056,8659623...
14	1	201	1217	121710001,00000000000	1	Argentera	0	270	390,000000000...	164,000000000...	5631,53904415...	236968,731872...
15	1	201	1218	121810001,00000000000	1	Rivarossa	1	286	817,000000000...	372,000000000...	7535,22325730...	308941,673637...
16	1	201	1219	121910001,00000000000	1	Bruere	0	320	331,000000000...	136,000000000...	7434,24897219...	388848,600342...
17	1	201	1220	122010001,00000000000	1	Robassomero	1	360	2449,00000000...	1054,00000000...	11131,7832573...	875269,747994...
18	1	201	1221	122110001,00000000000	1	Rocca Canavese	1	421	538,000000000...	265,000000000...	7216,36856033...	414317,066616...
19	1	201	1222	122210001,00000000000	1	Roletto	1	412	789,000000000...	347,000000000...	13250,5457497...	500482,171580...
20	1	201	1223	122310001,00000000000	1	Romano Canavese	1	270	2493,00000000...	1139,00000000...	18023,0011530...	1313003,02946...
21	1	201	1224	122410001,00000000000	1	Ronco Canavese	1	961	143,000000000...	75,00000000000	4734,50293816...	206842,987476...

(b) *Variables of the urban centre data*

Figure 10: *Visualization and variables of raw data about urban centre in Lazio*

In addition to the spatial filtering operation to select only urban centres located within the boundaries of the Lazio region, further filtering criteria were applied. In particular, a selection was made based on the variable TYPE_LOC, to exclude urban centres classified as industrial areas or as "scattered houses" and "industrial areas", which are not relevant for the purposes of the analysis focused on residential areas. The SQL query (OGR SQL) used to obtain the relevant subset of data was as follows:

```
("TIPO_LOC" = 1 OR "TIPO_LOC" = 2) AND "POP21" > 0
```

Where the condition `'POP21' > 0` ensures that only urban centres with a population greater than zero are included. The variable POP21, which represents the population of each urban centre, is particularly important, as it will be used in subsequent analyses to calculate accessibility indicators, weighting urban centres according to population density.

The last of the three specific objectives of this phase with spatial data is to generate a file containing the point centroids of urban centres. The centroids were initially calculated using the vector geometry tool integrated in QGIS, as described above for the municipalities. However, during a verification and validation phase, it emerged that many centroids were located outside the polygonal geometries of their respective urban centres. This phenomenon, highlighted, for example, in image 11, which shows the centroids calculated for the centres of "Centro" and "Forme di Suio-San Cataldo" di Castelforte (province of Latina), would have caused significant problems in the calculation of origin-destination matrices (OD matrices), particularly for the analysis of accessibility by public transport, as a centroid outside the geometry may not adequately represent the spatial position of the centre.

Figure 11: *Initial calculation of urban centre centroids – Municipality of Castelforte (LT)*

To resolve this issue, an analysis was conducted to realign the centroids using the road layer for the Lazio region as a reference. The road layer was imported into QGIS and filtered to retain only roads of territorial relevance within the region, using the following OGR SQL query:

```
"fclass" IN ('primary', 'secondary', 'residential', 'trunk', 'tertiary')
```

Next, using the 'Move points to nearest layer' tool in QGIS, the centroids outside the geometry were moved to the nearest road within the perimeter of the corresponding urban centre. This ensured that each centroid was correctly located within the geometry of the centre, thus improving the reliability of subsequent spatial analyses and accessibility calculations. Finally, the correct file containing the centroids of the urban centres was exported in GeoJSON format, thus ensuring optimal interoperability with the subsequent spatial analysis phases.

Next, using the 'Move points to nearest layer' tool in QGIS, the centroids outside the geometry were moved to the nearest road within the perimeter of the corresponding urban centre. This ensured that each centroid was correctly located within the geometry of the centre, thus improving the reliability of subsequent spatial analyses and accessibility calculations.

However, the validation process also revealed that several centroids, although technically located within the correct urban centre geometry, were placed in areas that were visibly disconnected from the actual built-up environment, such as isolated rural fields or agricultural zones. This became particularly evident during a visual inspection phase, in which all centroids were exported in KML format and overlaid on satellite imagery using Google Maps.

Although no further correction was applied in these cases, this observation highlights the inherent limitations in spatial data granularity and automated centroid generation. These anomalies in this work were accepted as part of the spatial heterogeneity of the dataset, and the analysis proceeded using the corrected centroids.

The validated file containing the urban centre centroids was exported in GeoJSON format, thus ensuring optimal interoperability with the subsequent spatial analysis phases.

At the end of this first data collection phase, the following main files were generated: two GeoJSON files, containing the centroids of the municipalities and urban centres respectively, and a JSON file associating each municipality with a list of neighbouring municipalities. The next phase of data collection, which will be shown, is dedicated to the acquisition and preparation of data relating to education services, describing how the data was obtained and the adjustments necessary for analysis.

3.1.2 Locations data

The objective of this other sub-phase of data collection is to prepare data relating to the public services considered, in order to make them ready for the subsequent phases of the methodology. In particular, the objective is to obtain the following datasets:

- A dataset containing the locations of the schools of interest, with a variable indicating their category (e.g., nursery school, primary school, secondary school);
- A dataset containing the locations of general practitioners (family doctors);
- A dataset containing the hospitals in the area;
- A dataset containing emergency facilities;

These datasets will then be used to calculate accessibility indicators and to integrate spatial information on the distribution of public services in the Lazio region.

Educational centre data

In order to obtain a consistent and specific dataset of schools of interest in the Lazio region, after downloading it from the MIUR website in CSV format, the raw data contained in a CSV file was filtered using Python with Pandas.

The first step was to select only schools located in the Lazio region, based on the "REGION" variable, which allowed the analysis to be restricted to the territory of interest.

The dataset was then further refined to include only schools belonging to certain levels of education relevant to the analysis. In particular, schools with the following types of education were selected: nursery schools, primary schools, lower secondary schools and comprehensive schools. Although comprehensive schools do not appear in the classification of public services previously presented, they are essential institutions as they encompass all the mentioned school levels, and they play a crucial role in assigning school services to municipalities lacking full service coverage. This will be detailed later in the methodology. This was achieved by filtering the variable that describes the type of the school, which classifies the different levels of education. The resulting dataset includes only the relevant schools.

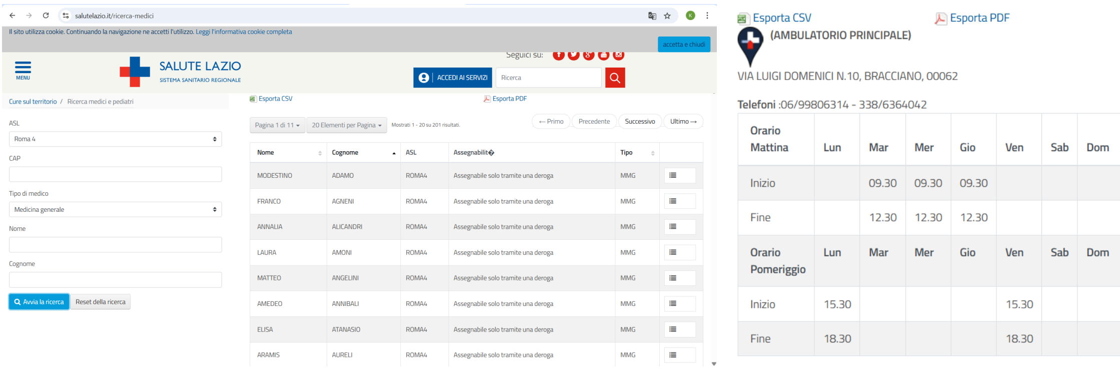
Finally, the data frame was exported to a new CSV file, ready for use in the next stages of analysis. This dataset is a fundamental part of the creation of accessibility indicators, as schools are one of the main public services of interest in the field of education.

Practitioners clinics data

Obtaining information on the locations of family doctors' surgeries was handled differently from other data, as the information was not available for direct download. In this case, it was necessary to apply a web scraping technique to extract the data from the family doctor search engine available on the ASL SALUTE LAZIO platform.

The complete source code used to perform the web scraping to get the family doctors clinics is available at: https://github.com/keving-ing/Distance-project/tree/main/Raw_data_processing/web_scraping_salute.

A preliminary consideration that is fundamental for the collection of health data in the Lazio region is that the health system is organised into health zones (or metropolitan ASLs), given that the analysis was carried out only with municipalities with fewer than 40,000 inhabitants. The aim is to collect data related to ASL ROMA 4, ASL ROMA 5, ASL ROMA 6, ASL VITERBO, ASL LATINA, ASL RIETI and ASL FROSINONE. Those from other health zones refer only to data from clinics in the municipality of Rome.



(a) *Search engine home page* (b) *Clinic address*

Figure 12: *Search engine of family doctors in the Lazio region - ASL SALUTE LAZIO platform*

The first step was to analyse the structure of the web page to understand how the data was organised. In particular, the HTML format of the page was examined. It can be observed in the image 12 that, after selecting the service type 'family doctors', a list of doctors appeared, and for each doctor there was a button that, when clicked, showed the details of the clinic, including the address and other relevant information.

The next step was to implement the web scraping process, which allowed us to automatically extract information from the HTML elements of the page, collecting the addresses of the family doctors' surgeries for each health unit. For this step, the browser automation tool `Selenium` was used, which allowed dynamic interaction with the SALUTE LAZIO platform, in particular with the family doctor search engine. The scraping methodology followed these steps:

1. Access to the platform: The web page containing the doctor search was opened via `Selenium`, using the Chrome browser driver. The URL for accessing the platform was the starting point for navigation;
2. Selection of the local health authority: The value corresponding to the relevant local health authority was selected iteratively from the drop-down menu dedicated to the selection of the local health authority. Interaction with the selection field was achieved using the Selenium command Select, which automated the selection of the correct value in the field with ID `_it_smc_laziocrea_sysmed_web_internal_portlet_SearchFormPortlet_asl`:
3. Starting the search: Once the ASL was selected, the "Start search" button was clicked, whose HTML element was identifiable by its ID `_it_smc_laziocrea_sysmed_web_internal_portlet_SearchFormPortlet_submit`. Clicking this button started the search for doctors, and the system loaded the results table;
4. Data extraction: Each doctor listed in the table was processed iteratively. The information was extracted from the rows of the table, identifiable by the CSS selector `.table-data tr`. For each row, the first name and surname data were retrieved from the table cells, located

with the tags `<td>`. To access the clinic details, the 'Details' button (element with class `page-link`) was identified, which opened a window with the clinic address. The addresses were extracted from all text elements in paragraphs with class `p.text-uppercase`, filtering out the relevant ones and ignoring information about opening hours.

5. Navigating between results pages: The resulting page showed a list of doctors divided into several pages. Using the 'Next' navigation link, identified by the id text `Next`, the bot was able to move to the next page and continue collecting data. This process was repeated until all results pages were covered;

6. Saving data: Once the extraction was complete, the data relating to the doctors (name, surname and address of the surgery) was organised into a data frame using the pandas library, removing any duplicates. Finally, the dataset was exported in CSV format using pandas methods for later use in analysis.

This process made it possible to automatically collect the addresses of family doctors' surgeries in the area of interest, creating a dataset useful for analysing the accessibility of healthcare services in the area.

Hospitals and emergency ambulatory data

With regard to data relating to hospitals and emergency facilities, no complex pre-processing operations were necessary, as the data, already available in CSV format, contained detailed information on hospitals in the Lazio region, including addresses and information relevant to accessibility analysis. As indicated, the dataset was acquired directly from official sources and, as it was already correctly structured, it was used without significant changes. The information in the dataset was then used for the subsequent stages of the accessibility analysis, contributing to the creation of origin-destination matrices for the calculation of indicators.

At the end of the data collection phase, all the datasets needed to analyse accessibility to public services in the Lazio region were obtained and prepared. Data on municipalities, urban centres and public services (education, healthcare, hospitals and emergency services) were extracted, filtered and integrated into formats compatible with the subsequent analytical phases. With this data ready, the next step in the methodology involves calculating origin-destination matrices, a key step in determining accessibility to various public services based on distances and travel times between urban centres and points of interest. The next chapter is dedicated to this phase, which will describe the methods used to construct and analyse the matrices, integrating geospatial and socio-demographic information into a territorial accessibility model.

3.2 Calculation of origin-destination matrices

The calculation of origin-destination matrices (OD matrix) is a crucial step in assessing accessibility to public services. In this phase, the calculation of distances is separated into two sub-phases, based on the classification of services as either local or metropolitan. In particular, local services, such as schools and general practitioners, will be analysed separately from metropolitan services, such as hospitals and emergency facilities, which may require different modes of access or allocation due to their distribution across the territory. The OD matrices for local and metropolitan services will then be calculated taking into account the specific characteristics and administrative constraints when it comes to assigning which service to which urban centre.

The methodology is not limited to simply calculating distances, but also includes a series of data preparation operations. This preparatory process is essential to ensure that distances are calculated correctly. In particular, factors such as the correct definition of origins (urban centres) and destinations (services) and the correct assignment of distances and travel times based on the type of service must be taken into account.

The methodology considers online resources such as Google Maps to assess travel times from the centroids of the considered urban centres to the services. Google Maps is a very valuable source of information, as the spatial data of the road network are enriched with temporal data on travel times in different traffic conditions, providing important insights for accessibility. Google Maps employs graph data structures to calculate the shortest path from the origin (point A) to the destination (point B) and an algorithm to find the shortest path between a given origin and destination. Google Maps determines the approximate arrival time taking into account factors such as the remaining distance, average speed due to real-time traffic conditions, historical data, and the officially recommended speed. Google Maps API requests are also made with the 'transit' mode to calculate the distance, in km, and time from an origin to a destination by public transport; It requests directions or distance using public transport routes (when available).

This work is based on realistic data, and Google Maps algorithms reflect more factors than a GIS transportation model based solely on infrastructure data. This section describes the details of data preparation and the methods used to calculate OD matrices, providing a comprehensive overview of the accessibility analysis process. The complete source code used to perform the OD matrix calculations is available at: https://github.com/keving-ing/Distance-project/tree/main/Script.

3.2.1 Distances to local public services: educational and general medicine

The process followed to construct the distance and time matrices from each urban centre to the most appropriate, educational and family doctors centre, in of their municipality, can be defined in four stages:

1. Identification of points of origin, municipalities and urban centres;
2. Identification of the service coordinates;
3. Assignment services to municipalities;
4. Call request to Google Distance Matrix API.

Identification of points of origin, municipalities and urban centres

- Municipalities with up to 40,000 inhabitants
  - First, the CSV file containing the population data for each municipality was uploaded, as shown in the data collection phase, obtained from the official ISTAT source. Using the pandas library in Python, the data was read and then filtered to select only municipalities with a population of less than 40,000 inhabitants. The result of this operation was a data frame containing a list of municipalities to be used for the next stage of analysis.
- Urban centres associated with each of these municipalities with population, except for those considered scattered and industrial zones
  - The GeoJSON file containing the spatial data of the urban centres' centroids, previously obtained during a pre-processing data collection phase, was uploaded. For each selected urban centre, the geographical coordinates were obtained in the UTM (Universal Transverse Mercator) reference system. Next, using a conversion function, the coordinates were transformed into the WGS84 geographic reference system (latitude and longitude), used for subsequent accessibility analyses. The final result is a dictionary that associates each urban centre with its ID and geographical coordinates (latitude and longitude).

Identification of the service coordinates

For some of the service locations, it was necessary to calculate the geographical coordinates (latitude and longitude), as the input data did not include this information in geographical format.

Which addresses and why will be shown in the paragraph below. The method used to carry out this process is briefly described here to provide the reader with the necessary context. The process of converting addresses into geographical coordinates was carried out using the Google Maps Geocoding API. The method used to obtain the coordinates is as follows:

- Preparing the API request
  - For each address, contained in the dataset of the public services considered, a request was made to the Google Maps API using the address parameter (containing the address of the service) and the API key. The request was sent using the HTTP GET method to the geocoding endpoint, which returns a response in JSON format containing the geographical information related to the address.
- Cache management
  - To optimise the process and reduce the number of repetitive calls in case of errors, a cache system has been implemented. Before sending the request to the API, the script checks whether the address has already been geocoded, extracting the data from the local cache (stored in a JSON file). If there is a hit in the cache, the coordinates are returned immediately, avoiding a new API request.
- Response processing
  - Once the response is received from the API, the programme checks that the request status is 'OK'. If the response is positive, the geographical coordinates (latitude and longitude) are extracted from the JSON structure returned by the API and saved in the cache. Otherwise, an error is logged, and a log message is written to an error file to monitor geocoding issues.
- Saving the results
  - Saving the results The geocoded data is then returned for use in subsequent spatial analyses.

If an error occurs during the geocoding process (e.g., address not found or problems with the API), an error log containing the message returned by the API is written so that the problematic data can be dealt.

Once the geographical coordinates had been obtained through the geocoding process, it was necessary to perform a spatial validation of the geolocated points to verify that each point corresponding to a service was actually located within the correct municipal boundaries. The validation process was carried out using QGIS, a geographic information system software that allows spatial analysis to be performed. Each geolocated point (latitude and longitude) was associated with a specific municipality, as indicated in the database, which contains the name of the municipality to which it belongs as an attribute.

- Data preparation in QGIS
  - The geolocated data, in the form of points, was imported into QGIS in Geo-JSON format, with information about the points of interest (e.g., addresses of clinics or schools) and the municipality associated with each location. Similarly, the geometric layer of municipal boundaries, previously prepared, was loaded, containing the polygons that define the boundaries of each municipality in the Lazio region.
- Spatial check
  - For each geolocated point, a spatial intersection operation was performed in QGIS to check whether the point was within the polygon of the corresponding municipality. This was done using the "Check point within polygon" function, which checks whether the coordinates of the point are within the geometric boundary of the specified municipality.

- Validation results
  - Points that were within the boundaries of the municipality were considered valid, while those outside were flagged as anomalies. The results of this spatial validation were used to exclude or correct problematic data, ensuring that all locations were geographically correct and actually represented a point within the municipality to which they belonged.
- Manual corrections
  - For points that were not validated, manual corrections were made or, in some cases, the problematic points were excluded from the analysis to avoid errors in the subsequent stages of calculating the origin-destination matrices.

This validation process ensured that all geolocated points were correctly associated with municipalities, minimising the risk of errors in subsequent spatial analyses and providing a solid basis for calculating accessibility indicators.

Assignment services to municipalities

The first step involves identifying each of the infrastructures located within a municipality and automatically assigning it to the relevant municipality. After completing this step, the result is a JSON file containing the services available in each municipality. If a service is not available in a municipality, its corresponding category will remain empty. The output should follow the JSON structure outlined below:

```
{
    "MUNICIPALITY": {
        "SERVICE_TYPE": [
            {
                "SERVICE_DENOMINATION": "SERVICE_DENOMINATION",
                "SERVICE_LOCATION": "SERVICE_LOCATION"
            }
        ],
        "SERVICE_TYPE": []
    }
}
```

If a municipality lacks services of a required type, a search is conducted to find this missing public service type in neighbouring municipalities. The method applied is as follows:

(a) Identify the neighbouring municipalities of the municipality under consideration;

(b) Reference the JSON file generated in the first step to check if the required service type is available in any of the neighbouring municipalities;

(c) Consider all the centres found for the required service in the neighbouring municipalities;

(d) Calculate the network road distance using Google Maps between the centroid of the municipality and the centres identified:
   - if more than 10 educational centres are found, only the 10 closest centres, based on Euclidean distance, will be considered for further distance calculations using Google Maps;
   - in these calculations, the latitudes and longitudes of the service's centres are used as the coordinates for the Euclidean distance calculation.

(e) The Google Maps API call to calculate distances is performed using the following parameters, without considering the traffic:

```
params = {
    "origins": origin_str,
    "destinations": destination_str,
    "key": GOOGLE_MAPS_API_KEY,
    "mode": "driving"
}
```

(f) Assign the closest centre (in terms of distance in km) to the municipality;

(g) If no centres are found in the neighbouring municipalities, extend the search to the neighbours of the neighbouring municipalities for services.

The final JSON file will now have each types centre category filled in for each municipality

```
{
    "MUNICIPALITY_1": {
        "SERVICE_TYPE_1": [
            {
                "SERVICE_DENOMINATION": "SERVICE_DENOMINATION",
                "SERVICE_LOCATION": "SERVICE_LOCATION"
            }
        ],
        "SERVICE_TYPE_2": [
            {
                "SERVICE_DENOMINATION": "SERVICE_DENOMINATION",
                "SERVICE_LOCATION": "SERVICE_LOCATION"
            }
        ],
    },
    "MUNICIPALITY_2": {
        "SERVICE_TYPE_1": [
            {
                "SERVICE_DENOMINATION": "SERVICE_DENOMINATION",
                "SERVICE_LOCATION": "SERVICE_LOCATION"
            }
        ],


        "SERVICE_TYPE_2": [
            {
                "SERVICE_DENOMINATION": "SERVICE_DENOMINATION",
                "SERVICE_LOCATION": "SERVICE_LOCATION"
            }
        ],
    }
}
```

Call request to Google Distance Matrix API

Before making the API call, to simplify the process, the first step is add to the JSON file with the municipality and the services of the municipality, also the urban centres of the corresponding municipality.

```
{
    "MUNICIPALITY": {
        "SERVICE_TYPE": [
            {
                "SERVICE_DENOMINATION": "SERVICE_DENOMINATION",
                "SERVICE_LOCATION": "SERVICE_LOCATION"
            }
        ]
        "URBAN_CENTRE": [
            URBAN_CENTRE_1,
            URBAN_CENTRE_2
        ]
    }
}
```

Once the JSON is built, the number of elements required to complete the matrix will be determined. For each urban centre centroid, the distance to the educational centres within its municipality will be calculated. The following configurations are considered for the calculation of educational distance:

- Distance in km and time, at 8:00 am considering traffic (DRIVING)

```
params = {
    "origins": "|".join(origins),
    "destinations": "|".join(destinations),
    "key": GOOGLE_MAPS_API_KEY,
    "mode": "driving",
    "departure_time": time_unix_period  # Consider current traffic
}
```

- Distance in km and time with public transport at 8:00 am

```
params = {
    "origins": "|".join(origins),
    "destinations": "|".join(destinations),
    "key": GOOGLE_MAPS_API_KEY,
    "mode": "transit",
    "departure_time": time_unix_period
}
```

The API call is then made, with the origins being the centres of a municipality and the destinations being the different educational centres. Care is taken to avoid calculating the same distance twice, for example, if a school appears in two different categories. The calls are batched to comply with the following restrictions:

- In a single call, the maximum number of elements (in this case, the number of centres x the number of educational centres) cannot exceed 100;
- The number of origins (centres) and destinations (educational centres) must not exceed 25.

The objective of this phase was to obtain a final JSON file containing the distance between the centroid of each urban centre and the local service of interest for that urban centre. If the municipality of the urban centre already has the requested service, the distance is calculated for all the service facilities within the municipality. If the municipality does not have the service, the nearest service among the neighbouring municipalities is first assigned. Next, the distance is calculated only for the single infrastructure of the assigned service, considering the location of the latter in relation to the urban centre.

3.2.2 Distances to metropolitan public services: hospitals and emergency facilities

The method for calculating the distances to public services classified as metropolitan services, which include hospitals and emergency services, slightly differs from the method described earlier for local services. This process can be defined in four phases, with the only difference being in the method used for phase three. The methods used for the other phases are the same as those described for local data.

1. Identification of points of origin, municipalities and urban centres (refer 3.2.1 for detailed methodology).
2. Identification of the service coordinates (refer 3.2.1 for detailed methodology);
3. Assignment of services to urban centres;
4. Call request to Google Distance Matrix API (refer 3.2.1 for detailed methodology).

At this stage, the focus shifts to the methods used for assigning services to urban centres. In accordance with the constraints of the Italian healthcare system, this phase is treated separately, as the assignment methodology differs for hospitals and emergency services.

Assignment hospitals to urban centres

Before illustrating the methods used for the assignment of hospitals, two fundamental considerations must be taken into account:

- Service classification: As highlighted in the literature, it would be more accurate to classify the service as "healthcare provision" rather than simply "hospitals". For example, the classification could include medical, surgical, paediatric, and neonatal services, among others. However, due to the lack of open data from official sources, the analysis in this work is limited to calculating accessibility to "hospital services". Distance data to hospitals is stored in cached structures, allowing the possibility of future integration without the need to recalculate distances already computed.
- Regional health system division: The healthcare system in the Lazio region is divided into "health zones", represented by Local Health Authorities (Aziende Sanitarie Locali, ASLs). These are public entities within the Italian public administration, responsible for delivering healthcare services within a defined geographical area, usually at the provincial level. The ASLs are responsible for implementing the National Health Service and other legal obligations within their respective territories. Consequently, for administrative and economic reasons, each resident of a municipality will have access to the services available within the ASL of which their municipality is a part.

The method for assigning hospitals to urban centres follows the steps outlined below:

(a) The Euclidean distance from the centroid of the urban centre is calculated to all the hospitals located within the health zone to which the municipality belongs. This step allows the identification of the proximity of hospitals to the urban centre in terms of straight-line distance;

(b) The calculated distances are sorted in ascending order, from the closest to the farthest hospital;

(c) Calculate the distance using Google Maps between the centroid of the municipality and the centres identified

```
params = {
    "origins": "|".join(origins),
    "destinations": "|".join(destinations),
    "key": GOOGLE_MAPS_API_KEY,
    "mode": "driving",
    "departure_time": time_unix_period
}
```

(d) The driving distance is automatically saved in a cache data structure, making the previous calculation redundant. As a result, the driving distance between the urban centre and the hospital is already computed during this step of assignment, and therefore, step 4 becomes unnecessary for future calculations;

(e) The urban centre is assigned to the closest hospital based on the calculated driving distance, ensuring that the service is linked to the nearest hospital;

(f) The result of this phase is a JSON file structured as follows:

```
{
    "MUNICIPALITY": {
        "urban_centres": [
            URBAN_CENTRE_1,
            URBAN_CENTRE_2
        ],
        "DISTANCE": {
            "URBAN_CENTRE_1": {
                "HOSPITAL_ASSIGNED": {
                    "distanza_m": distance,
                    "tempo_s": time
                }
            },
            "URBAN_CENTRE_2": {
                "HOSPITAL_ASSIGNED": {
                    "distanza_m": distance,
                    "tempo_s": time
                }
            }
        }
    }
}
```

Although step 4 is effectively redundant due to the caching mechanism, it is still carried out for the public transport distance and to ensure that the distances are added to the JSON file in a consistent format, making the file compatible with other data aggregation processes in the subsequent stages.

Assignment emergency facility to urban centres

In the case of emergencies, an individual in a state of emergency is not required to go only to the emergency facility within the ASL of their municipality of residence. Therefore, the following methodological steps are used for assigning emergency facilities to urban centres. As previously mentioned, the datasets for hospitals and emergency facilities are slightly different, as not all hospitals within the hospital service provide emergency services.

(a) The Euclidean distance from the centroid of the urban centre is calculated to all emergency facilities in the Lazio region. This allows identification of the proximity of emergency services in terms of straight-line distance;

(b) The calculated distances are sorted in ascending order, from the closest to the farthest emergency facility;

(c) The driving distance from the centroid of the urban centre to the 3 closest emergency facilities, based on Euclidean distance, is calculated using the Google Maps API;

```
params = {
    "origins": "|".join(origins),
    "destinations": "|".join(destinations),
    "key": GOOGLE_MAPS_API_KEY,
    "mode": "driving",
    "departure_time": time_unix_period
}
```

(d) The driving distance is automatically saved in a cache data structure, making the previous calculation redundant. Therefore, the driving distance between the urban centre and the emergency facility is already computed during this step, and step 4 becomes unnecessary for future calculations;

(e) The urban centre is assigned to the 3 closest emergency facilities based on the calculated driving distance, ensuring that emergency services are assigned to the nearest facilities;

(f) The result of this phase is a JSON file structured as the previous for hospitals.

It is important to note that, in this phase, the public transport distance is not calculated, as it would not be appropriate to use public transport in an emergency situation.

3.2.3 Distances with public transport

As outlined in the methodology, the distance values are obtained using the Google Maps API, both for car travel and public transport. However, as previously mentioned and as will be further detailed in the chapter on results, some public transport distances could not be obtained. This issue is assumed to be caused by the fact that Google Maps does not contain updated data for all transport companies, the possibility of service unavailability during selected hours, or areas that are not covered by the public transport network.

Given these limitations, an alternative methodology was implemented to obtain the missing distances for public transport. This section outlines the alternative methods used, specifically:

- OpenTripPlanner (OTP) for routing: The first alternative method involved creating a routing network using OpenTripPlanner (OTP), which is capable of handling bus schedules and routes by integrating GTFS data (General Transit Feed Specification). OTP allows for accurate calculation of distances by considering transport schedules and route details. This solution helps overcome the limitations of Google Maps by providing more precise transit information for specific routes and times.

- Web Scraping with Moovit application: Another method used to obtain public transport distances involved scraping data from the Moovit application. Moovit provides real-time transit information, including bus schedules, routes, and travel times. By scraping the results from Moovit for the desired origin and destination points, it was possible to obtain public transport distances and travel times for cases where Google Maps failed to provide accurate data.

OpenTripPlanner (OTP) for routing

The first alternative methodology for calculating public transport distances is based on the use of OpenTripPlanner (OTP), a widely used open-source tool designed for multi-modal transport routing. OTP can process various types of data such as GTFS (General Transit Feed Specification) data and road networks to create a routing graph. This allows for the calculation of travel itineraries that combine walking and public transport, addressing the need for accessibility analysis across different transport modes. The process is divided into the following key steps:

(a) The first step in the process involves preparing the necessary data for OTP. This includes downloading and organizing the GTFS data, which provides schedules and stop locations for public transportation, as well as the road network data for walking routes. The GTFS data is typically made available by local transit authorities, while the road network is either provided directly or can be generated from available geographical data sources. Once all the data is gathered and placed in a single directory, the OTP graph is built using the following command:

```
java -Xmx4G -jar otp-2.4.0-shaded.jar --build data/
```

This command initializes the OTP application and processes the GTFS and road network data, creating the underlying graph used for routing. The `-Xmx4G`option ensures that OTP is allocated sufficient memory (4 GB in this case) to handle the large datasets.

(b) Once the graph is built, the OTP server is started to handle routing requests:

```
java -Xmx4G -jar otp-2.4.0-shaded.jar --serve --load data/
```

The `-serve` flag starts the OTP server, and the –load flag ensures that the data is loaded into memory. The server runs on localhost by default, providing an API endpoint to handle requests for routing between origin and destination points.

(c) A Python script is then used to interact with the OTP server. This script sends requests to the OTP API for calculating routes between an origin and a destination, both of which are specified by their geographical coordinates (latitude and longitude). The script uses the `requests` library to interact with the OTP server's API. The key parameters in the request are as follows:

- `fromPlace` and `toPlace`: These specify the origin and destination coordinates for the routing request.
- `time` and `date`: These parameters define the departure time for the journey
- `mode`: The mode is set to 'TRANSIT, WALK', indicating that OTP should find a route that combines both walking and public transit.
- `maxWalkDistance`: This parameter limits the maximum walking distance to 1 km for each journey leg.
- `arriveBy`: This is set to 'false', meaning the route is calculated assuming departure at the specified time, rather than arrival.

If the route is found, the script processes the itinerary, which includes walking legs and transit connections. The total duration and distance for the itinerary are calculated and saved to the log file.

(d) If OTP is unable to find a route (e.g., no connection exists between the origin and destination), the script logs this as an error with the message: error: No transit connection found between origin and destination. This ensures that any failed requests are captured for further analysis.

(e) The script processes multiple origin-destination pairs in a batch, reading them from a log file. The regular expression pattern in the script extracts latitude and longitude coordinates for both the origin and destination from the log file.

(f) The results of the routing request, including travel times and distances, are logged in a file. The output includes details such as the travel time, distance, and individual route steps (e.g., walking and transit segments) with information on the transport route, agency, and the specific stops or stations involved.

(g) If a successful route is found, the results are formatted and saved as follows:

```
OTP found: ## minutes, ## km
1. WALK from 'Location A' to 'Stop 1'
2. BUS from 'Stop 1' to 'Stop 2' with route 'Line 1'
3. WALK from 'Stop 2' to 'Destination B'
```

This method is used for all distances that need to be calculated, allowing for comparison between them and with the results obtained using the Google Maps API.

This alternative method was primarily implemented as a validation step, in order to assess whether different routing engines (in this case, OpenTripPlanner and Google Maps) would produce divergent results under the same input conditions. However, contrary to initial expectations, the results obtained using OTP did not reveal substantial differences when compared to those retrieved via the Google Maps API.

According to the official Google documentation, the platform already integrates all publicly available GTFS feeds when generating public transport itineraries. As a consequence, the accessibility calculations based on Google Maps can be considered reliable and inclusive of local transit data. Therefore, the OTP implementation served as a methodological cross-check rather than a core component of the final accessibility analysis.

Distance Calculation Using Moovit

To complement the OpenTripPlanner approach, an alternative method for obtaining public transport distance data involves using Moovit, a widely used mobile application for real-time public transport information. This method relies on web scraping to collect data directly from the Moovit interface, which provides the most accurate and up-to-date transit information available. The process for calculating public transport distances using Moovit is detailed below.

Moovit does not provide an official API for free, so the data was collected using web scraping techniques. The following steps describe how the scraping process works:

(a) Initially, the address of the origin and destination are provided as inputs to the Moovit application. These address can either be extracted from a previously processed dataset (such as centroids of urban centres) or entered manually. The system accepts not only latitude and longitude as parameters.

(b) he script navigates to the Moovit website or app interface using a headless browser (e.g., Selenium), where the coordinates are entered into the input fields for "Origin" and "Destination".

(c) The time of travel and the transport modes (e.g., bus, metro) are also specified. The user can set the desired time of arrival or departure. By default, the current time is used, but this can be adjusted for different scheduling scenarios.

(d) The script simulates a user clicking the "Find Route" button on the Moovit interface. This action triggers the back-end calculation for the public transport itinerary, which includes walk, bus, metro, or other modes of transport. Moovit uses real-time data to generate this itinerary.

(e) Once the request is processed, the results are displayed on the screen. The script waits for the page to load and then extracts the relevant data (i.e., travel time and distance). This is done using web scraping tools to parse the HTML and extract the required information from the page elements.

(f) After scraping the travel time and distance data, the information is stored in a structured format, such as a CSV or JSON file, for further analysis and integration with other datasets (e.g., for OD Matrix calculation).

In this phase, the distances between urban centres and public transport services have been calculated using alternative methods to Google Maps, specifically leveraging OpenTripPlanner (OTP) and Moovit for more accurate and comprehensive routing data. These methods have allowed the integration of real-time transport information, improving the reliability of distance and time calculations for public transport services.

The results of these calculations are stored in a structured format (JSON), including both distance and travel times, for each origin-destination pair. The data is now prepared for further processing, ensuring consistency and compatibility with other datasets.

With the distances and travel times calculated and stored, the next step in the methodology is the aggregation of the data. In this next phase, the individual data points will be aggregated at before at urban centre level, then at the municipality level, combining all relevant data to derive meaningful indicators of accessibility. This will allow for the calculation of weighted averages, standard deviations, and other necessary metrics for further analysis of accessibility across different public services.

3.3 Data Aggregation for indicator creation

Once the distances have been calculated as described in the previous step of the methodology, these data must be aggregated and combined to achieve the primary objective: to obtain an accessibility indicator for each type of public service, classified in Table 3, with a value representing the distance in kilometres and travel time in minutes, both for car access and public transport.

This section of the methodology outlines the methods used for data aggregation, distinguishing between two phases: aggregation of data related to distances calculated using the driving mode and those based on public transport. The need for this separation arises from the fact that certain results (which will be detailed in the results section) for public transport distances could not be obtained due to the lack of publicly available data, both from the Google Maps API and from the GTFS data of the transport companies operating in the Lazio region. The source code referring to the methods in this part of the methodology can be found at the link: https://github.com/keving-ing/Distance-project/tree/main/Raw_data_processing/Script.

3.3.1 Driving distance data aggregations

After obtaining all the origin-destination distance matrices, the first step in the aggregation process is to compute the values for each urban centre, followed by aggregation at the municipal level.

- **Aggregation at the urban centre level**: Aggregation at the urban centre level is necessary because multiple infrastructure locations of the same service may have distances calculated for a single urban centre. Aggregation is performed by calculating the mean distances and standard deviations for each service type (e.g., early kindergarten, primary education, secondary education, family doctors, hospitals, and emergency services). As a result, a data frames is obtained for each urban centre, with the following structure:

Table 5: Structure of the data frame with data aggregated by urban centre - method: DR

		Public services	
Urban centres		**Mean**	**St. Dv**
	Distance [Km]		
	Distance [min]		

The mean distance (mean_km) and mean time (mean_min) for each urban centre are calculated using the following unweighted formulas:

$$\text{mean\_km} = \frac{\sum_{i=1}^{N} x_i}{N}$$

$$\text{mean\_min} = \frac{\sum_{i=1}^{N} y_i}{N}$$

Where:

- x_i and y_i are the distances and durations, respectively;
- N is the total number of distances or durations for the given urban centre.

The standard deviation is calculated using the pandas library, following the formula:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

Where:

- σ is the standard deviation.
- x_i represents each individual distance or duration value;
- μ is the mean of the distances or durations;
- N is the total number of distances or durations.

If, for a given service in a specific urban centre, there is only one infrastructure available, the standard deviation cannot be calculated. In this case, the standard deviation values for that service type will be omitted from the results.

- **Aggregation at municipality level**: After aggregating the data at the urban centre level, the next step is to aggregate the data by municipality. This involves calculating the average distance and time for each urban centre (the data aggregated in point 1) of the municipality, for each public service. Using these values, the weighted average and the combined weighted standard deviation are calculated for each municipality, with the total population of urban centres serving as the weight. The dataset of the urban centres contains for each urban centre its population, so this information was used. This results in a data frame with the following structure:

Table 6: Structure of the data frame with data aggregated by municipality - method: DR

		Public services	
Munic ipality		**Mean**	**St. Dv**
	Distance [Km]		
	Distance [min]		

The weighted average for each service type (e.g., mean distance to the service and mean travel time) was calculated for each municipality. The weight used in the calculation was the population of the urban centres within the municipality. For each service type, the weighted average is computed as follows:

$$\text{mean\_km} = \frac{\sum_{i=1}^{N}(x_i \cdot w_i)}{\sum_{i=1}^{N} w_i}$$

$$\text{mean\_min} = \frac{\sum_{i=1}^{N}(y_i \cdot w_i)}{\sum_{i=1}^{N} w_i}$$

Where:

- x_i and y_i represent the distances and durations for the urban centre i;
- w_i is the population of the corresponding urban centre;
- N is the total number of urban centres in the municipality.

The standard deviation for each service type is computed at the municipality level to capture the variation in distances and times across the urban centres. The formula for the standard deviation is:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

Where:

- σ is the standard deviation,
- x_i represents each individual distance or duration value;
- μ is the weighted mean of the distances or durations;
- N is the total number of distances or durations.

3.3.2 Public transport data aggregations

Before moving on to see how the data on access distances by public transport were aggregated, it is important to note that some distances could not be calculated and are therefore missing, due to a lack of data. A process of imputation of values was necessary.

- **Aggregation at the urban centre level**: The aggregation for public transport data is practically the same as that described in the method for aggregation at the level of car distance data (see point 3.3.1), except that the number of services available for the distances is taken into account and specified. Therefore, in addition to the average and standard deviation information, the data frame will also contain the number of facilities for the service out of the total number in that urban centre that were considered. The structure of the data frame is as follows:

Table 7: Structure of the data frame with data aggregated by urban centre - method: TP

		Public services		
Urban centres		**Mean**	**St. Dv**	**Present / Total**
	Distance [Km]			
	Distance [min]			

So now, for the types of services where there is no value for that service in public transport, assigning this value is necessary, as missing data compromises the accuracy of aggregation by municipality. The following section illustrates how values were assigned for the urban cores that do not have public transport distance data for a given type of service.

Imputation of public transport distances missing values

Given the spatial nature of the data and the lack of strong predictive variables available at the urban centres level, regression-based imputation methods were deemed inappropriate. Instead, a conservative penalized approach was chosen, leveraging the maximum observed value at the municipality level as a bounding reference. The penalty was scaled based on the proportion of available data and the absence of services in the area, following principles commonly used in territorial accessibility studies. (Ballas et al., 2006; Longley et al., 2015) This method avoids underestimation and allows a flexible control of uncertainty via a tunable parameter λ. Missing public transport distances (km and minutes) values were imputed using a penalized maximum-based approach, parametrized by λ. This process was applied independently for each public service type $c \in \{\text{Kindergarten}, \text{Primary}, \text{Secondary}, \text{General\_medicine}, \text{Hospitals}\}$, and for each urban centres i.

Let:

- $T_i^{(c)}$ be the observed (possibly missing) travel time for category c in urban centre i,
- $T_{\max}^{(m,c)}$ be the maximum observed (non-missing) value for category c in municipality m,
- $p^{(m,c)}$ be the proportion of non-missing values for category c in municipality m,
- $\lambda \in \mathbb{R}^+$ be the penalization factor,
- $\delta_i^{(c)} \in \{0, 1\}$ be a binary indicator: $\delta_i^{(c)} = 1$ if the urban centre has a school of category c, 0 otherwise.

Imputation is performed only if:

- $T_i^{(c)}$ is missing,
- $T_{\max}^{(m,c)}$ is available (i.e., at least one valid value in the municipality),
- $\delta_i^{(c)} = 0$ (i.e., the urban centre does not have a school of that category).

The imputed value is then computed as:

$$\hat{T}_i^{(c,\lambda)} = T_{\max}^{(m,c)} \cdot \left[(1 - p^{(m,c)}) \cdot \lambda \cdot \beta + p^{(m,c)} \right]$$

where $\beta = 1.5$ is a fixed penalty multiplier applied to control the degree of imputation. No imputation is performed if the above conditions are not met.

- **Aggregation at municipality level**

  Given that the original dataset provides accessibility indicators at the *urban centre* level, values were aggregated to the *municipality* level for spatial consistency and interpretability. For each school category $c \in \{\text{Kindergarten}, \text{Primary}, \text{Secondary}, \text{General\_medicine}, \text{Hospitals}\}$, and for each imputation scenario defined by a parameter λ, the population-weighted average travel time for municipality m was computed as:

$$\bar{T}_m^{(c,\lambda)} = \frac{\sum_{i \in \mathcal{N}_m} T_i^{(c,\lambda)} \cdot P_i}{\sum_{i \in \mathcal{N}_m} P_i}$$

where:

- $\mathcal{N}_m$ is the set of urban centres belonging to municipality m,
- $T_i^{(c,\lambda)}$ is the (possibly imputed) mean travel time for urban centre i, category c, and lambda value λ,
- P_i is the population associated with urban centres i.

The methodology outlined in the chapter, from the identification of urban centres and the calculation of distances to the assignment of services, has culminated in the creation of a structured file containing these indicators in CSV format. This file is ready for use in the next stages of analysis and visualization.

Statistical analysis of the obtained indicators has been conducted to examine the distributions and to address the research questions outlined in the objectives of the study. These analyses provide valuable insights into the accessibility of different public services, enabling a deeper understanding of the spatial inequalities in service availability.

The results of this methodology will be integrated into a dataset and displayed through the visualization techniques described in the subsequent chapters. This will allow decision-makers and stakeholders to easily interpret and act upon the insights generated from the data.

3.4 Publication of the data

The final step in the methodology, after collecting, processing and finally aggregating the data to define the indicators, is to publish the data infrastructure that has been created.

First step, to ensure that the datasets are compliant with Open Data publication standards, particularly those required for integration into the European Data Portal, the data must be transformed into RDF (Resource Description Framework) format. RDF enables data to be linked, semantically described, and easily consumed across distributed systems, in line with the specifications defined in the DCAT (Data Catalogue Vocabulary) standard maintained by the W3C (W3C, 2022).

For this purpose, a dedicated RDF transformation pipeline developed by the Department of Computer Engineering at University Carlos III of Madrid (UC3M) was adopted. This pipeline is part of an ongoing initiative for publishing accessibility indicators for the Community of Madrid, and it has been adapted in this work for datasets concerning the Lazio Region.

The pipeline takes as input a series of structured CSV files and produces RDF-compliant metadata and indicator values, serialized for publication. However, the transformation process requires that the input files strictly conform to a predefined schema, aligned with the accessibility data model adopted by the Madrid Region. This includes using standardized field names, consistent indicator coding, and proper formatting of geographic and statistical values.

As a result, prior to execution, all CSV files produced in this project were restructured to meet the required specifications. This step ensures full interoperability with the RDF pipeline and guarantees that the output complies with the standards defined for Linked Open Data across Europe.

In parallel with DCAT-based RDF serialisation, the semantic structure of the dataset has been modelled according to the ontology defined in the LoDCOREMadrid project 13. This ontology describes the entities and relationships that are fundamental for representing territorial indicators in an interoperable manner.

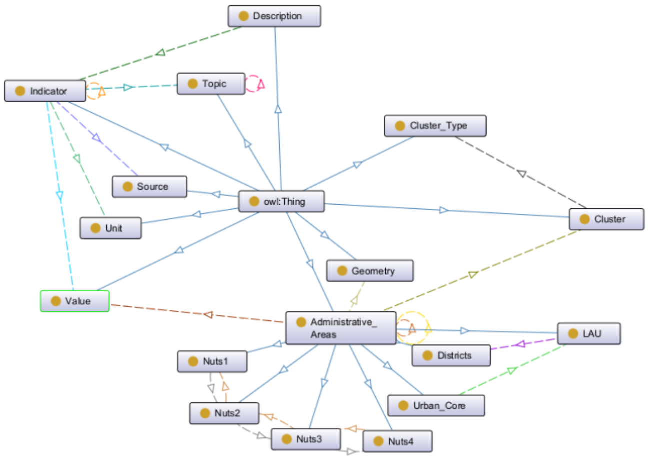


Figure 13: *Ontology as defined in the LoDCOREMadrid project (UC3M, 2024)*

Specifically:

- Each indicator is an instance of the `Indicator` class and is described by a `Description`.
- It is associated with an `Administrative_Area` (e.g. LAU or Urban Centre).
- It has a `Unit`, a `Value`, and a `Source` from which it was obtained (`is_obtained`).
- It is organised into `Topics` through the `belongs_to` property.

This ontological approach allows for a formal representation of data that facilitates both integration with existing datasets and semantic querying via SPARQL, ensuring full compliance with Linked Open Data publication standards.

Furthermore, the RDF output is also published via a Virtuoso triple store, which enables SPARQL-based querying and integration into the European data infrastructure. Virtuoso is a high-performance, scalable graph database engine designed to store and manage RDF data. It supports the SPARQL query language and provides a public endpoint for Linked Open Data consumption. Publishing the accessibility indicators through Virtuoso ensures that the data is not only compliant with semantic web standards but also queryable via SPARQL endpoints, discoverable by third-party platforms, and aligned with the best practices for Linked Data publication.

The entire pubblic publication process has been adapted to the LoDCOREMadrid framework, an initiative of the Department of Computer Engineering at the Carlos III University of Madrid for the publication of Linked Open Data relating to territorial indicators on attractiveness and accessibility. The resulting files have been described using RDF triples, using standardised prefixes and properties consistent with the LoDCORE project specifications.

Below is a simplified example of the description of one of the datasets produced, relating to the average time taken to access educational services by public transport in Roma:

```
@prefix dcat: <http://www.w3.org/ns/dcat#> .
@prefix dct: <http://purl.org/dc/terms/> .
<http://lodcoremadrid.uc3m.es/dataset/educacion_tiempo_tp_normalizado_lazio>
a dcat:Dataset;
    dct:title "Educación - Tiempo transporte público normalizado por población
    (Lazio)"@es;
    dct:identifier "3.98";
    dct:publisher <https://lodcoremadrid.org> ;
    dcat:theme <http://lodcoremadrid.uc3m.es/theme/educacion>;
    dcat:distribution <https://example.org/data/educacion_tiempo_tp
    \_normalizado\_lazio.csv>
```

To ensure compliance with European metadata standards and enable publication in Linked Open Data portals, each indicator is accompanied by a dedicated metadata record, which includes the essential descriptive and semantic fields. Each metadata file is structured with fields shown in the table 8.

Table 8: Structure of metadata fields, aligned with DCAT and Linked Open Data standards

Field Name	**Description**
`Block Code`	Internal code identifying the thematic block or group (e.g., `SERVICE_ACCESSIBILITY`)
`Variable Code`	Unique identifier assigned to the specific indicator, matching the internal ID (e.g., `IND301`)
`Block Name`	Full name of the thematic domain the indicator belongs to (e.g., *Access to Education*, *Access to Healthcare*)
`Variable Name`	Short descriptive title of the indicator (e.g., *Time to reach hospitals by car*)
`Variable Description`	A brief explanation of what the indicator measures, how it is calculated, and its purpose
`Source`	The computational formula or methodological source used to derive the indicator
`Unit of Measure`	The unit in which the indicator value is expressed (e.g., *minutes*, *kilometres*)
`Identifier (URI)`	A persistent URI or unique code for the indicator, used in RDF serialization and Linked Data publication
`Spatial Coverage`	The geographical extent the indicator refers to (e.g., Lazio Region, municipalities, urban centres)
`Temporal Coverage`	The reference time period of the data (e.g., 2025)
`Date of Creation / Update`	Date when the indicator was generated or last updated
`Publisher`	Entity responsible for making the dataset publicly available (e.g., UC3M Department of Computer Engineering)
`License`	Terms of reuse and redistribution (e.g., `Creative Commons Attribution 4.0`)
`Data Format`	The file format in which the dataset is made available (e.g., `CSV`, `RDF/Turtle`)
`Distribution URL`	The web address where the dataset or distribution file can be accessed or downloaded

These fields ensure semantic clarity and reusability of the data in open government contexts. Overall this approach not only ensures publication in accordance with Linked Open Data standards, but also allows for future semantic interoperability with similar datasets, facilitating integration into visualisation or automated analysis tools.

3.5 Web application design and implementation

After describing the methodology used to generate the dataset containing accessibility indicators, the next step involves making these data accessible and understandable to both technical and non-technical users. To this end, a web-based cartographic application was developed as part of the methodological framework. The interactive map viewer serves as the primary interface for visualising spatial accessibility data, allowing users to explore and compare indicators dynamically. The development process involved defining system requirements, building the front-end interface, and implementing the back-end infrastructure to manage data access and interaction.

To ensure a robust, user-oriented, functional and technically sound implementation, the requirements are categorized into three main groups:

- User requirements: This section identifies the needs and expectations of end-users interacting with the platform, including intuitiveness, responsiveness, and usability criteria relevant to different user profiles (e.g., policymakers, researchers, citizens, government).
- Functional requirements: This section specifies the concrete operations and behaviours the system must support, such as layer toggling, data filtering, dynamic tooltips, and indicator-based styling.
- Technical Requirements: This section outlines the technical constraints and standards necessary to guarantee system performance, scalability, data integrity.

This structured specification provides the foundation for the development and evaluation of the map-based visualization component.

User requirements

The web interface must be accessible from all major browsers to ensure compatibility across different systems and devices. The layout must follow modern UX/UI design principles, offering an intuitive and easy-to-navigate user experience, with clearly labelled elements and guided workflows.

Table 9: User Requirements

ID	Requirement Description	Component
UR-01	The web interface must be accessible from all major browsers (Chrome, Firefox, Edge, Safari) to ensure compatibility.	Frontend
UR-02	The layout must follow modern UX/UI design principles, providing an intuitive and easy-to-navigate experience.	Frontend
UR-03	Users must be able to select the accessibility indicator they wish to visualize through a clear and guided process.	Frontend
UR-04	Users must be able to visualize accessibility indicators on an interactive map with a responsive interface.	Map Viewer
UR-05	The system must include a legend to explain the colour scale or symbols used in each indicator.	Map Viewer
UR-06	Users must receive contextual instructions and responsive feedback to facilitate interpretation of the data.	Frontend
UR-07	Users must be able to apply filters to compare indicator results (e.g. by population size or province).	Frontend
UR-08	The interface must support full multilingual functionality in Italian, English, and Spanish.	Frontend
UR-09	The interface must provide an inclusive and informative consultation experience for international stakeholders.	Frontend

When the term front-end is used in the component table, it refers to all elements of the web page such as buttons and menus. In contrast, MapViewer specifically refers to the interactive map embedded within the web page.

These requirements are designed to ensure that the system provides an inclusive, user-friendly, and informative consultation experience for a diverse audience, including stakeholders from various countries and sectors.

Functional requirements

The system must store all relevant data related to accessibility indicators, including:

- Indicator values at both urban centre and municipality levels;
- Accessibility data for driving (with traffic) and public transport modes;
- Population data for each municipality;
- Geospatial layers for both administrative boundaries and accessibility zones.

These data, as described in the previous chapter, are stored in a structured database designed to support efficient querying and retrieval, while ensuring data integrity through validation rules and update constraints.

Table 10: Functional Requirements

ID	Requirement Description	Component
FR-01	The system must store indicator values at both urban centre and municipality levels.	Database
FR-02	The system must store accessibility data for both driving (with traffic) and public transport modes.	Database
FR-03	The system must store population data for each municipality.	Database
FR-04	The system must store geospatial layers for administrative boundaries and accessibility zones.	Database
FR-05	Data must be stored in a structured database with validation rules and update constraints to ensure integrity.	Database
FR-06	The system must allow users to dynamically extract and visualize accessibility results based on customizable filters.	Map Viewer
FR-07	The viewer must support selection of accessibility type (Education or Healthcare).	Map Viewer
FR-08	The viewer must support service type selection: Kindergarten, Primary, Secondary, General Practitioners, Hospitals, Emergency Care.	Map Viewer
FR-09	The viewer must support access mode selection: Driving (with traffic), Public Transport, or their difference.	Map Viewer
FR-10	The viewer must support metric selection: Distance or Time.	Map Viewer
FR-11	The map must colour municipalities or urban centres based on indicator values with continuous or categorized scales.	Map Viewer
FR-12	A legend must clearly explain the visual representation of the data.	Map Viewer
FR-13	Users must be able to filter results based on population range or provincial boundaries.	Frontend
FR-14	The system must allow highlighting of outliers, especially in differences between car and public transport accessibility.	Map Viewer
FR-15	The map must be interactive and responsive: clicking on a municipality reveals associated urban centres.	Map Viewer
FR-16	The interface must ensure that results are both technically accurate and easily interpretable.	Frontend

Technical requirements

The system must meet the following technical specifications to ensure robustness, scalability, and maintainability.

Table 11: Technical Requirements

ID	Requirement Description	Component
TR-01	The application shall follow a client-server architecture with a decoupled frontend (JavaScript framework) and a RESTful back-end.	Architecture
TR-02	The interactive map must be implemented using OpenLayers to efficiently handle geospatial layers.	Map Viewer
TR-03	All spatial and tabular data must be stored in a proper data structure with spatial indices for optimized query performance.	Database
TR-04	Data integrity must be ensured through foreign key constraints and validation checks during insert/update operations.	Database
TR-05	The map viewer must load visible layers and render accessibility results within a few seconds under standard network conditions.	Performance
TR-06	The system must handle concurrent user requests without performance degradation, leveraging caching where applicable.	Performance
TR-07	The platform must support dynamic language switching between Italian, English, and Spanish using i18n libraries or translation files.	Frontend
TR-08	The interface must be responsive and interactive, allowing users to click on a municipality to reveal related urban centres.	Map Viewer
TR-09	The application must be fully compatible with modern browsers (Chrome, Firefox, Safari, Edge) and support responsive design across devices.	Frontend

This part has detailed the user, functional, and technical requirements essential for the development of the interactive map viewer. These specifications provide a comprehensive foundation to guide the subsequent design and implementation phases. By aligning user needs with system capabilities, these measures aim to ensure an application that is both effective and user-centric. The subsequent section examines the architectural and interface design strategies adopted to translate these requirements into a cohesive and functional system.

To translate the previously defined requirements into a functional web application, it is essential to outline the architectural structure and the design principles adopted. This chapter describes the overall structure of the map viewer system, focusing on the interaction between components and the technologies used for implementation. The source code for the front-end web page is available at the link https://github.com/keving-ing/Distance-project/tree/main/OpenLayers/my-openlayers-project, while that for the web server (back-end) is at: https://github.com/keving-ing/Distance-project/tree/main/WEB_SERVER.

In general, the architecture of the map viewer can be divided into three main layers:

- **Front-End**: This layer represents what the user interacts with. It consists of a web application developed using HTML, CSS, and JavaScript, with OpenLayers as the core library for dynamic map visualization. It handles the visual presentation, user interaction, and triggers for core functionalities.
- **Back-End**: This layer is responsible for processing logic and data handling. It is built using Python and the Flask framework, which manages HTTP requests from the frontend and returns appropriate responses. It also handles access to the underlying open data files containing geospatial information and indicator values, structured for efficient retrieval and integration into the map interface.
- **Functions Layer**: The application is centred around three main user-facing functions:
  - Indicator selection: allowing users to choose the accessibility metric to be displayed.
  - Result filtering: providing dynamic filters based on population or administrative units.
  - Language change: enabling users to switch between supported languages (Italian, English, Spanish) for a multilingual experience.

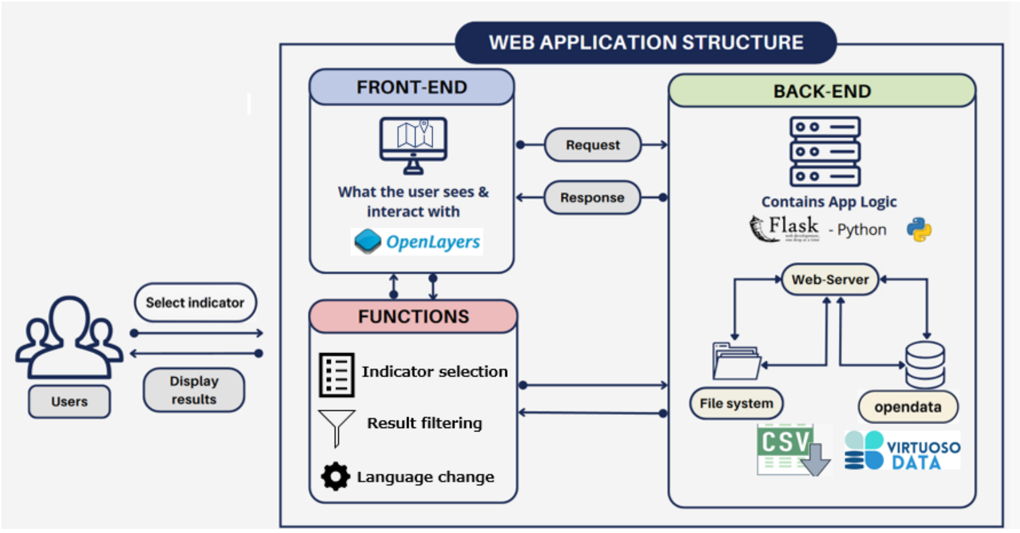


Figure 14: *General architecture of the map viewer*

Before diving into the implementation details of the front-end and back-end components, it is useful to understand the structure of the project files. The codebase is organized to ensure modularity, maintainability, and a clear separation of concerns.

The main components are structured as follows:

- `index.html`: The entry point of the web application. It initializes the base HTML structure and includes references to style sheets, JavaScript scripts, and the OpenLayers map container.
- `style.css`: Defines the visual style of the interface, ensuring consistency and responsiveness..
- src/ folder: This directory contains the core source code written in JavaScript and CSS. The structure inside this folder reflects the key functionalities of the application:
  - Map Rendering and logic:
    - `map.js`, `mapRoma.js`: Set up the OpenLayers map, define layers, views, and coordinate transformations.
    - `interactions_map.js`: Handles user interactions with the map, such as clicks and hover effects.
  - Data processing:
    - `municipalityProcessor.js`, `urbanCentreProcessor.js` and `differenceProcessor.js`: Process and format indicator data at the municipality and urban centre levels based on the data.
  - UI and Functional modules:
    - `filters.js`: Implements the logic for filtering results based on user-selected criteria.
    - `uiEvents.js`: Manages UI controls and event-driven interactions.
    - `i18n.js`: Supports multilingual functionality for Italian, English, and Spanish.
  - Utilities
    - `utilities.js`, `counter.js`: Contain helper functions and reusable logic across the application.
- node_modules/: Contains the installed dependencies managed via NPM, as defined in `package.json`.
- Flask back-end receives the request and performs a query to the database.

3.5.1 Front-end development

The front-end of the application is responsible for rendering the interactive user interface and managing all client-side logic. It serves as the visual and functional bridge between the user and the underlying data, allowing for dynamic exploration of accessibility indicators on a spatial map.

Preliminary interface design and navigation logic

Following the definition of all user, functional, and technical requirements, a detailed analysis of the page structure is performed to outline how the user will interact with the system.

To support the design phase, a preliminary sketch of the interface is created and complemented by a low-level navigation diagram (figure 15), which illustrates the internal logic of how the user navigates within the single-page application.

This diagram 15 is based on the key functional components of the interface, including:

- Macro-area selection menus (e.g., Health, Education, Public Transport);
- The main interactive map with toggles and filters;
- A contextual information panel;
- Result visualization with a colour-coded map and a reset mechanism;
- Access to detailed urban centre data by clicking on a municipality.

Diagram Legend:

- Rounded rectangles represent interface states or menu panels.
- Bold uppercase labels indicate user-triggered actions or transitions.
- Arrows define the navigation flow based on user interactions.

This structured visualization provides a comprehensive overview of how the interface responds to user actions, ensuring a coherent and predictable user experience aligned with the specified requirements.

High-level design

The front-end is developed using HTML, CSS, and vanilla JavaScript, with OpenLayers as the main library for geospatial visualization. The application is entirely browser-based and does not require installation, ensuring accessibility and responsiveness across devices. Key responsibilities of the front-end include:

- Initializing and rendering the map interface;
- Loading and displaying geospatial layers based on user selection;
- Handling user interactions such as indicator selection, filtering, and tooltip display;
- Managing multilingual content through dynamic language switching;
- Providing visual aids such as legends, colour scales, and pop-up summaries to improve interpretability.

In the this section, interface and a description of the layout and functionality of each component are detailed.

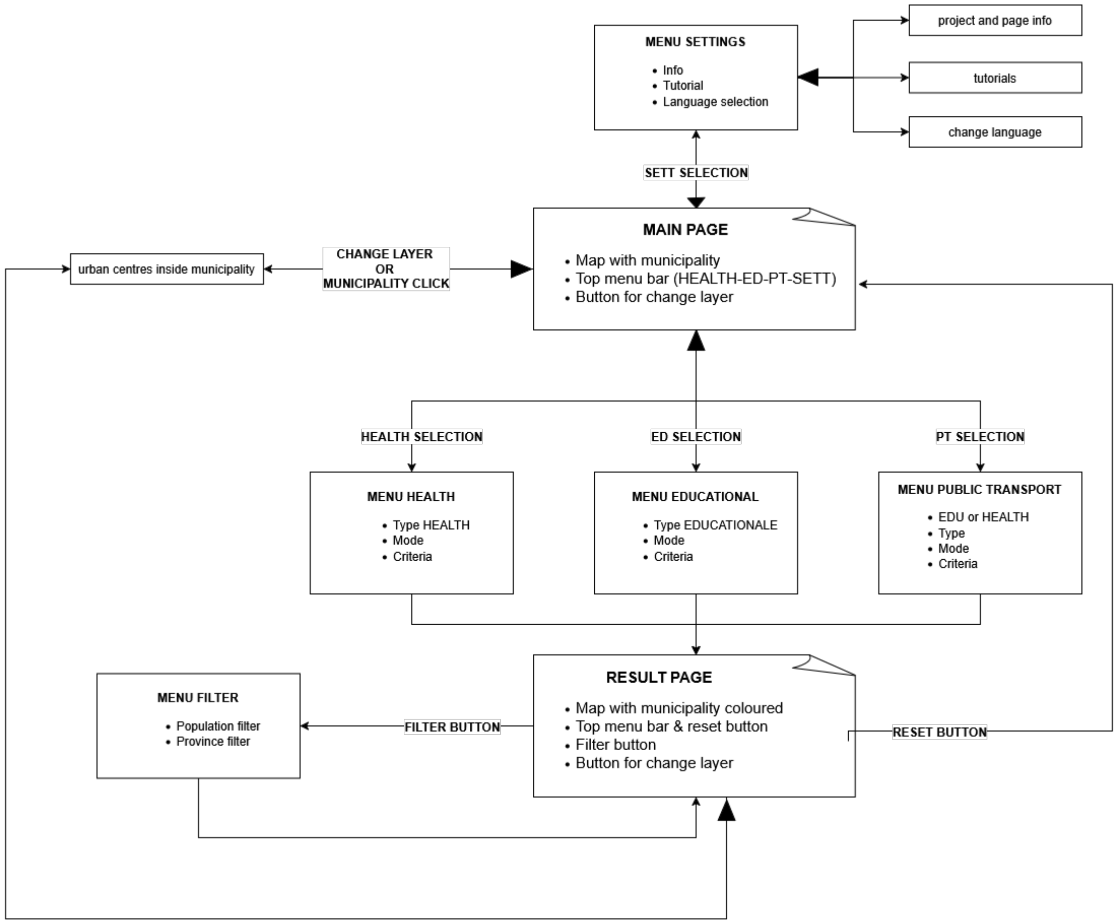


Figure 15: *Preliminary website navigation flow*

The web application is implemented as a single-page interface (commonly referred to as a Single Page Application, or SPA), meaning that all user interactions dynamically update the content of the page without requiring full page reloads. This approach improves responsiveness and enhances the user experience.

The home page, shown in figure 16, is the initial view presented to the user upon accessing the platform through a web browser. It displays a map centred on the Lazio region, with clearly outlined municipal boundaries.

The layout includes several interface elements that remain visible and accessible across all stages of interaction:

- A top navigation bar, providing consistent access to the main sections of the application for selecting which accessibility indicators to visualize;
- A settings menu, accessible via an icon located in the top-right corner of the interface in the top navigation bar;
- A radio button group positioned on the left side of the screen, allowing the user to toggle between viewing data at the municipality level or the more granular urban centre level.

This consistent structural layout ensures a smooth and intuitive navigation experience while allowing the content of the map and menus to update dynamically in response to user input.

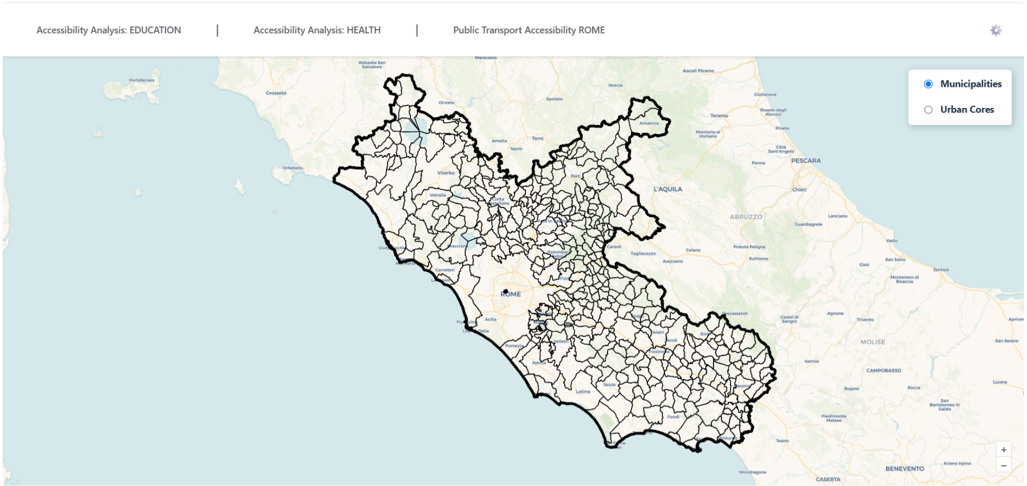
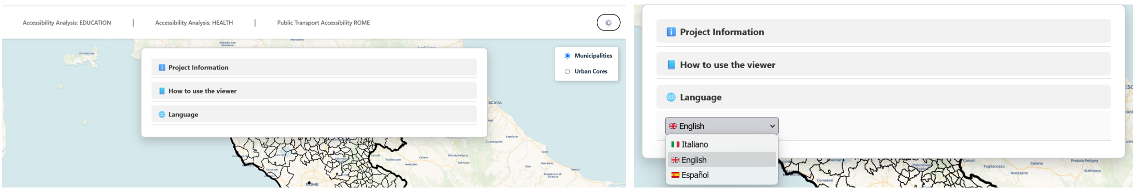


Figure 16: *Viewer design - home page*

By clicking on the information button, the user gains access to an expandable container offering three key options:

- Project information: A brief overview explaining the purpose of the application, the rationale behind its development, and a summary of the data sources and indicators displayed;
- How to use the viewer: A concise tutorial that assists users in understanding how to interact with the viewer, including navigation tips, indicator selection, and interpretation of results;
- Language: A panel allowing users to switch between the available interface languages: Italian, English, and Spanish (shown in figure 17b).



(a) *Setting panel options* (b) *Language change functionality*

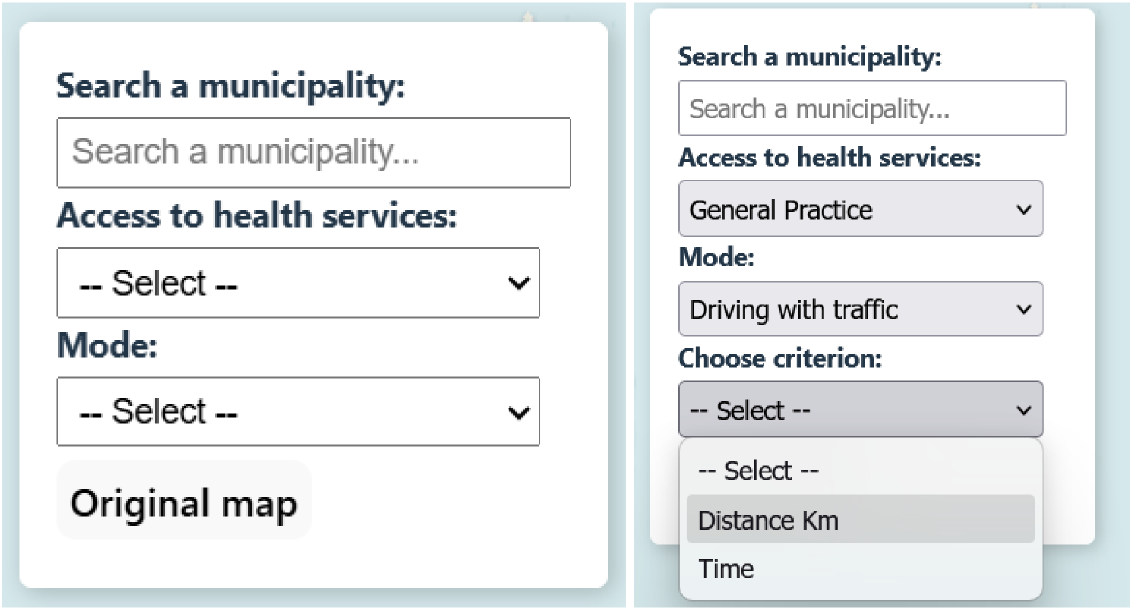
Figure 17: *Viewer design - settings panel*

This functionality enhances accessibility and transparency, offering contextual guidance and language flexibility for a broader audience.

Clicking on one of the macro-areas in the top navigation bar (e.g., Education or Healthcare) opens a contextual menu that allows users to configure the parameters of the indicator they wish to visualize. This menu includes several key components:

- A search field for locating a specific municipality. When the user types a municipality name and presses Enter, the map automatically zooms to the corresponding area.
- A reset button, which resets the map view and selected parameters to their default state.
- A step-by-step selection process, where input fields appear progressively based on previous selections (As shown in the corresponding figure 18b):
  - Service type selection, depending on the macro-area (e.g., General medicine, Hospitals, or Emergency Services);

 – Mode of access, such as driving (with traffic) or public transport;
 – Criteria selection, where the user can choose to display either distance in kilometres or travel time in minutes.



(a) *Menu* (b) *Cascade selection options*

Figure 18: *Viewer design - menu for choosing parameters of the indicators*

Overall, the interface follows a cascading logic: each new selection field becomes visible only after the previous one is filled, thus guiding the user in a structured and intuitive way through the indicator configuration process.

Once the user completes the selection of all required parameters, the map is automatically updated using a colour scale that visually represents the differences in accessibility values across municipalities, based on the selected criteria. During the result display phase, the interface includes two key visual aids to support interpretation:

- A legend, located in the bottom-left corner of the map, which explains the meaning of the colour gradient. The scale typically ranges from dark blue, representing the most favourable accessibility values, to dark red, indicating the least favourable ones. The legend also displays the minimum and maximum values observed in the current dataset.
- A hover information box, positioned next to the legend, which dynamically shows the exact accessibility value for the municipality currently under the user's cursor.

This combination of visual encoding and interactive feedback ensures that the accessibility data is both easily interpretable and contextually precise, facilitating immediate comparison across the region.

Since the analysis is fundamentally conducted at the urban centre level, the accessibility values displayed for each municipality in the main view are the result of an aggregation of the values computed for its constituent urban centres. To provide greater insight into the internal variability within municipalities, the application allows users to explore intra-municipal accessibility patterns.

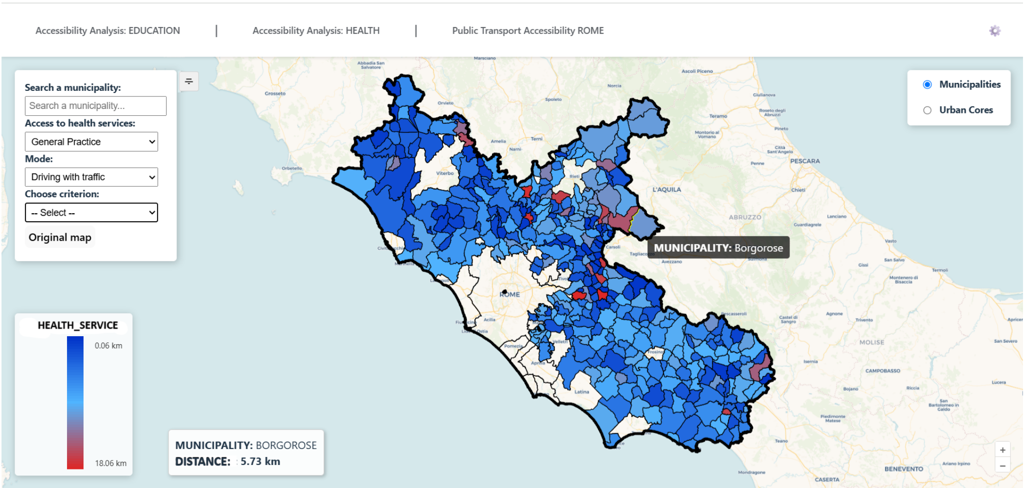


Figure 19: *Viewer design - results*

By simply clicking on a municipality, the interface dynamically updates to display its internal composition. The map 20 highlights the selected municipal boundary and visualizes the individual urban centres, each coloured according to their own accessibility value, using the same colour scale shown in the main legend.

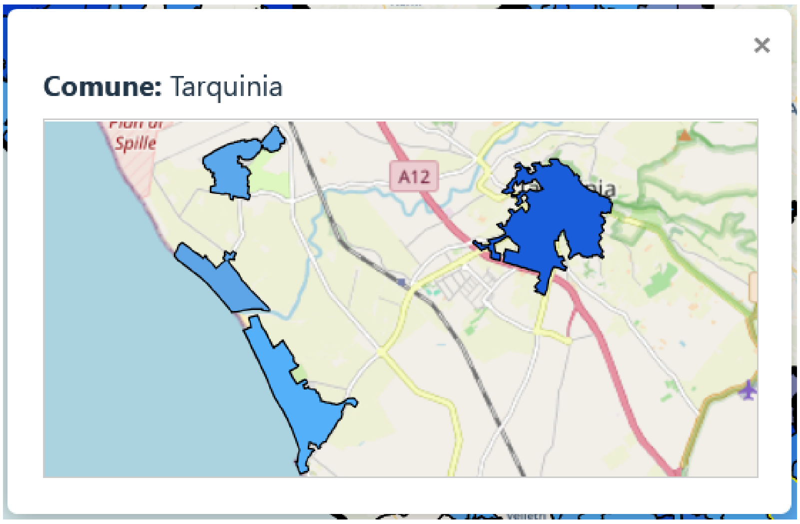


Figure 20: *Viewer design - urban centres visualizations results*

This feature allows users to understand the origin of the aggregated value and to identify whether a municipality's score is consistent across its territory or influenced by disparities among its internal areas.

In addition, the result page includes a filtering tool accessible via an icon located to the right of the menu bar. By clicking this icon, users can apply advanced filters to the displayed results. Specifically, they can:

- Set a population threshold (up to 40,000 inhabitants) to focus the analysis on smaller municipalities;
- Select one or more specific provinces, narrowing the view to a geographic subset of interest.

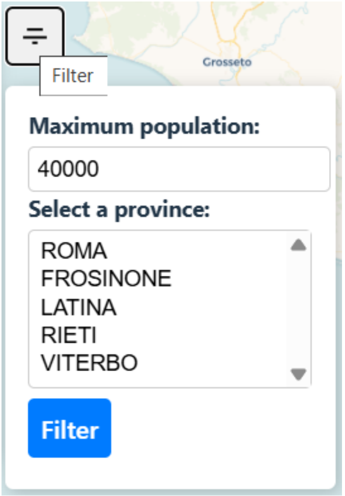


Figure 21: *Viewer design - filter menu*

These filters allow for targeted exploration of accessibility disparities across demographically or administratively defined areas, enhancing the analytical flexibility of the viewer.

3.5.2 Back-end development

The back-end component of the application is implemented using Flask, a lightweight Python web framework suitable for creating RESTful APIs and micro services. The final architecture of the web application is designed to delegate all data management to a dedicated micro service, ensuring better modularity, maintainability, and scalability.

The back-end acts as a data provisioning layer between the user interface and the source datasets (currently stored as structured CSV files). It exposes a set of HTTP endpoints that respond to parametrized requests from the front-end and return the appropriate dataset in JSON or file format.

- **Core responsibilities**:
  - Receive client requests containing input parameters such as:
    - Indicator type (e.g., EDUCATIONAL or HEALTH);
    - Service category (e.g., kindergarten, emergency care);
    - Access mode (e.g., driving, public transport);
    - Metric (e.g., distance or time);
  - Choose, filter and process the corresponding CSV dataset accordingly;
  - Return the filtered data to the client as csv.
- **Main used libraries**:
  - `flask`: to define routes, manage request and response cycles;
  - `flask_cors`: to allow cross-origin requests from the frontend;
  - `send_file` / `Response`: for sending data files to the front-end;
- **Integration with front-end**: The front-end application would send HTTP requests to the Flask server using JavaScript `fetch()`, passing the user-selected parameters as query strings. The Flask service would then return the corresponding dataset, which the front-end uses to render or update the map visualization.

- **Endpoint flow**: The following snippet illustrates the logic of the back-end endpoint responsible for dynamically returning filtered indicator data in CSV format, based on parameters received from the front-end.

```
@app.route('/get_data', methods=['GET'])
def get_data():
    # Receive input parameters
    indicator = request.args.get('indicator')    # e.g., 'hospital'
    mode = request.args.get('mode')              # 'driving' or 'public_transport'
    metric = request.args.get('metric')          # 'time' or 'distance'
    level = request.args.get('level')            # 'municipality' or 'urban centre'

    # Select appropriate filename
    if level == 'municipality':
        filename = f"aggregated_{indicator}_by_municipality.xlsx"
    elif level == 'urban centre':
        filename = f"aggregated_{indicator}_distances_weighted.xlsx"
    else:
        return jsonify({'error': 'Invalid level'}), 400

    # Read and filter data
    df = pd.read_excel(f"data/{filename}")
    selected_cols = [col for col in df.columns if metric in col.lower()
    or 'name' in col.lower()]
    filtered_df = df[selected_cols]

    # Return CSV response
    output = io.StringIO()
    filtered_df.to_csv(output, index=False)
    output.seek(0)

    return send_file(
        io.BytesIO(output.getvalue().encode()),
        mimetype='text/csv',
        as_attachment=True,
        download_name=f"{indicator}_{level}_{metric}.csv"
    )
```

3.5.3 Function development

The functionality of the platform is structured around three key user actions, which define the interactive behaviour of the interface: indicator selection, result filtering, and language switching. These functionalities are developed in the client side with JavaScript.

Indicator selection

When the user selects a combination of accessibility parameters (indicator, service type, access mode, metric, and geographic level) via the user interface, the application captures these values directly from the corresponding HTML elements. This process is handled using native JavaScript, which listens for changes in drop downs or confirmation buttons.

Once all required fields are selected, a JavaScript function formats these inputs into query parameters and sends a GET request to the back-end micro service via `fetch()`. The server, built with Flask, processes the request and returns the appropriate dataset in CSV format, which is then parsed and rendered on the map. This dynamic request-response cycle ensures that the displayed data always reflects the user's current configuration.

Result filtering

After the accessibility results are loaded and visualized on the map, the user has the option to apply post-processing filters to narrow down the view based on population size and administrative boundaries (provinces).

When the user selects a population threshold or specific provinces, the front-end, which already has access to local datasets containing population and province codes, applies the filter client-side, without making a new request to the server.

The filtering mechanism is implemented by temporarily hiding map features (e.g., municipalities or urban centres) that do not satisfy the selected criteria. Technically, this is achieved by setting the fill colour of excluded features to white, effectively removing them from the visual layer while preserving the rest of the map's structure and legend scale. This approach allows for fast, responsive filtering without reloading data or refreshing the map.

Language switching

The application supports multilingual content through a lightweight, modular internationalization (i18n) system implemented in JavaScript.

At startup, the application calls the `loadTranslations()` function, which fetches a JSON file `/translations.json` containing all language keys and values. The current language is loaded from local storage or defaults to Italian ("it"). The loaded translations are then applied to the interface via two functions:

- applyStaticTranslations(): Updates the inner content or placeholders of static elements such as buttons, titles, labels, and tooltips.
- applySelectTranslations(): Replaces the option labels within drop-down menus based on the selected language.

Language changes are triggered through a selection menu, which calls the setLanguage(lang) function. This updates both the internal state and the local storage value, and re-applies the translations to the page dynamically.

The t(key) utility function is used to retrieve the translated value for any key at runtime, optionally replacing placeholders in the text. This modular and declarative translation system allows the interface to support additional languages with minimal changes.

This chapter has provided a detailed overview of the architectural and functional design of the web application developed to visualize accessibility indicators. The system is structured around a single-page interface that integrates interactive mapping, dynamic data retrieval, filtering mechanisms, and multilingual support.

The architecture is composed of a modular front-end, responsible for rendering the interface and managing user interactions; a lightweight Flask-based back-end, designed to serve indicator data dynamically; and a set of core functionalities, including indicator selection, result filtering, and language switching, that define the interactive experience of the user.

Through the combination of structured client-server communication, flexible data handling, and a user-centric interface design, the application is able to deliver a responsive and intuitive environment for exploring complex spatial accessibility patterns.

In the next chapter, the results obtained through the application are presented and discussed, highlighting spatial differences and accessibility disparities across municipalities and urban centres.

4. Results

Following the methodology described in the previous chapters, a set of accessibility indicators was computed and structured in a format suitable for subsequent analysis. However, to ensure that the outcomes of this research are interpretable and actionable even by stakeholders without advanced statistical or technical backgrounds, a geospatial data viewer was developed. This tool was iteratively refined through multiple meetings and feedback sessions with urban planners and territorial management officials from the Community of Madrid, who provided critical insight into which visualizations and metrics would be most useful for practical urban and regional planning purposes.

This chapter presents the main findings of the project. The results are organised in such a way as to present the dataset of indicators obtained, which reflects the overall objective of the methodology. In addition, the chapter aims to present, on a case-by-case basis, the accessibility conditions obtained for the various public services in the contexts considered. The most part of the chapter focuses on aggregated data analyses, illustrating how accessibility metrics vary across space and transport modalities, and the kind of insights these variations reveal.

To complement this textual analysis, representative screenshots from the developed map viewer are included, alongside a selection of summary statistics and plots that illustrate key patterns and anomalies. While not all results are presented exhaustively within this chapter, complete data can be accessed via the public web viewer at http://lodcoremadrid-viewers.sel.inf.uc3m.es:8083/, and all statistical scripts and analytical notebooks are available in the dedicated results folder of the project repository at https://github.com/keving-ing/Distance-project/tree/main/Script/RESULTS.

4.1 Dataset description

In the previous chapter, the methodology adopted to compute the accessibility indicators was presented in detail. This section of the results focuses on the structure and content of the datasets generated through that process.

First, a comprehensive overview of each dataset is first provided, including the definition of the indicators and the structure of the associated data tables. The aim is to make the meaning and format of each field explicit, facilitating both interpretation and reuse. This chapter also includes the explication of metadata, in this context, refer to the structured descriptive information that characterizes each dataset or indicator. These include not only basic fields like title or unit of measure, but also semantic relationships and categorization elements that allow for automated interpretation, discovery, and reuse in open data portals and linked data ecosystems.

4.1.1 Indicator inventory

To support the structuring of metadata and facilitate integration with Linked Open Data standards, each accessibility indicator computed, are shown in this project with a unique internal identifier.

Table 12 provides the complete list of indicators. Each row corresponds to a distinct indicator defined by: domain (Education or Healthcare), service category mode of access (Driving or Public Transport) and metric type (Distance or Time). The Indicator ID is a numeric code starting from 3.91, following an internal naming convention adopted for this project. These identifiers will be also used as stable references in metadata definitions and RDF serialization.

While the inventory presented in table 12 lists all unique combinations of domain, service category, access mode, and metric, it is important to note that each indicator is computed at two spatial levels: municipality and urban centre level. This structured metadata system guarantees the find-ability, accessibility, interoperability, and reusability (FAIR) of the indicators published. Moreover, it enables their seamless integration into semantic web platforms through RDF serialization based on the DCAT standard and the LoDCOREMadrid ontology.

Table 12: Detailed inventory of accessibility indicators

Indicator ID	Domain	Service Category	Access Mode	Metric Type
3.91.1	Education	Kindergarten	Driving	Distance
3.91.2	Education	Kindergarten	Driving	Time
3.91.3	Education	Kindergarten	Public Transport	Distance
3.91.4	Education	Kindergarten	Public Transport	Time
3.92.1	Education	Primary School	Driving	Distance
3.92.2	Education	Primary School	Driving	Time
3.92.3	Education	Primary School	Public Transport	Distance
3.92.4	Education	Primary School	Public Transport	Time
3.93.1	Education	Secondary School	Driving	Distance
3.93.2	Education	Secondary School	Driving	Time
3.93.3	Education	Secondary School	Public Transport	Distance
3.93.4	Education	Secondary School	Public Transport	Time
3.99.1	Healthcare	General Practice	Driving	Distance
3.99.2	Healthcare	General Practice	Driving	Time
3.99.3	Healthcare	General Practice	Public Transport	Distance
3.99.4	Healthcare	General Practice	Public Transport	Time
3.100.1	Healthcare	Hospitals	Driving	Distance
3.100.2	Healthcare	Hospitals	Driving	Time
3.100.3	Healthcare	Hospitals	Public Transport	Distance
3.100.4	Healthcare	Hospitals	Public Transport	Time
3.101.1	Healthcare	Emergency Care	Driving	Distance
3.101.2	Healthcare	Emergency Care	Driving	Time
3.101.3	Healthcare	Emergency Care	Public Transport	Distance
3.101.4	Healthcare	Emergency Care	Public Transport	Time

4.1.2 Data structure

After defining the indicators and their publication strategy, the focus shifts to the actual datasets derived through the methodology described in the previous section. All computations were initially carried out and validated in memory as structured data frames, and subsequently exported in CSV format to ensure interoperability with analytical tools and open data platforms.

Education accessibility data indicators

The dataset related to Education accessibility comprises four CSV files, each representing a different combination of geographic aggregation and transportation mode:

- `aggregated_school_distances_weighted.csv`
- `aggregated_school_by_municipality.csv`
- `aggregated_school_distances_transit_weighted.csv`
- `aggregated_school_transit_by_municipality.csv`

Each file contains indicator values for access to three education levels:

- SI: Kindergarten
- SP: Primary school
- SS: Secondary school

The files `aggregated_school_distances_weighted.csv` and `aggregated_school_distances_transit_weighted.csv` contain accessibility data aggregated at the urban centre level.

Each row corresponds to a single urban centre within a municipality, and includes both distance (km) and travel time (min) indicators, along with their respective standard deviations. The structure of these files is as follows:

Table 13: Structure of urban centres-level education accessibility datasets

Column Name	**Description**
Municipality	Name of the municipality
UrbanCentre_ID	Unique identifier of the urban centre
Population	Population of the urban centres
SI_mean_km	Mean distance to kindergarten
SI_St.Dv_km	Std. deviation of distance to kindergarten
SI_mean_min	Mean travel time to kindergarten
SI_St.Dv_min	Std. deviation of time to kindergarten
SP_mean_km	Mean distance to primary school
SP_St.Dv_km	Std. deviation of distance to primary school
SP_mean_min	Mean travel time to primary school
SP_St.Dv_min	Std. deviation of time to primary school
SS_mean_km	Mean distance to secondary school
SS_St.Dv_km	Std. deviation of distance to secondary school
SS_mean_min	Mean travel time to secondary school
SS_St.Dv_min	Std. deviation of time to secondary school

The files `aggregated_school_by_municipality.csv` and `aggregated_school_transit_by_municipality.csv` represent the same indicators but aggregated at the municipality level. In this case, each row corresponds to a municipality, and the indicators are calculated as population-weighted averages of the values of its urban centres. The columns are nearly identical, except that the UrbanCentre_ID is removed, and the field Total_population represents the sum of populations across all urban centres in the municipality.

Table 14: Structure of municipality-level education accessibility datasets

Column Name	**Description**
Municipality	Name of the municipality
Total_population	Total population of the municipality
SI_mean_km	Mean distance to preschool
SI_St.Dv_km	Std. deviation of distance to kindergarten
SI_mean_min	Mean travel time to kindergarten
SI_St.Dv_min	Std. deviation of time to kindergarten
SP_mean_km	Mean distance to primary school
SP_St.Dv_km	Std. deviation of distance to primary school
SP_mean_min	Mean travel time to primary school
SP_St.Dv_min	Std. deviation of time to primary school
SS_mean_km	Mean distance to secondary school
SS_St.Dv_km	Std. deviation of distance to secondary school
SS_mean_min	Mean travel time to secondary school
SS_St.Dv_min	Std. deviation of time to secondary school

Healthcare accessibility data indicators

For healthcare services, the accessibility indicators are organized into separate files for each service category:

- General practice (doctors)
- Hospitals
- Emergency care

Unlike the education dataset, where multiple categories are combined into single files, here each service type is stored in its own set of files, grouped by transportation mode and aggregation level. Each category includes the following four files:

- aggregated_<category>_distances_weighted.csv
- aggregated_<category>_by_municipality.csv
- aggregated_<category>_distances_transit_weighted.csv
- aggregated_<category>_transit_by_municipality.csv

Where <category> is one of:

- doctors → General practitioners
- hospital → Hospitals
- emergency → Emergency services

Each file contains accessibility indicators either by urban centres or by municipality, and distinguishes between private vehicle (driving) and public transport modes. Files ending with _distances_weighted.csv or _distances_transit_weighted.csv report data at the urban centres level. Each row corresponds to a specific urban centre within a municipality.

Table 15: Structure of urban centres-level healthcare accessibility datasets

Column Name	**Description**
Municipality	Name of the municipality
Urban_centre_ID	Unique identifier of the urban centre
Population	Number of residents in the urban centre
mean_km	Mean distance to the selected healthcare service
St.Dv_km	Standard deviation of distance
mean_min	Mean travel time to the service
St.Dv_min	Standard deviation of travel time

Files ending with _by_municipality.csv or _transit_by_municipality.csv contain the same metrics aggregated by municipality, based on weighted averages of the constituent urban centres.

Table 16: Structure of municipality-level healthcare accessibility datasets

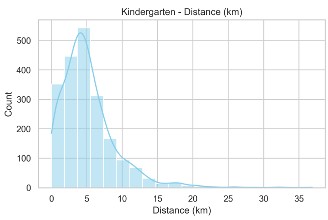
Column Name	**Description**
Municipality	Name of the municipality
Total_population	Total population of the municipality
mean_km	Mean distance to the selected healthcare service
St.Dv_km	Standard deviation of distance
mean_min	Mean travel time to the service
St.Dv_min	Standard deviation of travel time

This chapter has provided a structured overview of the datasets generated through the methodology described earlier, focusing on accessibility indicators across the education and healthcare domains. The datasets were presented in CSV format, with clearly defined structures depending on the level of geographic aggregation (municipality or urban centre) and the transportation mode (private vehicle or public transport).

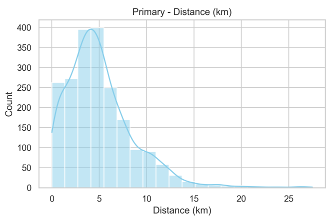
4.2 Education accessibility results

The first stage of the indicator analysis focuses on the accessibility to educational services, starting from the data aggregated at the level of urban centres. This level of granularity provides a raw and spatially detailed representation of accessibility, based on direct distance and travel time to the nearest available educational facilities. Such indicators are particularly valuable for identifying intra-municipal disparities and for capturing localized accessibility gaps that may not emerge when aggregating data at broader administrative levels. For each urban centres, the following basic indicators were calculated: distance to the nearest school (in kilometres), travel time to the nearest school (in minutes), for each of the three education levels considered: kindergarten, primary, and secondary education.

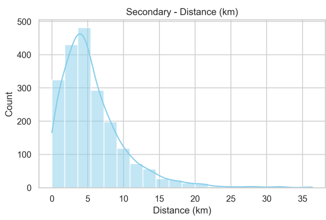
Before interpreting spatial patterns or drawing comparisons across territories, it is essential to examine the distribution of these indicators. This section presents the empirical distributions of the raw accessibility indicators (distance and time) across all 2,521 urban centres in the Lazio region. These distributions provide an initial understanding of the variability in access to educational services and allow for the identification of outliers and areas with critical accessibility limitations. The results are illustrated through graphs, offering a comparative view of how accessibility varies between kindergarten, primary, and secondary education.



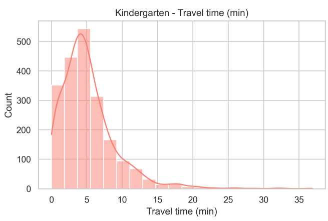
(a) *Kindergarten – Distance (km)*



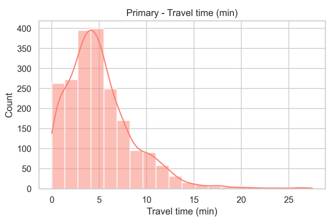
(b) *Primary – Distance (km)*



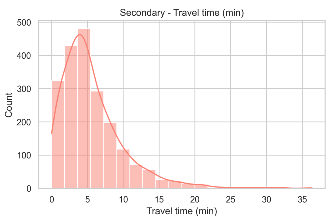
(c) *Secondary – Distance (km)*



(d) *Kindergarten – Time (min)*



(e) *Primary – Time (min)*



(f) *Secondary – Time (min)*

Figure 22: *Distribution of average distances for each school services at the urban centre level*

Table 17: Summary statistics of school accessibility – Distance (km)

Level	Mean	Std Dev	Min	25%	Median	Max
Kindergarten	5.19	3.91	0.002	2.70	4.47	36.87
Primary	5.08	3.59	0.002	2.67	4.46	27.48
Secondary	5.69	4.52	0.001	2.73	4.67	36.36

Table 18: Summary statistics of school accessibility – Travel time (min)

Level	**Mean**	**Std Dev**	**Min**	**25%**	**Median**	**Max**
Kindergarten	8.14	5.28	0.01	4.82	7.40	56.85
Primary	8.01	4.94	0.01	4.82	7.32	42.68
Secondary	8.76	5.70	0.01	4.97	7.78	56.43

From the statistical summaries, it can be observed that primary schools exhibit the lowest average distance and time, suggesting slightly better accessibility in comparison to kindergarten and secondary schools. Notably, secondary schools present the highest average distance (5.69 km) and travel time (8.76 minutes), along with the widest spread, as indicated by the standard deviation. These results are consistent with the fact that secondary education facilities tend to be more centralized, typically located in larger urban areas.

Table 19: Top and bottom 5 urban centres by distance to nearest kindergarten (km)

Best 5 (Shortest Distance)		**Worst 5 (Longest Distance)**	
Municipality (ID)	**Dist. (km)**	**Municipality (ID)**	**Dist. (km)**
PROSSEDI (5902)	0.002	COLLALTO SABINO (5701)	36.868
BOMARZO (5600)	0.011	NESPOLO (5704)	32.361
POGGIO BUSTONE (5705)	0.011	CAMERATA NUOVA (5801)	31.990
BARBARANO ROMANO (5600)	0.016	TURANIA (5707)	28.623
AUSONIA (6001)	0.019	COLLALTO SABINO (5701)	26.411

Table 20: Top and bottom 5 urban centres by travel time to nearest kindergarten (min)

Best 5 (Shortest Time)		**Worst 5 (Longest Time)**	
Municipality (ID)	**Time (min)**	**Municipality (ID)**	**Time (min)**
POGGIO BUSTONE (5705)	0.001	COLLALTO SABINO (5701)	56.850
PROSSEDI (5902)	0.001	NESPOLO (5704)	47.783
BARBARANO ROMANO (5600)	0.017	MARCETELLI (5703)	40.400
BOMARZO (5600)	0.017	COLLALTO SABINO (5701)	39.350
AUSONIA (6001)	0.033	COLLEGIOVE (5702)	38.700

Table 21: Top and bottom 5 urban centres by distance to nearest primary school (km)

Best 5 (Shortest Distance)		**Worst 5 (Longest Distance)**	
Municipality (ID)	**Dist. (km)**	**Municipality (ID)**	**Dist. (km)**
PROSSEDI (5902)	0.002	STIMIGLIANO (6603)	32.482
CORCHIANO (5602)	0.023	STIMIGLIANO (0001)	29.324
TORRICELLA SABINA (5706)	0.025	PESCOROCCHIANO (5704)	26.391
GALLINARO (6004)	0.054	TOLFA (5810)	25.734
BELLEGRA (5801)	0.054	MARCETELLI (5703)	24.111

Table 22: Top and bottom 5 urban centres by travel time to nearest primary school (min)

Best 5 (Shortest Time)		**Worst 5 (Longest Time)**	
Municipality (ID)	**Time (min)**	**Municipality (ID)**	**Time (min)**
PROSSEDI (5902)	0.001	STIMIGLIANO (5706)	42.683
TORRICELLA SABINA (5706)	0.033	PESCOROCCHIANO (5704)	41.475
CORCHIANO (5602)	0.050	STIMIGLIANO (5706)	38.883
GALLINARO (6004)	0.133	MARCETELLI (5703)	38.067
COLLE SAN MAGNO (6002)	0.167	VARCO SABINO (5707)	37.567

Table 23: Top and bottom 5 urban centres by distance to nearest secondary school (km)

Best 5 (Shortest Distance)		**Worst 5 (Longest Distance)**	
Municipality (ID)	**Dist. (km)**	**Municipality (ID)**	**Dist. (km)**
CANALE MONTERANO (5801)	0.001	COLLALTO SABINO (5701)	36.357
ARSOLI (5801)	0.025	COLLEGIOVE (5702)	35.952
CARPINETO ROMANO (5802)	0.025	CAMERATA NUOVA (5801)	31.990
GAVIGNANO (5804)	0.032	MONTE S.G. IN SABINA (5704)	31.895
VEJANO (5605)	0.041	NESPOLO (5704)	31.850

Table 24: Top and bottom 5 urban centres by travel time to nearest secondary school (min)

Best 5 (Shortest Time)		**Worst 5 (Longest Time)**	
Municipality (ID)	**Time (min)**	**Municipality (ID)**	**Time (min)**
CANALE MONTERANO (5801)	0.001	COLLALTO SABINO (5701)	56.433
CARPINETO ROMANO (5802)	0.050	COLLEGIOVE (5702)	55.500
GAVIGNANO (5804)	0.067	NESPOLO (5704)	47.350
ARSOLI (5801)	0.117	MARCETELLI (5703)	39.033
ARNARA (6000)	0.183	COLLALTO SABINO (5701)	38.950

A ranking of urban centres was compiled based on their average accessibility to the nearest school facility, both in terms of distance and travel time, for each educational level. From the results, several trends can be observed:

- Prossedi and Poggio Bustone consistently appear among the top-ranked urban centres with shortest distances and travel times, particularly for kindergarten and primary education, suggesting excellent local availability of services.
- In contrast, Collalto Sabino appears multiple times among the bottom-ranked urban centres, with the highest distances and travel times across all educational levels. This suggests severe accessibility limitations, likely due to its geographic isolation or lack of nearby infrastructure.
- Secondary schools show the greatest variation, with several urban centres exceeding 30 km or 50 minutes of travel time, highlighting the centralization of secondary education facilities and the resulting spatial disparities.
- Several urban centres (e.g., Marcetelli, Nespolo, Collegiove) appear repeatedly in the bottom rankings, identifying them as priority areas for potential educational service enhancement.

These rankings not only help validate the computed indicators but also provide actionable insights for regional policy-making aimed at improving equitable access to educational services.

4.2.1 Population in education accessibility

After illustrating the distribution of distances and access times for each school level, it is useful to investigate which factors may influence these variations. One of the most frequently analysed aspects in the literature is the relationship between population size and accessibility to services. For this reason, a correlation analysis was conducted between the number of inhabitants in each urban centre and the average distance/time to reach schools.

To investigate the relationship between population size and accessibility to educational services, the Pearson correlation coefficient between the population of each urban centre and its average distance/time to reach schools was calculated. Particularly it was calculated separately for each school level (Kindergarten, Primary, and Secondary) and for both distance in kilometres and travel time in minutes. This helps to verify whether more populated areas generally enjoy better accessibility (i.e., shorter distances or times to schools), as commonly hypothesized in accessibility studies.

Table 25: Pearson correlation between population size and average accessibility to schools

School Level	**Accessibility Measure**	**Pearson r**	**p-value**
Kindergarten	Distance (km)	-0.128	4.69×10^{-9}
Kindergarten	Time (min)	-0.096	1.20×10^{-5}
Primary	Distance (km)	-0.126	9.07×10^{-9}
Primary	Time (min)	-0.100	5.36×10^{-6}
Secondary	Distance (km)	-0.147	1.92×10^{-11}
Secondary	Time (min)	-0.121	3.13×10^{-8}

All correlation coefficients are negative, indicating that higher population is weakly associated with better accessibility (i.e., shorter distances and travel times). While the correlations are relatively low in magnitude (ranging from -0.096 to -0.147), they are statistically significant, as indicated by the very low p-values. The strongest correlation is for Secondary schools (distance) with $r = -0.147$, suggesting that less populated urban centres tend to be slightly farther away from secondary schools. These results align with the general accessibility theory: urban centres with low population densities tend to have fewer nearby services, and thus worse accessibility.

4.2.2 Traffic in education accessibility

To investigate whether travel distance and travel time show consistent behaviour across different school levels, the data were reshaped into long format by melting the distance and time columns separately. The reshaped tables were then merged to align distance and time values for each urban centre and school level. After filtering out any missing values, the Pearson correlation coefficient was calculated to quantify the overall linear relationship between distance (in kilometres) and time (in minutes). The scatter plot 23 visually summarizes the relationship between distance and travel time across all urban centres and school levels.

It shows a strong positive linear relationship between distance and time, as expected. This is confirmed by the overall Pearson correlation coefficient, `r = 0.948 (p-value = 0.012e+00)`. The black trend line represents the best linear fit across all school levels and shows a consistent increase in time as distance grows. Also some outliers can be observed, a few urban centres exhibit unusually high travel times even for moderate distances (e.g., 10–15 km resulting in 30+ minutes), those points may indicate infrastructure or traffic problems in certain urban areas. Others show unexpectedly short travel times for relatively long distances, possibly due to higher-speed road infrastructure or data inaccuracies. These anomalies may indicate local infrastructure disparities, limitations in estimating travel times, or the presence of particularly remote or poorly connected areas. It is therefore important to pay attention to these critical points that appear as outliers.

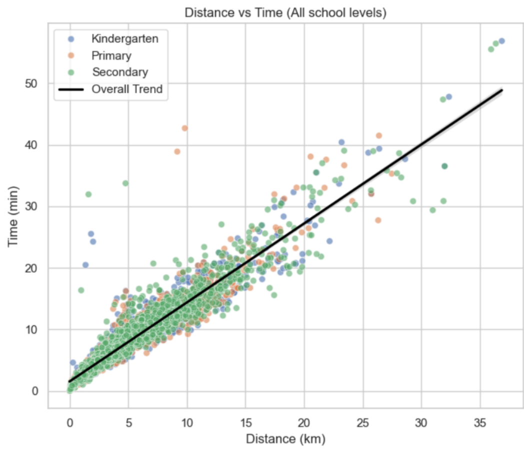


Figure 23: *Relationship between distance and travel time in education*

This was done by finding the points with the maximum residual, i.e. those for which the observed time is much higher or lower than that predicted by the regression.

Table 26: Top 15 urban centres with the highest positive residuals (time vs distance)

Municipality	**UrbanCentre**	**School Level**	**Distance (km)**	**Residual (min)**
STIMIGLIANO	5706627002	Primary	9.83	+28.53
STIMIGLIANO	5706610001	Secondary	1.61	+28.33
STIMIGLIANO	5706620001	Secondary	4.77	+26.07
STIMIGLIANO	5706610001	Primary	9.18	+25.56
STIMIGLIANO	5706620001	Kindergarten	1.83	+21.63
STIMIGLIANO	5706627002	Kindergarten	2.01	+20.14
STIMIGLIANO	5706610001	Kindergarten	1.36	+17.18
STIMIGLIANO	5706627002	Secondary	0.99	+13.55
MARCETELLI	5703610001	Primary	20.56	+10.15
CIAMPINO	5811810001	Secondary	7.14	+10.00
MARCETELLI	5703610001	Kindergarten	23.20	+9.09
SONNINO	5902927001	Secondary	12.95	+8.89
GROTTAFERRATA	5804627001	Kindergarten	4.79	+8.59
GROTTAFERRATA	5804610001	Primary	4.81	+8.50
GROTTAFERRATA	5804627001	Kindergarten	4.04	+8.32

The residuals between observed time and predicted time from distance shown in the table are extremely useful for identifying anomalous urban centres, i.e. positive residuals (time higher than predicted) are possible signs of poor road infrastructure, winding or slow roads, or a lack of direct connections.

Among the main outliers in the table are: Stimigliano, which dominates the ranking with seven entries across all school levels. This suggests a critical infrastructure condition or a particularly disadvantageous geographical configuration. The times are more than double what would be expected given the distance travelled. Marcetelli shows high values with distances of around 20 km but times of over 35 minutes: this reinforces the hypothesis of a lack of fast roads or indirect routes. Grottaferrata, although not a mountain municipality, also shows positive residual values for short distances (4–5 km), suggesting possible critical issues due to urban traffic, traffic lights or complex topography. Sonnino is an interesting case to be verified in the field to understand whether times are influenced by non-linear factors (hills, bends, physical barriers).

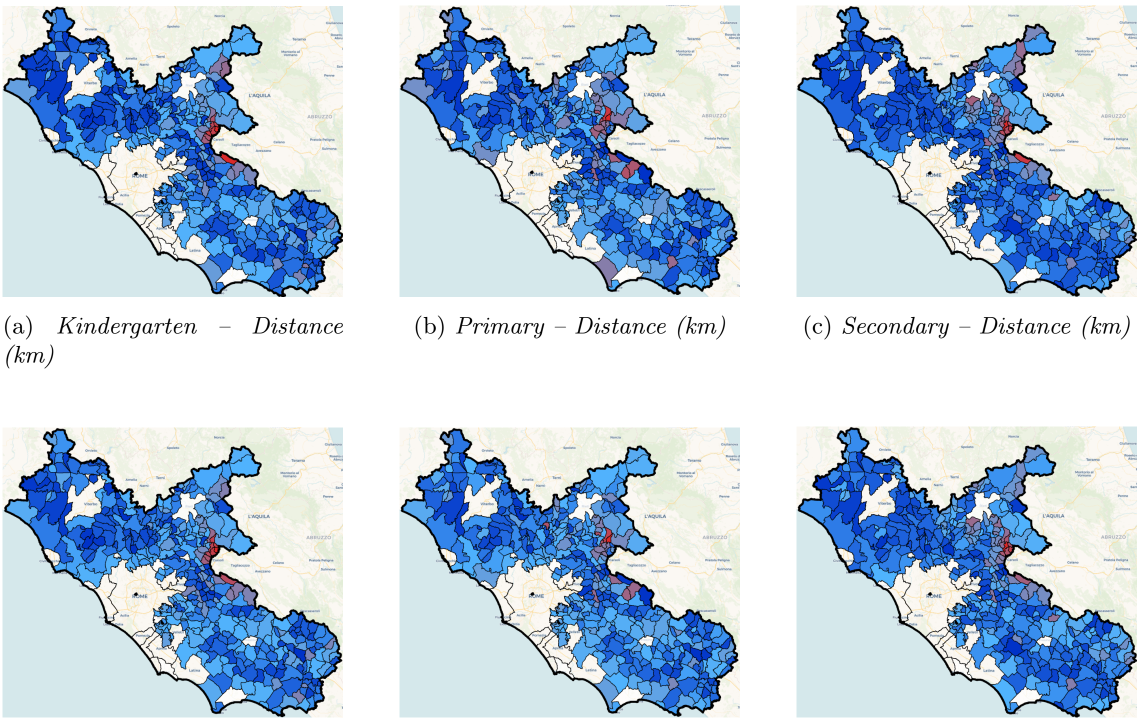
4.2.3 Accessibility education results aggregated for municipality

The objectives of the next showed results are to evaluate whether aggregating accessibility indicators at the municipality level preserves the same ranking trends observed at the urban centre level. This helps assess whether the aggregation masks important disparities.

After ranking accessibility at the urban centres level, the same analysis was replicated at the municipal level to assess whether trends observed in local areas are reflected at the broader administrative scale. This comparison is useful to evaluate whether low accessibility in specific urban centres significantly affects the overall performance of the municipality, or if spatial heterogeneity within the municipality compensates for local deficiencies. From the results, it is evident that some municipalities consistently perform well across both levels. For example:

- Corchiano and Colle San Magno appear in the top 5 for kindergarten and primary accessibility, both at the urban centre and municipal level, indicating good spatial equity within their territories.
- On the contrary, Collalto Sabino, Nespolo, and Collegiove emerge as critical cases in both classifications, with extremely high distances and travel times, confirming the presence of structural isolation.

However, in some cases, discrepancies arise. Rocca Santo Stefano appears among the best municipalities but does not appear in the urban centre level top ranks, suggesting that overall good accessibility is driven by an average performance across all its urban centres, rather than the presence of exceptionally well-connected ones. This comparison highlights the importance of multi-scale analysis: relying solely on municipal averages may mask intra-municipal disparities, while focusing only on urban centres may overlook systemic patterns. The results of the distance and time indicators aggregated by municipality can also be seen in image 24.



(a) *Kindergarten – Distance (km)* (b) *Primary – Distance (km)* (c) *Secondary – Distance (km)*

(d) *Kindergarten – Time (min)* (e) *Primary – Time (min)* (f) *Secondary – Time (min)*

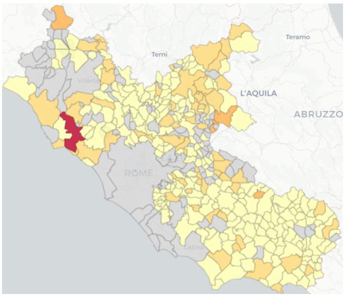
Figure 24: *Distance for each school services at the municipality level - viewer visualization*

While aggregated accessibility indicators at the municipal level provide a useful overview, they can obscure significant internal variations among the urban centres within the same municipality.

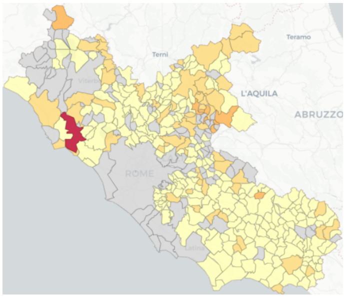
To better understand inequities that are spatially localized, it is essential to go beyond averages and investigate how accessibility is distributed internally.

The intra-municipal disparities in accessibility were assessed by analysing the standard deviation of distances and travel times between urban centres within each municipality. High standard deviation values highlight cases where some urban centres have excellent access to services, while others are significantly disadvantaged. These municipalities exhibit the greatest internal inequalities and may require targeted interventions at the sub-municipal level. The analysis of intra-municipal disparities was conducted by identifying the municipalities with the highest standard deviation of accessibility indicators among their constituent urban centres. A high standard deviation in either travel distance or time indicates that some urban centres within the same municipality enjoy good accessibility to education services, while others are significantly disadvantaged.

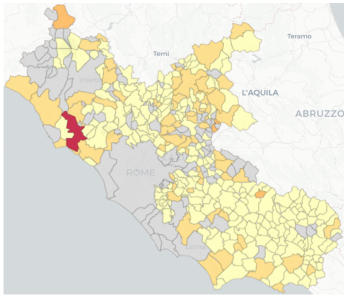
These findings are crucial for fine-grained policy design, as they point to municipalities where inequity is internal rather than between municipalities. Such disparities are often masked when relying only on aggregated values.



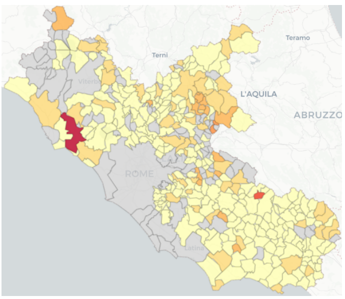
(a) *Kindergarten – St.Dv Distance (km)*



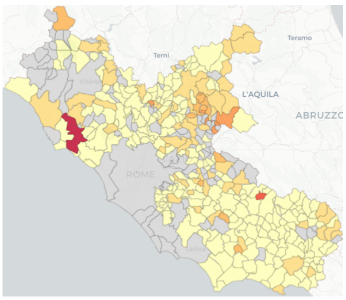
(b) *Primary – St.Dv Distance (km)*



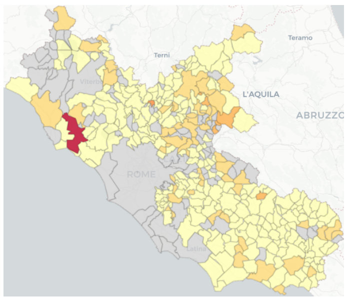
(c) *Secondary – St.Dv Distance (km)*



(d) *Kindergarten – St.Dv Time (min)*



(e) *Primary – St.Dv Time (min)*



(f) *Secondary – St.Dv Time (min)*

Figure 25: *Intra-Municipal disparities based on standard deviation - educational services*

The analysis of intra-municipal disparities reveals a number of municipalities where accessibility to educational services varies dramatically between urban centres within the same administrative boundary. The following insights summarize the most significant findings:

- Tolfa stands out as the municipality with by far the highest internal disparity, both in terms of distance and time, across all school levels. For secondary school accessibility, the standard deviation reaches nearly 19 km and 25 minutes, indicating that while some parts of the municipality enjoy direct access, others are highly isolated.
- Collepardo consistently ranks among the top 3 municipalities with the highest internal disparities, for both kindergarten and secondary education. This suggests a fragmented spatial distribution of services that disproportionately affects some areas.

- Collalto Sabino, already identified as one of the municipalities with the worst overall accessibility, also suffers from high intra-municipal disparity, particularly in secondary school access. This reinforces its position as a critical area for policy intervention.
- Affile, Rocca di Papa, and Cevara di Roma also emerge in multiple rankings, highlighting that some of their urban centres are well-served while others are substantially disadvantaged, a situation that could be masked by municipal averages.
- Stimigliano and Pescorocchiano, already present among the top residual outliers in the distance–time correlation, are confirmed here as internally inconsistent, suggesting possible infrastructural or topographic challenges affecting service accessibility differently across their territory.

These results are particularly valuable for fine-grained territorial planning, as they indicate that the problem is not just which municipalities are disadvantaged, but where within them accessibility gaps exist. Spatial targeting of resources within these municipalities could reduce inequalities far more effectively than broad administrative strategies. In fact, through the viewer, administrators or those in charge can see exactly which urban areas are the most problematic, even within municipalities with greater disparities, and perhaps even see their population.

4.2.4 Accessibility education aggregated for province

Having explored disparities both between municipalities and within them, the next step of the analysis scales up to the provincial level to evaluate broader territorial patterns. While intra-municipal disparities reveal localized inequities, a provincial comparison allows for a more systemic understanding of how accessibility is distributed across the region.

The provincial ranking of accessibility was calculated using population-weighted averages of municipality-level indicators. This approach ensures that larger municipalities contribute proportionally to the overall score, reflecting their demographic importance.

Additionally, the standard deviation of accessibility values between municipalities within each province highlights inter-municipal inequality. Higher values indicate provinces where access to education services varies significantly from one municipality to another, which may reflect uneven spatial distribution of facilities or differing levels of infrastructural development. This aggregated analysis supports strategic policy-making by identifying not only which provinces have lower average accessibility, but also how consistently services are distributed within each province.

Table 27: Kindergarten – Provincial accessibility and inter-municipal disparities

Province	**Weighted Mean (km)**	**Weighted Mean (min)**	**Std Dev (km)**	**Std Dev (min)**
VITERBO	1.83	3.35	1.98	2.49
ROMA	3.10	6.29	4.80	6.66
FROSINONE	3.20	6.03	2.74	4.18
LATINA	4.00	7.07	2.95	4.04
RIETI	5.17	8.49	6.99	10.11

Table 28: Primary School – Provincial accessibility and inter-municipal disparities

Province	**Weighted Mean (km)**	**Weighted Mean (min)**	**Std Dev (km)**	**Std Dev (min)**
VITERBO	1.85	3.40	2.28	2.85
FROSINONE	3.02	5.78	1.63	2.73
ROMA	3.03	5.95	3.02	4.56
LATINA	4.33	7.23	2.66	3.21
RIETI	6.59	10.00	4.37	7.48

Table 29: Secondary School – Provincial accessibility and inter-municipal disparities

Province	Weighted Mean (km)	Weighted Mean (min)	Std Dev (km)	Std Dev (min)
VITERBO	1.63	3.26	2.32	2.98
ROMA	2.69	5.81	5.06	6.94
FROSINONE	2.96	5.91	3.28	5.17
LATINA	3.71	6.98	3.02	4.82
RIETI	7.87	11.30	7.84	11.19

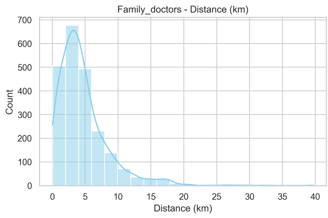
The results clearly show that the province of Rieti has the highest average values for distance and time for all school levels, accompanied by the highest standard deviations. This suggests a more problematic and uneven accessibility situation, where differences between municipalities are particularly marked. In contrast, Viterbo has the lowest averages and relatively low deviations, indicating good accessibility and territorial uniformity. Rome, despite being a central and urbanised province, shows high internal disparities, especially in secondary schools, with a standard deviation of over 5 km and 6.9 minutes. Latina and Frosinone are in an intermediate position, with moderate values in both weighted averages and dispersion, although Frosinone stands out for its fairly high internal homogeneity, especially in primary schools. These data highlight the provinces where rebalancing policies should be implemented on a provincial scale (e.g. Rieti), but also those where targeted intra-provincial interventions are needed (e.g. Rome).

Table 30: Pearson correlation between population size and average accessibility to schools

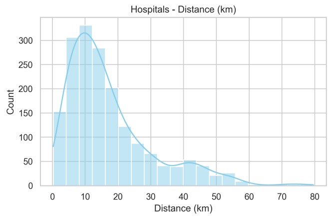
School Level	Accessibility Measure	Pearson r	p-value
Kindergarten	Distance (km)	-0.128	4.69×10^{-9}
Kindergarten	Time (min)	-0.096	1.20×10^{-5}
Primary	Distance (km)	-0.126	9.07×10^{-9}
Primary	Time (min)	-0.100	5.36×10^{-6}
Secondary	Distance (km)	-0.147	1.92×10^{-11}
Secondary	Time (min)	-0.121	3.13×10^{-8}

4.3 Healthcare accessibility results

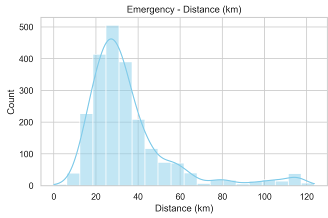
Up to this point, the accessibility analysis of the results has focused exclusively on the education system, considering the three school levels (kindergarten, primary and secondary) and assessing the average distance and time required to reach services, intra-municipal inequalities and provincial differences. Starting from this section, the same analytical methodology is applied to the healthcare context, with a particular focus on accessibility to hospitals and emergency services. The aim is to, compare access dynamics between different sectors (education vs. healthcare), identify any specific critical issues in access to healthcare services and highlight areas of the territory with the most significant shortcomings in terms of equity and distance from service delivery points. The results will be structured using the same approach adopted for the education sector, so as to ensure consistency in interpretation and ease of comparison between services.



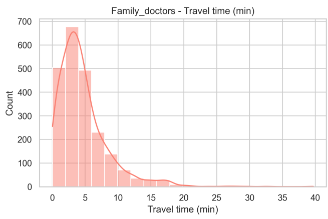
(a) *Kindergarten – Distance (km)*



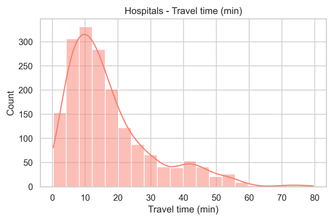
(b) *Primary – Distance (km)*



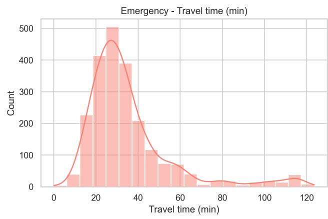
(c) *Secondary – Distance (km)*



(d) *Kindergarten – Time (min)*



(e) *Primary – Time (min)*



(f) *Secondary – Time (min)*

Figure 26: *Distribution of average distances for each healthcare services at the urban centre level*

Distance (km) and time (min) statistics for each health service type: Family Doctors, Hospitals, and Emergency Services are showed.

Table 31: Descriptive statistics for distance to health services (km)

Service	**Mean**	**Std. Dev.**	**Min**	**25%**	**50%**	**75%**	**Max**
Family Doctors	4.66	3.83	0.03	2.18	3.81	5.94	39.68
Hospitals	17.29	13.10	0.28	8.10	13.63	22.26	79.55
Emergency	35.26	20.48	0.06	23.12	29.93	39.67	123.62

Table 32: Descriptive statistics for average time to health services (min)

Service	**Mean**	**Std. Dev.**	**Min**	**25%**	**50%**	**75%**	**Max**
Family Doctors	7.29	5.14	0.08	4.09	6.37	9.11	60.00
Hospitals	20.92	12.73	0.90	12.10	17.75	26.44	74.75
Emergency	38.14	17.74	0.18	27.66	33.81	43.05	118.19

As expected, the distance and time increase from local services (family doctors) to more centralized (metropolitan) services. Family Doctors accessibility is very good, with a median travel distance of $\tilde{3}$.8 km and time $\tilde{6}$.4 min, comparable to or even better than kindergarten accessibility. Definitely from the results family doctors services are generally more accessible than any school level, reinforcing their distributed nature. Hospitals, roughly 4–5 times farther on average than family doctors, with much wider dispersion. Still, median values ($\tilde{1}$3.6 km, $\tilde{1}$7.8 min) suggest reasonably centralized coverage in many municipalities. While emergency services are are the least accessible, with average distances 35 km and times 38 min. The range is also the widest (max $>$ 120 km / 118 min) (fig. 27). This reflects the concentration of emergency infrastructure in major hubs only.

The map clearly highlights a spatial gradient in accessibility to emergency healthcare services across the Lazio region. It visualizes the average driving distance (with traffic) to the nearest emergency facility from each municipality. This map (figure 27) provides strong visual evidence that distance to emergency services increases as one moves away from Rome, with a notable concentration of disadvantage in rural and mountainous municipalities.

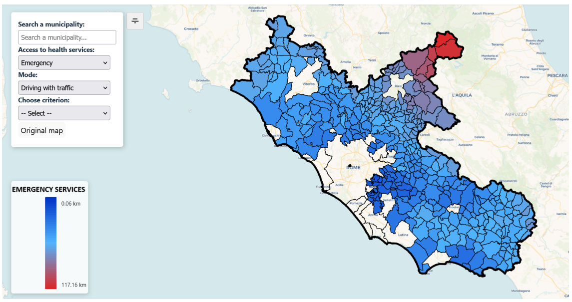


Figure 27: *Emergency distances (km) with traffic*

These areas face potential delays in time-critical healthcare interventions due to geographic isolation and infrastructure centralization.

Such spatial inequality in access to emergency care poses a serious public health concern, especially for residents in regions where the combination of distance, topography, and traffic can drastically extend emergency response times.

Like for educational services, the rankings for health service accessibility across Lazio's urban centres provide clear insights into how geographic and infrastructural disparities affect residents' access to care.

Table 33: Best and worst urban centres for health service accessibility (distance and time)

Service	Best UrbanCentre	Value	Worst UrbanCentre	Value
Family Doctors	Barbarano R. (5600410001)	0.034 km	Micigliano (5703724901)	39.681 km
	Prossedi (5902010002)	0.083 min	Villa S. S. (6009026710)	60.000 min
Hospitals	Sabaudia (5902410005)	0.276 km	Ponzano R. (5808010001)	79.549 km
	Sabaudia (5902410005)	0.900 min	Collegiove (5702010001)	74.750 min
Emergency	Ventotene (5903310001)	0.056 km	Accumoli (5700126608)	123.621 km
	Ventotene (5903310001)	0.183 min	Accumoli (5700126608)	118.189 min

Below is a synthesis of the patterns that emerge:

- Family doctors
  - Best Access (distance & time): Urban centres like Barbarano Romano, Civitella San Paolo, and Prossedi show exceptional accessibility, with distances as low as 34–43 meters and travel times under 10 seconds in some cases.
  - Worst Access: The municipality of Pescorocchiano dominates the bottom rankings with multiple urban centres exceeding 25–30 km and travel times over 40 minutes. Micigliano reaches nearly 40 km, the highest recorded.

  These results confirm that family doctors are generally well-distributed, but mountainous and sparsely populated areas still face major accessibility gaps.

- Hospitals
  - Best Access: Urban centres like Sabaudia, Amatrice, and Colleferro stand out for proximity to hospitals, with travel distances below 0.5 km and under 2 minutes.
  - Worst Access: The Tevere Valley area (e.g., Ponzano Romano, Torrita Tiberina, Filacciano, Nazzano) consistently ranks among the least accessible. Some urban centres face distances of 70–80 km and travel times over 70 minutes.

  The presence of multiple urban centres from the same municipalities in the worst list (especially Torrita Tiberina) indicates internal disparities within those municipalities.

- Emergency services
  - Best Access: The island municipalities of Ventotene and Ponza exhibit exceptional proximity, with some urban centres less than 0.1 km from emergency service locations. Even Monte Porzio Catone performs well.
  - Worst Access: Accumulate and Amatrice, already known for their isolation and vulnerability, clearly suffer from severe delays in access to emergency care, with distances over 120 km and times approaching 2 hours.

These are critical gaps, particularly for emergency services, where every minute is life-saving. The data reveals an urgent need for targeted solutions in these extreme outlier areas.

The islands of Ponza and Ventotene are special cases. Both islands achieve exceptional results in terms of accessibility to emergency services. This is due to the presence of two emergency outpatient facilities managed by the local health authority of Latina, which serve as local response points for urgent care. The small size and proximity of urban centres to these medical facilities result in some of the best scores in the region in terms of access to emergency services: 0.06 km and 11 minutes. It is also very important to note that the official registers of the official sources used do not list any hospitals on these islands, which is why no results are shown for accessibility to hospital services.

4.3.1 Population in healthcare accessibility

As already observed for educational services, it is important to assess whether a similar relationship exists between population size and accessibility to health services. This analysis helps determine whether sparsely populated areas are generally at a disadvantage in terms of reaching essential medical infrastructure, such as family doctors, hospitals, and emergency services.To explore this aspect, Pearson correlation coefficients were computed between the population of each urban centre and both travel distance and time to the nearest health service. The goal is to understand whether lower population density correlates with worse accessibility, reinforcing the hypothesis of territorial health inequality.

Table 34: Pearson correlation - population size and average accessibility to healthcare services

Service Type	Accessibility Measure	Pearson r	p-value
Family Doctors	Distance (km)	-0.189	2.86×10^{-19}
Family Doctors	Time (min)	-0.172	3.06×10^{-16}
Hospitals	Distance (km)	-0.105	7.71×10^{-6}
Hospitals	Time (min)	-0.114	1.21×10^{-6}
Emergency Services	Distance (km)	-0.151	6.70×10^{-13}
Emergency Services	Time (min)	-0.152	5.88×10^{-13}

All correlations are negative, suggesting that more populated urban centres generally have better access to healthcare services (i.e., shorter travel distances and times). The strongest correlation is observed for Family Doctors, especially in terms of distance ($r = -0.189$), indicating a moderate but statistically significant tendency for smaller towns to be farther from primary care. Emergency services and hospitals follow similar patterns but with weaker coefficients. Despite the low magnitude of the Pearson r values (between -0.105 and -0.189), the very low p-values indicate high statistical significance, confirming the trend observed across the health domain.

4.3.2 Traffic in healthcare accessibility

To understand how spatial barriers affect access to health services, the average travel distance and travel time for each urban centre in the region were compared across three categories of healthcare access: family doctors, hospitals, and emergency services.

This type of analysis helps assess whether there is a consistent relationship between physical distance and time required, or if certain services or territories are particularly affected by infrastructure inefficiencies or geographic constraints.

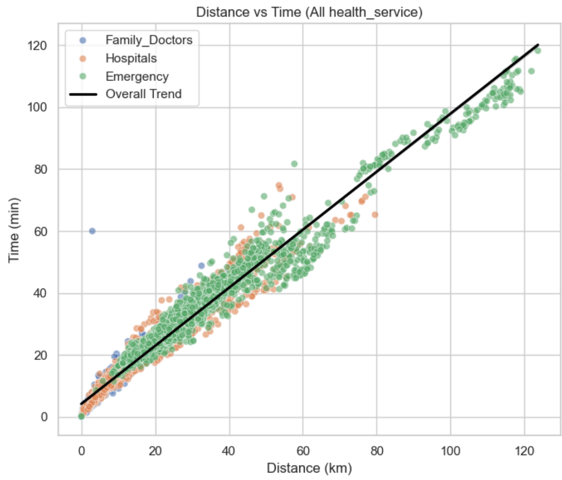


Figure 28: *Relationship between distance and travel time in healthcare*

The plot 28 shows a nearly linear relationship between distance (km) and travel time (min) across all types of health services. The Pearson correlation coefficient is high: $r=0.983$, $p<0.001$.

This result confirms that, on average, travel time increases proportionally with distance. The correlation is almost perfect, which is expected in a dataset where driving time is the dominant factor and where the road infrastructure plays a central role. Emergency services (green points) are present across the full spectrum of distances, reflecting their concentration in specific hubs. Family doctors (blue) tend to cluster at shorter distances and times. Some outliers do exist, especially for hospitals and emergency services, indicating delays not fully explained by distance, potentially due to road conditions, terrain, or poor connections.

The following table 35 shows the 10 urban centres with the highest absolute residuals (residuals from distance–time model), meaning that their observed travel times significantly diverge from the expected time based on their physical distance. These cases are critical to identify potential inefficiencies or anomalies in accessibility.

Table 35: Top urban centres with unexpected travel times to health services

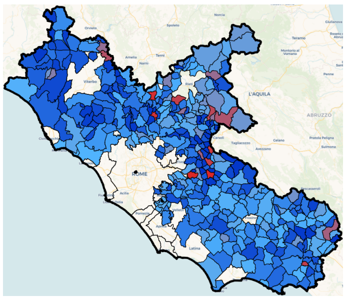
Health Service	**Urban Centre (ID)**	**Residual (min)**
Family Doctors	Villa S. S. (60090) – 3.00 km, 60.00 min	+52.98
Emergency (Filettino)	Filettino (24901) – 57.70 km, 81.69 min	+23.41
	Filettino (26602) – 49.50 km, 71.22 min	+20.62
	Filettino (10001) – 46.19 km, 66.87 min	+19.37
Hospitals	Collegiove (57020) – 53.52 km, 74.75 min	+20.39
	Collalto S. (57018) – 53.84 km, 73.67 min	+19.00
	Pescorocchiano (57049) – 48.76 km, 65.03 min	+15.13
Emergency (others)	Subiaco (58103) – 41.28 km, 57.40 min	+14.50
	Petrella S. (57050) – 60.57 km, 46.32 min	**–14.65**

These urban centres with a residual value > 0, exhibit travel times that are much longer than what would be expected for their distance, possibly due to poor road connectivity or mountainous terrain, indirect routes or limited infrastructure and traffic or delays not captured by the distance alone. Most notably, Villa Santo Stefano shows an extreme mismatch: only 3 km to the nearest family doctor, but an average travel time of 60 minutes, with a residual of nearly +53 minutes, clearly indicating a critical accessibility issue, possibly due to classification errors or severe local transport conditions. In contrast Petrella Salto shows a travel time that is significantly faster than expected given the long distance. This could be due to high-speed road access (e.g., a nearby highway) or low traffic congestion.

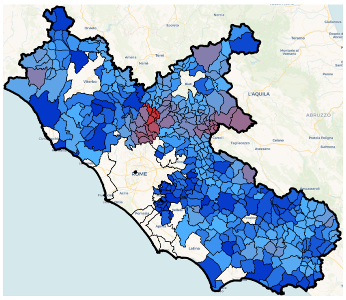
Again, outliers like these are crucial in planning, because they do not follow the general accessibility trend and may indicate structural barriers (infrastructure, geography), service distribution gaps, or data inconsistencies worth investigating further. They also highlight that distance alone is not sufficient to describe accessibility, and that localized conditions must be considered in policy and infrastructure decisions.

4.3.3 Accessibility healthcare results aggregated for municipality

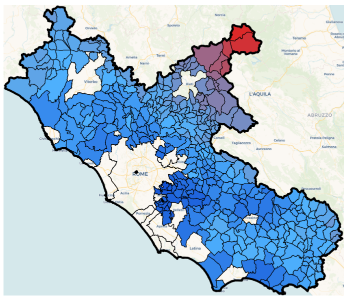
As for educational purposes, the results for municipalities in terms of accessibility to various health services are shown (figure 29). Focusing on hospital and emergency services, anomalies can be noted that suggest inconsistencies between the results. This is not what it seems, however, as it should be remembered that the methodology for obtaining emergency data is slightly different from that for hospitals (which takes into account the municipality's health area when allocating infrastructure). This indicates that some of the areas shown in red may refer to municipalities that are located in health areas that are centralised in another part of the territory.



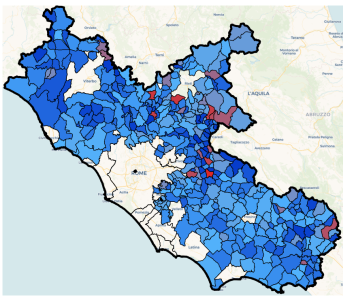
(a) *Family doctors – Distance (km)*



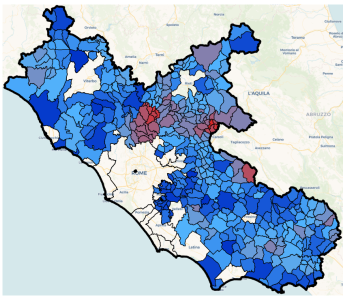
(b) *Hospitals – Distance (km)*



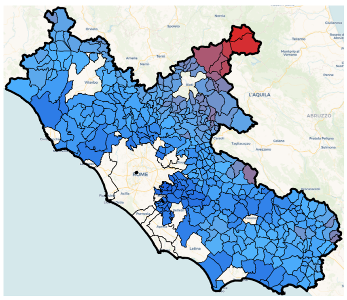
(c) *Emergency services – Distance (km)*



(d) *Family doctors – Time (min)*



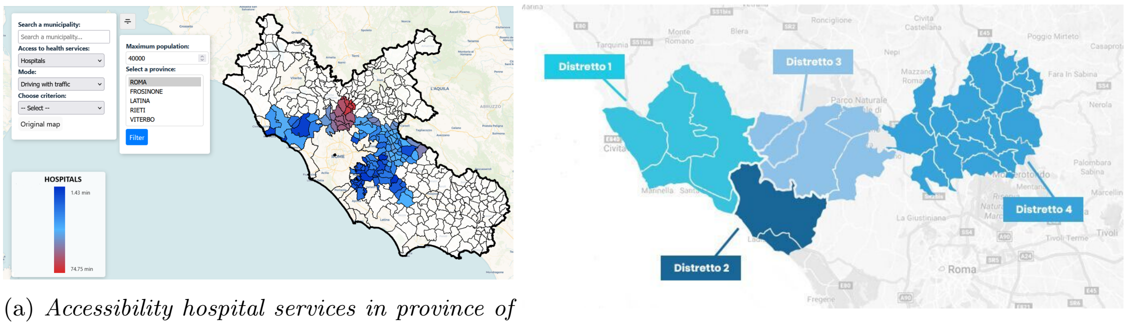
(e) *Hospitals – Time (min)*



(f) *Emergency services – Time (min)*

Figure 29: *Distribution of average distances for each healthcare services at the urban centre level*

Thanks to the viewer's features 30a, an in-depth data validation analysis was carried out on the area where outliers values were found in terms of hospital accessibility. The municipalities in red that can be seen on maps 29b and 29e, they are part of the ASL ROMA 4 health zone, the structure of which can be seen in the figure opposite. Even after careful research, this confirms that the hospitals in ASL Roma 4 are concentrated in districts 1 and 2 (see figure 30b), thus showing poor results for hospital accessibility in those municipalities. As there are no hospitals, this area contains several smaller facilities and, as shown by the emergency data, hospitals belonging to other health zones, but the problem remains for citizens who need to go to hospital for non-urgent scheduled appointments.



(a) *Accessibility hospital services in province of Rome*

(b) *ASL ROMA 4 - health zones*

Figure 30: *Accessibility of hospitals in Rome - ASL ROMA 4*

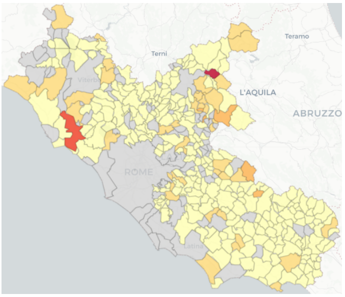
In general the aggregated results by municipality confirm the same spatial trends already observed at the urban centre level:

- Ventotene consistently ranks first across all health service types and both metrics (distance and time), just as in the detailed urban centre level analysis. This validates the presence of its local outpatient emergency facility.
- Worst cases like Nazzano, Torrita Tiberina, Filacciano, Civitella San Paolo, and especially Accumoli and Amatrice also reappear across both analyses, confirming severe accessibility challenges.
- The same structural inequalities observed in the urban centres, based rankings persist after aggregation, indicating that municipality-level averages still reflect underlying accessibility gaps, especially in mountainous and peripheral areas.
- Minor variations (e.g., Poli being among the worst for family doctors but best for hospitals) can be explained by within-municipality disparities and service-specific infrastructure.

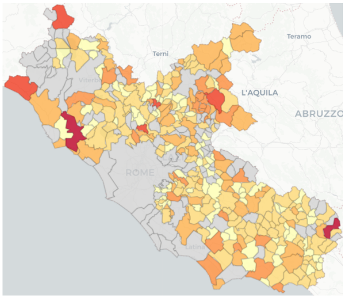
In conclusion the municipality-level analysis reinforces the findings at the finer urban-centre scale, demonstrating robustness of results and validating the use of aggregated indicators for strategic planning when detailed spatial granularity is not available.

Now the disparities in access within municipalities are shown (figure 31). The image highlights a difference between hospital and emergency services and those provided by family doctors, but do not be misled by the colours, because the shade of red corresponds to the maximum for each category and not overall.

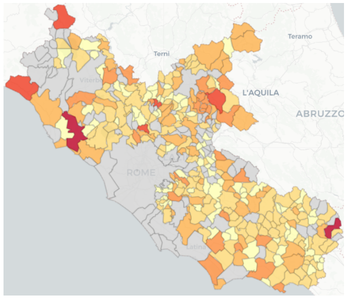
The results show that, compared to education, the levels of internal inequality between households within the same municipality are generally less marked in health services, especially for hospitals and emergency services. This is consistent with the metropolitan or supra-municipal nature of these services: they are not distributed as widely as nursery or primary schools, but are centralised in reference facilities that serve larger areas.



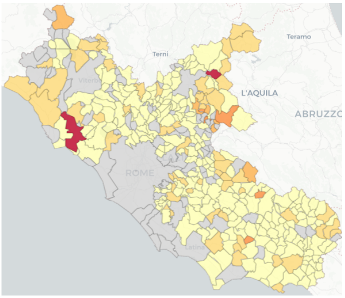
(a) *Kindergarten – Distance (km)*



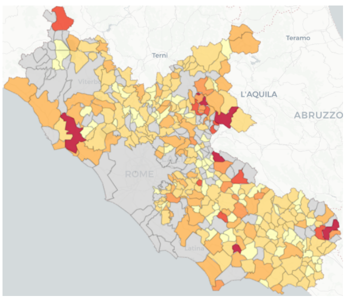
(b) *Primary – Distance (km)*



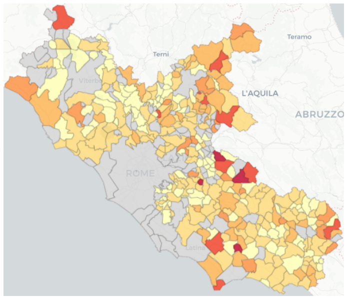
(c) *Secondary – Distance (km)*



(d) *Kindergarten – Time (min)*



(e) *Primary – Time (min)*



(f) *Secondary – Time (min)*

Figure 31: *Intra-Municipal disparities based on standard deviation - healthcare services*

In the few municipalities where high levels of inequality are found (e.g. Tolfa, Vallerotonda, Pescorocchiano, Acquapendente), these are large or geographically fragmented areas, often mountainous, where the physical distance between centres is significant and road links can vary greatly. However, the overall scale of values is more limited than in education, where inequalities of 10–15 km between centres were much more frequent. In summary, health services reflect a more centralised model of access, and as a result, intra-municipal inequalities are less pronounced in most cases.

4.3.4 Accessibility healthcare aggregated for province

After analysing accessibility to healthcare services at the individual municipality and urban centre level, it is useful to scale up and evaluate the aggregate results by province, always considering the population when calculating the weighted average, in the same way as when the data was aggregated by municipality. This summary allows for the rapid identification of structural territorial issues and provides an objective starting point for guiding rebalancing policies and infrastructure investments. The weighted average values represent the general situation of the resident population, while the standard deviations between municipalities indicate internal territorial disparities within the province. Together, these indicators help to identify the provinces that require greater attention, both in terms of distance and average access time and in terms of inequalities between internal territories.

Table 36: Provincial indicators of access to family doctors

Province	**Mean Distance (km)**	**Mean Time (min)**	**Std. Dev. Dist.**	**Std. Dev. Time**
VT (56)	1.36	2.79	2.38	2.92
LT (59)	1.88	3.84	1.19	1.61
RM (58)	2.06	4.20	3.61	4.63
FR (60)	2.17	4.46	2.13	3.02
RI (57)	3.67	5.79	3.68	5.35

Table 37: Provincial indicators of accessibility to hospitals

Province	Mean Distance (km)	Mean Time (min)	Std. Dev. Dist.	Std. Dev. Time
LT (59)	8.61	12.35	7.59	9.88
FR (60)	11.31	15.67	8.05	9.72
VT (56)	12.76	14.80	8.58	7.97
RM (58)	12.89	16.39	17.78	16.42
RI (57)	29.40	31.18	12.34	14.10

Table 38: Provincial indicators of accessibility to emergency services

Province	Mean Distance (km)	Mean Time (min)	Std. Dev. Dist.	Std. Dev. Time
RM (58)	18.06	23.09	9.38	8.60
LT (59)	23.18	28.31	8.70	9.67
FR (60)	26.31	30.80	6.57	7.77
VT (56)	31.96	34.81	7.81	6.45
RI (57)	49.58	49.34	16.92	13.80

The province of Rieti (RI) emerges as the most disadvantaged in terms of average distance and time, thus demonstrating that the trend is repeated for both educational and health services. This may be due to the lack of homogeneity between municipalities, and also confirms the hypothesis of a province distant from the capital.

On the contrary, Viterbo (VT) and Latina (LT) show more favourable and homogeneous conditions in all services. Surprising, however, is the strong heterogeneity of Rome (RM), which, despite having a good average, has the highest internal inequality, a sign of strong imbalances between metropolitan and peripheral areas, which also confirms our deduction that health areas are poorly organised.

Emergency services show a very uneven distribution, with the province of Rieti (RI) once again lagging far behind. Rome (RM) shows better average access, due to the large number of facilities in the area, but this is accompanied by significant territorial disparities.

4.4 Public transport accessibility results

Up until this point, the results presented focused on accessibility by private car, taking traffic conditions into account. This section shifts the focus to accessibility indicators related to public transportation.

Before delving into the results, it is important to clarify that the following analyses are limited to the Province of Rome. This decision was made due to a significant number of null values returned when querying the Google Maps API for travel distances via public transport. Several factors may have contributed to this issue:

- Lack of data from operators: many transport companies in the region are private, and their services are not always integrated into Google Maps;
- API request configurations: despite careful attention to departure time parameters, some routes may not have been served at the time or day of the request;
- Data coverage limitations: Google Maps typically uses public GTFS datasets, which are the same sources used when attempting routing via OTP (OpenTripPlanner). As a result, missing data affected both methods;
- Geographic dispersion: although centroids were adjusted to fall within residential areas, some remain in sparsely populated or rural locations where public transport is unavailable.

Given these limitations, the analysis was focused on the province of Rome, where data availability is more consistent. Some missing values were recovered using the MOOVIT API.

For the purpose of this report, detailed public transport accessibility results by service are not presented, as was done for private car. These can be explored through the interactive online map, where users can filter results based on service type, mode of transport, and accessibility metric. An example of this interface is shown in the figure 32. This section analyses which urban centres are not served by public transportation, providing key details on the nature and extent of this exclusion.

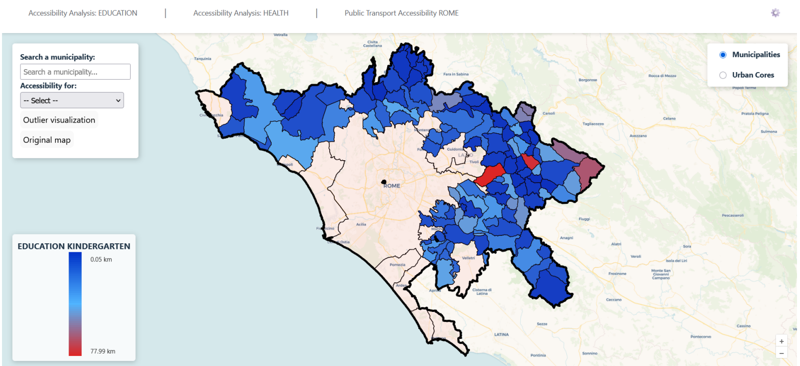


Figure 32: *Kindergarten accessibility indicator with public transport*

The image 32 shows that in the viewer, the section for obtaining public transport data is slightly different from the others. The user will view the results by selecting the parameters, starting with choosing whether to display education or health and then selecting the parameters as for the other sections.

4.4.1 imputation of missing data distances

The focus now shifts to the distances for which no public transport data were available and which, as explained in the methodology section, were imputed to avoid aggregation issues when calculating accessibility indicators at the municipal level.

This part presents the education and health results side by side, with a specific focus on hospital services within the health sector. Emergency services were not considered in this analysis, as public transport accessibility to emergency care is generally not meaningful given the urgent nature of such services.

No null values were encountered in the public transport results for family doctors. This is likely due to the local nature of general medical practices in Italy, where the distance to the assigned doctor is typically short. In fact, the Google Maps API often returns a valid result (even under the public transport mode) by interpreting short trips as walkable, and thus provides a result rather than a null value.

This behaviour helped ensure complete data coverage for family doctors, while hospitals and schools, due to their more centralized and sparser distribution, exhibited significantly more missing values, which required imputation for a complete and reliable territorial analysis.

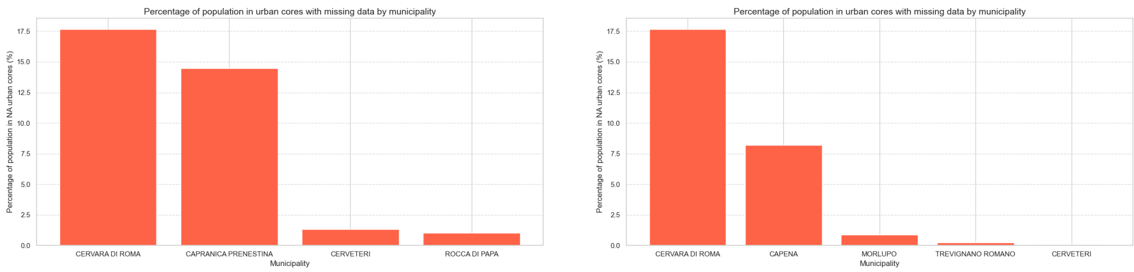
The tables 39 present the urban centres in the province of Rome for which no public transport access to key services could be calculated.A distinction is made between educational services (Kindergarten, Primary, Secondary) and health services, which in this context are limited to hospitals only.

Table 39: Urban centres without public transport access

Service	UrbanCentre ID	Population
Education	5802910000	299
	5808610000	183
	5802920000	179
	5802820001	82
	5801910000	46
	5802926612	30
Hospitals	5801820001	888
	5802820001	82
	5806820003	75
	5802926612	30
	5810726610	13

These data are important as they highlight clusters of population potentially excluded from essential services when using public transportation.

To better understand the territorial impact of missing public transport data, the analysis examined the proportion of each municipality's population residing in urban cores without accessible public transport routes to hospitals or educational services.



(a) *Education services – NA population by municipality*

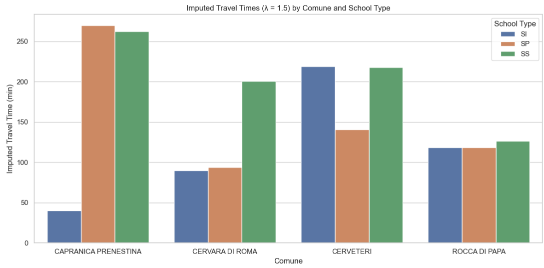
(b) *Hospital services – NA population by municipality*

Figure 33: *Population in urban cores without public transport access to services*

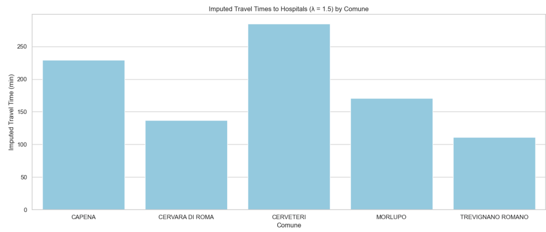
As shown in the two charts, Cervara di Roma stands out in both cases, with over 17% of its population located in urban cores for which no public transport route was retrievable. Capena and Capranica Prenestina also exhibit significant shares of population excluded from the network. In contrast, larger municipalities like Cerveteri and Rocca di Papa show missing values limited to marginal fractions of their population, indicating more robust network coverage in denser areas. These cases suggest that partial exclusion from services via public transport tends to affect small or fragmented municipalities, especially those with hilly or peripheral settlements.

Based on the imputation method described in the methodology section, the final aggregated accessibility values at the municipal level, after including the imputed values for urban cores where public transport data was missing, are presented. The imputation approach incorporates a penalization factor (λ) to reflect the uncertainty and potential inaccessibility of these zones, while still allowing them to contribute proportionally to the municipality's overall accessibility indicators.

Figure 34 displays the corrected accessibility values by municipality (after imputation). As expected, the impact of imputed data on the final municipal values is relatively small in most cases, especially where the excluded population was marginal. However, the penalization ensures that municipalities with a significant share of population in unreachable areas are appropriately penalized in the final score.



(a) *Education services – imputed and aggregated values by municipality*



(b) *Hospital services – imputed and aggregated values by municipality*

Figure 34: *Imputed value - result by municipality*

This approach guarantees data completeness for aggregation and comparison, a fair and conservative estimate that discourages underestimating accessibility gaps and continuity of analysis between complete and partially incomplete municipalities.

4.4.2 Comparing accessibility by Public transport vs. private car

To better understand the effectiveness of public transportation in delivering equitable access to essential services, the difference in travel times between public transport and private car is now compared. Given the similar trend in the results obtained, for simplicity's sake, only the results obtained from the analysis for educational services for the three levels: Kindergarten, Primary, and Secondary, will be shown.

This analysis highlights the accessibility gap experienced by users of public transport and quantifies how much longer, on average, it takes to reach educational services without a private vehicle. The comparison is based on the difference in average travel time (in minutes) between the two modes, aggregated at the municipality level within the Province of Rome.

Table 40: Difference in travel time between public transport and private car – education level

School Level	Count	Mean	Std Dev	Min	25%	Median	75%	Max
Kindergarten	107	24.16	36.75	0.32	5.37	10.79	26.11	194.28
Primary	108	20.59	30.73	–2.68	4.80	10.21	26.31	194.28
Secondary	108	25.95	34.11	–0.50	6.06	13.27	32.13	194.28

The results clearly show that travel times are significantly longer when using public transportation compared to private cars, across all education levels:

- On average, public transport users face 21–26 minutes longer trips to reach educational services.
- The spread of values is large, with standard deviations over 30 minutes, and maximum differences approaching 3 hours.
- Median values remain below 15 minutes, indicating that half the municipalities experience a moderate gap, but the upper quartile shows major penalties for users without a car.
- Negative values (minimums) are rare and likely represent edge cases where transit lines are more direct or where traffic severely delays driving routes.

This disparity is especially relevant for kindergarten and secondary schools, where the average time gap exceeds 24 minutes.

These findings emphasize the need for enhanced public transport planning, particularly in outer or semi-rural areas, to reduce reliance on private vehicles and support equitable access to education.

While the average difference between public transport and private car accessibility is informative, it is equally important to understand whether these differences are consistent across services or if there are anomalies that indicate strong intra-municipal variability.

To this end, the standard deviation of the difference in travel time across the three education levels (Kindergarten, Primary, Secondary) for each municipality were computed. This indicator helps identify where public transport performance is erratic, possibly due to fragmented infrastructure, limited coverage, or service irregularities.

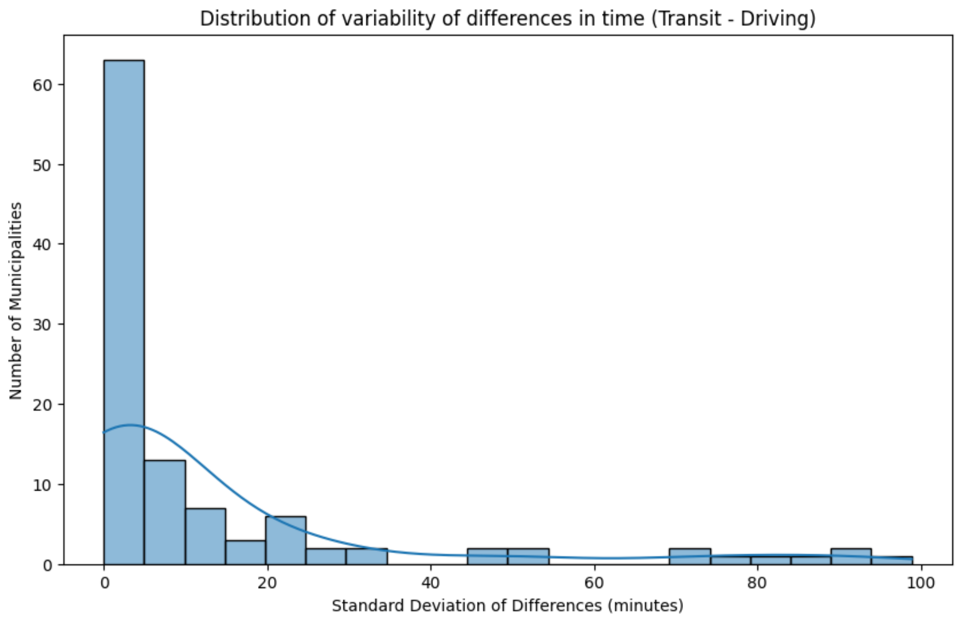


Figure 35: *Standard deviation of travel time difference (public transport vs car) - school levels*

The results (fig. 35) highlighted that, municipalities of Rocca Canterano, San Gregorio da Sassola, and Jenne show extremely high variability (over 89 minutes of standard deviation), suggesting that while access to one school type might be manageable, other services are poorly connected or entirely unreachable by public transport. This inconsistency may result from, limited bus/train routes only serving certain destinations, non-overlapping transit schedules, or disconnected local networks for different service types. Such findings highlight the need for a more integrated and balanced transport planning, ensuring that all education levels are accessible within reasonable time frames for residents without access to private vehicles.

To complement the previous analysis of differences in accessibility between private car and public transport, s correlation matrix was computed across key variables related to travel distance and time. This analysis helps assess to what extent travel time differences can be explained by physical distance or are more strongly tied to the structural inefficiencies of public transport.

The results from the correlation heat map shows that the strongest correlation is between public transport time and the accessibility gap (r = 0.99). This confirms that long public transport durations are the main drivers of accessibility inequality. In contrast, driving times and driving distances are weakly correlated with the gap (r = 0.38), suggesting that poor car accessibility is rarely the issue. Public transport distances show a moderate correlation with the gap (r = 0.71), which may reflect circuitous or poorly connected routes in some areas.

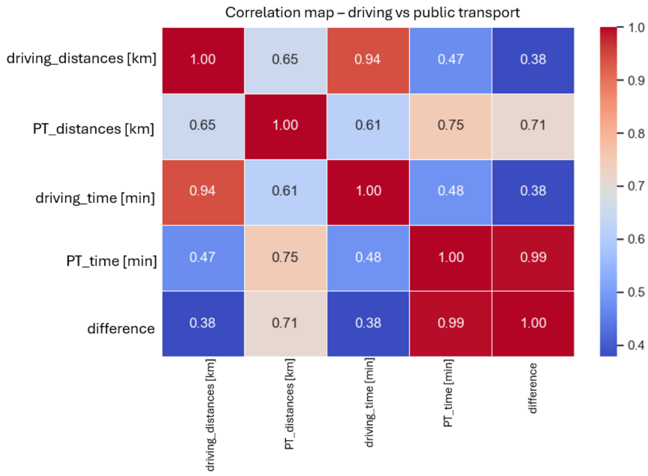


Figure 36: *Correlation Heat Map - driving vs public transport*

The correlation between driving distances and public transport distances (r = 0.65) is modest, indicating that the two modes often follow different spatial paths, especially in rural or disconnected zones.

This matrix reinforces the earlier conclusion: the accessibility gap is primarily driven by inefficiencies and coverage issues in the public transport network, not by physical remoteness or poor road infrastructure. Therefore, policy interventions should focus on improving transit reach, frequency, and integration, especially in areas where driving is currently the only viable option.

4.5 Comparative analysis: Lazio region vs. Community of Madrid

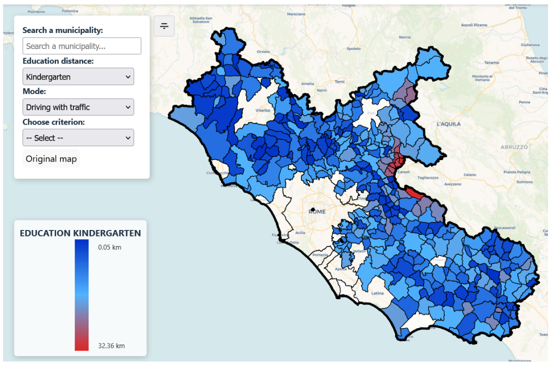
This final section of the results provides a descriptive comparison between two distinct territorial contexts: the Lazio Region in Italy and the Community of Madrid in Spain. Rather than relying solely on aggregated statistics, the comparison is visually supported through screenshots from each region's interactive map viewer. The objective is to highlight the most relevant spatial patterns and accessibility differences between the two areas, with a focus on public service coverage, transport network structure, and urban–rural disparities.

This qualitative exploration offers a basis for cross-regional insights and helps contextualize the results of Lazio within a broader European framework.

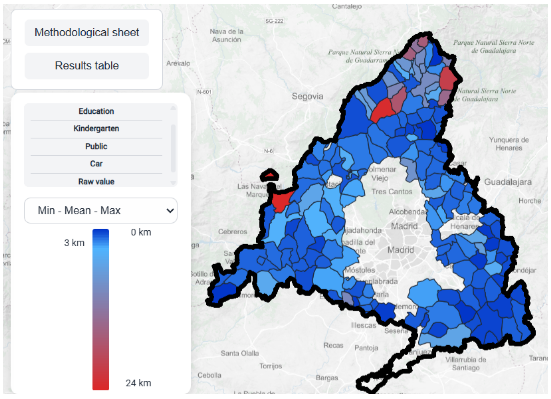
Both Lazio and the Community of Madrid show very similar accessibility patterns for educational services. As visible in the maps:

- The core metropolitan areas (Rome and Madrid) consistently have better accessibility across all school levels, with minimal driving distances.
- Accessibility tends to deteriorate as one moves towards peripheral or mountainous zones, especially when the distance from the capital increase.
- Importantly, secondary schools show wider red zones and greater spatial disparity, confirming a key finding already observed in Lazio: the higher the school level, the lower the territorial coverage.

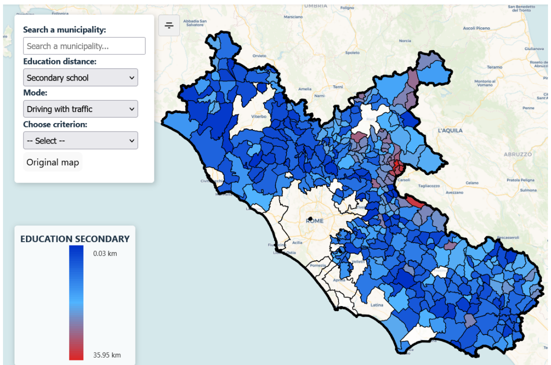
This cross-regional consistency strengthens the conclusion that accessibility to higher-level education services is structurally weaker, and thus should be prioritized in transport and urban policy.



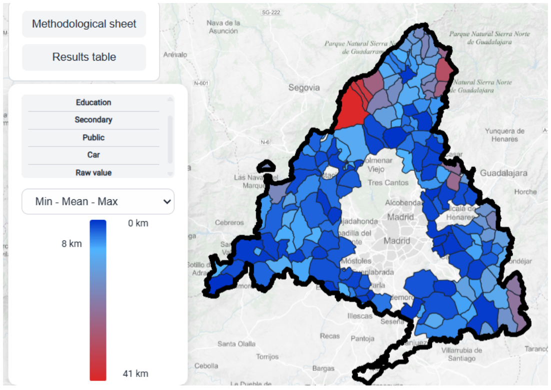
(a) *Lazio – Kindergarten (Driving)*



(b) *Madrid – Kindergarten (Driving)*



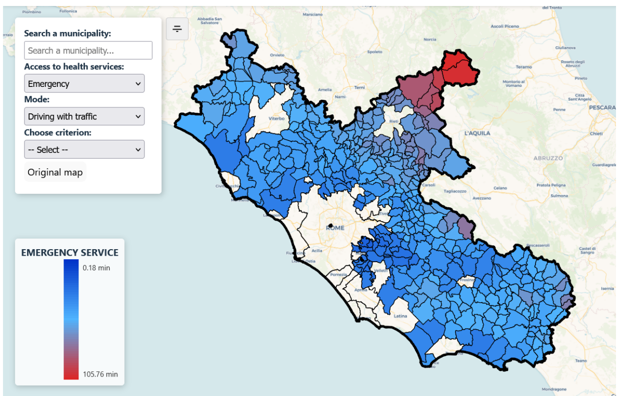
(c) *Lazio – Secondary School (Driving)*



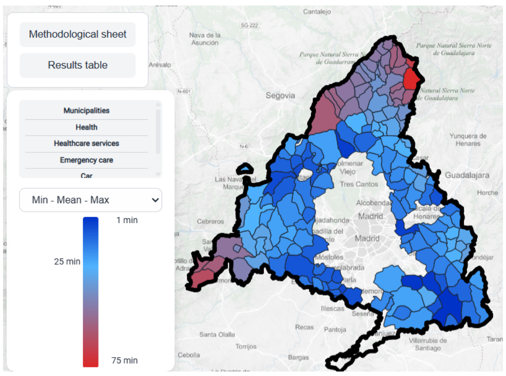
(d) *Madrid – Secondary School (Driving)*

Figure 37: *Comparison of accessibility to education between Lazio and Madrid by school level*

Some results about healthcare services are shown. The spatial distribution of emergency service accessibility shows a clear geographical pattern in both Lazio and the Community of Madrid: the farther you move from the capital city, the longer it takes to access emergency services.



(a) *Lazio – Emergency Service Accessibility (Time)*



(b) *Madrid – Emergency Service Accessibility (Time)*

Figure 38: *Accessibility to emergency healthcare by car in Lazio and Community of Madrid*

In both regions, areas near Rome and Madrid show travel times close to 0.5–1 minute, highlighting dense emergency service coverage in the capitals. As you approach border municipalities, the situation changes drastically. The maximum driving time to reach an emergency service in: Lazio reaches 105–118 minutes, with critical areas along the north-east and eastern Apennines boundary. While Madrid peaks at about 75–79 minutes, primarily in the northern and southwestern rural edges. Despite differences in territorial extent and topography, the trend is remarkably similar, underlining the common challenge of ensuring equitable emergency access across mountainous or peripheral areas. This visualization reinforces the need for decentralized emergency structures in border zones and faster intervention strategies such as telemedicine or air transport in hard-to-reach areas.

In the final part of the results, a comparison is made between two similar public health services: general practice in Italy and paediatrics in the Community of Madrid. Although they serve similar purposes, providing basic and preventive health services, their organizational structures differ notably. In Italy, general practitioners and paediatricians operate from distributed local medical offices, often close to residential areas. In contrast, in Spain (specifically the Madrid region), these services are centralized within healthcare centres, which are fewer and more spatially concentrated.

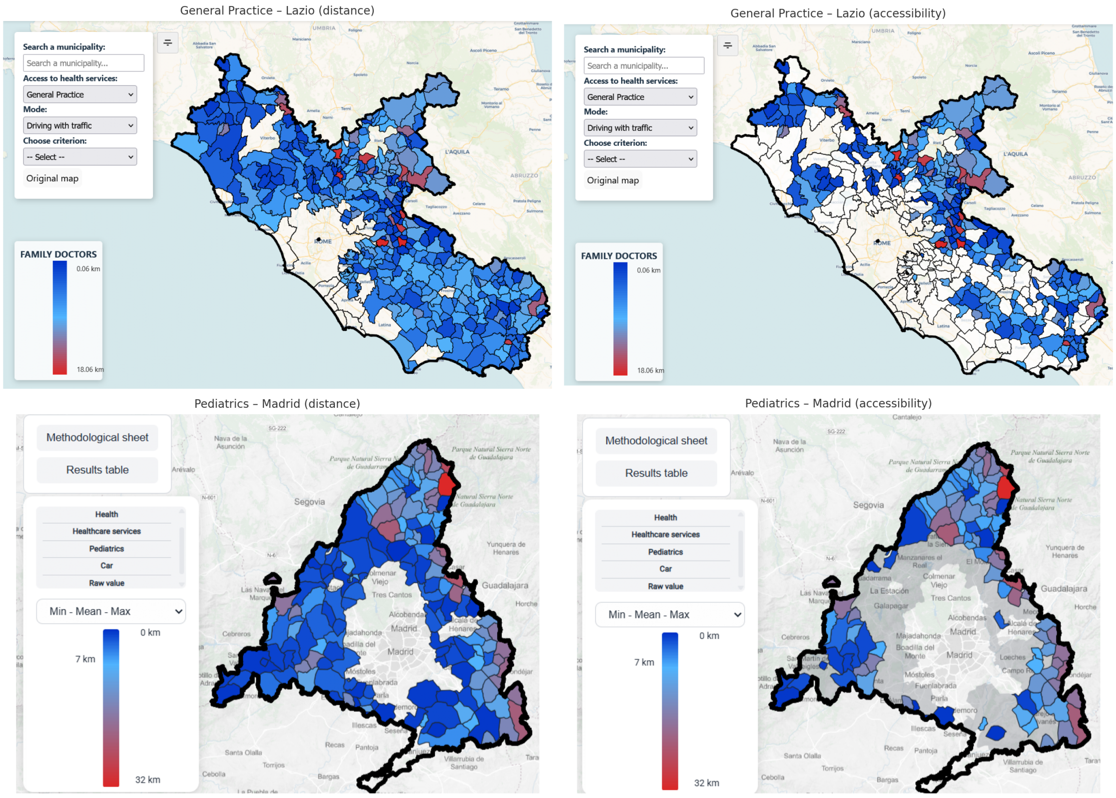


Figure 39: *Comparison between general medicine accessibility in Lazio and in Madrid*

As visualized in the maps 39, this difference in service distribution directly affects accessibility. The Lazio region demonstrates overall better accessibility in terms of distance to general practice services compared to paediatric care in Madrid, where long distances (highlighted in red) are more frequent.

Furthermore, in both regions, poor accessibility is generally associated with sparsely populated municipalities, often situated in peripheral areas. This reinforces what was already shown earlier: population size plays a crucial role in shaping accessibility outcomes, confirming a widely discussed hypothesis in the literature.

Final consideration

During this chapter, a comprehensive and detailed analysis was conducted on the accessibility to educational and healthcare services across the Lazio region. This section summarizes the most relevant findings by addressing the key analytical questions that guided the study.

Significant intra-municipal disparities in accessibility were observed, particularly in municipalities comprising multiple urban centres. The most critical disparities were found in mountainous or peripheral territories where topography and low service coverage amplify inequalities between urban centres. In this last phase, it is best to focus on the characteristics of the municipalities or urban centres under consideration, rather than on which they are, like presented in the detailed analysis carried out in this chapter.

Which urban centres exhibit the best accessibility for each analysed service, and is this also reflected at the municipality level?
Urban centres with the best accessibility are typically located in dense, centrally located municipalities, often the main town, where services are physically concentrated. These conditions generally reflect positively at the municipality level, confirming that internal centrality often translates into better aggregate performance.

Which urban centres exhibit the worst accessibility for each analysed service, and is this condition also observable at the municipality level?
Urban centres with the worst accessibility tend to be isolated, rural, or sparsely populated, particularly in internal areas of Rieti and southern Lazio. These weaknesses often impact the overall municipal average, although in some cases they may be masked when aggregated.

Which municipalities show the greatest intra-municipal disparities in accessibility, both in terms of time and distance, between their urban centres?
Municipalities exhibiting the greatest internal inequalities are generally composed of highly fragmented territories, where certain urban centres benefit from proximity to major roads or service nodes, while others remain disconnected. These disparities were measured through the standard deviation of accessibility indicators and highlight municipalities in need of sub-municipal interventions.

How do the provinces rank, in terms of accessibility, for each analysed service?
At the provincial level, Rome shows the best average accessibility across all services, due to its dense service network. Viterbo, Latina, and Frosinone show moderate accessibility, while Rieti consistently ranks last. The standard deviation among municipalities within provinces reveals that Rieti and Rome also present the greatest inter-municipal disparities, pointing to internal structural inequalities.

What are the differences in accessibility, both in terms of distance and time, in the province of Rome?
In the province of Rome, central urban centres exhibit very good accessibility, whereas peripheral or mountainous zones face significantly longer distances and travel times. This contrast highlights the importance of intra-provincial geography and service distribution.

Are the differences between car access (considering traffic) and public transport access to a given service consistent across the region, or are there significant outliers?
While central areas of the region show relatively consistent differences between private and public transport access, significant outliers exist in peripheral areas where public transport infrastructure is weaker. These gaps lead to strong inequalities in temporal access and reinforce mobility disadvantage in rural contexts.

Is there a correlation in accessibility between municipalities with similar population densities?
A weak but statistically significant negative correlation was observed between population and accessibility. More populated areas tend to enjoy better access to services. This finding is aligned with accessibility theory but highlights that population is not the only driver, topography, infrastructure, and spatial planning also contribute significantly.

What is the relationship between the most and least accessible services in the two case studies (Community of Madrid and Lazio)?
Direct comparison with the Community of Madrid is essential to contextualize the results from Lazio. The main insight, Madrid's heath-care centralized model, leads to more uniform accessibility levels, while Lazio's dispersed settlement structure results in higher variability. Others accessibility has the same trend of Italy depending of which public services is considered, the traffic affect less Madrid than the Lazio region. This comparison allows a better understanding of what constitutes "good" or "poor" access and demonstrates how spatial configuration shapes outcomes.

The results shown throughout this chapter were intended to provide a broad and comparative view of accessibility conditions, using descriptive indicators to confirm or challenge common hypotheses found in academic literature.

While these aggregate results offer valuable insights at the regional level, more detailed and case-specific analyses can be conducted using the interactive web dashboard or the data notebooks made available with this study. These tools allow for in-depth exploration of each municipality and urban core, offering stakeholders a flexible framework to guide policy planning, infrastructure design, or service redistribution strategies.

5. Conclusions

The results obtained throughout this study indicate that the majority of the initial objectives have been successfully achieved. The primary aim was to design and implement a data infrastructure containing reliable information that could serve as a foundation for future analyses. As previously defined, open data represents a fundamental resource, enabling universal access to valuable datasets. This work has established a solid foundation of accessibility indicators for essential public services, namely education and healthcare, with the hope that these indicators may support future improvements and inform planning decisions that ultimately benefit all residents of the Lazio region.

Nevertheless, achieving these results was not without challenges. One notable issue that remains unresolved concerns the unavailability of comprehensive public transport data for the entire Lazio region. During the course of the project, it was not possible to retrieve complete GTFS datasets from all local transport providers, revealing a persistent gap in data publication practices in some areas of Italy. This highlights the fact that several regions, particularly those managed by smaller or fragmented transport agencies, have yet to align with current open data standards. As a result, the public transport accessibility analysis was limited to the province of Rome, and only part of the original objective in this domain was fulfilled. Despite these limitations, the work lays the groundwork for future developments, offering a replicable framework that can be extended and refined as more data becomes available.

Looking ahead, several opportunities for improvement and expansion have been identified, which could enhance the accuracy, scope, and relevance of the platform in future iterations. While the current work provides a solid foundation for the analysis of accessibility to public services, several areas have been identified for potential improvement and expansion. These enhancements aim to increase the accuracy, completeness, and policy relevance of the platform, ensuring its adaptability to future planning needs and datasets.

1. **Temporal variability in distance calculations** The current distance and travel time metrics were obtained using the Google Maps API, assuming a fixed departure time. Although this approach provides an initial approximation of accessibility, it does not capture the temporal variability caused by changing traffic conditions throughout the day. Accessibility is inherently dynamic, and peak traffic hours can significantly affect travel times. A more robust methodology would involve calculating distances and durations across different time windows (e.g., morning rush hour, midday, evening), then aggregating the results, typically by averaging, to generate more realistic indicators. This enhancement can be integrated into future versions of the platform, leveraging the fact that all raw and aggregated data have been preserved and can be extended without the need to recalculate existing values.

2. **Normalisation based on demographic and spatial factors** To enhance the interpretability and equity of accessibility indicators, future work should explore more refined normalisation strategies. Until now, results have been primarily normalised by total population. However, additional layers of analysis could consider population density, proximity to the regional capital, and age, specific population groups. For instance, educational service indicators could be normalised by the number of residents aged 6–14, offering a more targeted assessment of how well the educational infrastructure serves its intended population. Similarly, incorporating spatial normalisation based on distance from central areas could help identify whether peripheral or rural zones are systematically underserved. These approaches would improve the contextual relevance of the indicators and allow for more nuanced and actionable interpretations.

3. **Accessibility also as a function of demand and supply** Accessibility cannot be fully captured by spatial distance or travel time alone. A more comprehensive approach should consider the balance between the demand for services, i.e., the number of people needing access, and the supply, i.e., the actual capacity of the services available. For example, a school that is geographically close but overcapacity may offer lower effective accessibility

than a more distant one with available space. Integrating this supply-demand perspective introduces a form of impedance that reflects congestion or scarcity, offering a more realistic and equitable measure of access. Future development could involve linking demographic data and facility capacity metrics to refine the accessibility indicators accordingly.

4. **Development of a composite accessibility index** Currently, separate indicators have been calculated for car-based and public transport-based accessibility. While this allows for independent assessments, a more informative metric would be a composite accessibility index that synthesizes both modes into a single score. This index would provide an overarching view of how accessible a given urban centres or municipality is, regardless of the transport mode. However, building such an index requires integrating data on both demand and supply: the number of individuals who use each transport mode and the level of service available. This information is essential to assign appropriate weights and ensure that the index reflects actual mobility patterns and not just theoretical service availability.

5. **Addressing the dispersion of urban centres** A technical but crucial aspect that requires further analysis is the spatial dispersion of urban centres representing population zones. In some cases, urban centres fall outside urbanized areas, potentially misrepresenting real travel dynamics and skewing accessibility metrics. Future work should explore systematic methods to reposition or redistribute centroids of these urban centres, based on population density, built-up areas, or connectivity to road and transport networks. This would enhance the spatial accuracy of the indicators and improve their interpretability for policy applications.

The platform developed through this project offers a versatile and valuable tool for multiple stakeholders, including urban planners, public administrators, and researchers. Its usefulness becomes especially evident when integrated with demand side data information about how citizens actually use or need access to public services. Although collecting such data is complex, incorporating it would allow for more informed and targeted territorial planning. By simulating the introduction of new infrastructures, such as schools, hospitals, or public transport lines, it becomes possible to model how accessibility would change across different areas, supporting proactive and evidence, based decision-making.

Moreover, the platform holds significant potential if adopted and maintained by regional authorities. When enriched with additional indicators, such as air quality, cost of living, or housing availability, it could evolve into a comprehensive urban attractiveness index. This would not only support policy development but also serve as a valuable resource for citizens, helping them make informed decisions about where to live based on service accessibility and overall quality of life.

Ultimately, the true value of this project lies not only in the technical results achieved, but in the possibility of continuation and expansion. It is essential that the work carried out so far be seen as a foundation upon which further developments can be built. Establishing collaborations with local institutions, municipal governments, and public service providers is crucial to ensure that the platform remains relevant, up-to-date, and responsive to real-world needs. By doing so, it can contribute meaningfully to more equitable, sustainable, and data-driven territorial planning, benefiting both public authorities and the citizens they serve.

In conclusion, this project demonstrates the potential of data-driven approaches to improve the understanding and management of territorial accessibility. By offering a structured, open, and extensible framework, it provides a concrete basis for future analyses, institutional collaborations, and citizen-oriented tools. While the current outcomes already represent a meaningful contribution, their real value will be fully realized only if the work is continued, enriched, and adapted over time. With the right investments and partnerships, this platform can evolve into a powerful instrument for promoting more inclusive and informed urban and regional development, reminding us that even data, when well used, can shape better places to live.

References

UNE 178601: Smart Territories. Definition, Attributes and Requirements, 2016. URL https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0069862. Accessed: 2025-05-26.

Gustavo Arciniegas, Rocco Janssen, and Niek Omtzigt. Testing the usability of a spatial decision support system for spatial planning. *Computers, Environment and Urban Systems*, 41:93–103, 2013.

ASL LAZIO. Ricerca medici e pediatri. https://www.salutelazio.it/ricerca-medici, 2025a. Accessed: 2025-03-12.

ASL LAZIO. I pronto soccorso nel lazio. https://www.salutelazio.it/pronto-soccorso-nel-lazio, 2025b. Accessed: 2025-04-22.

D. Ballas, G. Clarke, D. Dorling, H. Eyre, B. Thomas, and D. Rossiter. *Geocomputation: A Primer*. Routledge, London, 2006.

Tim Berners-Lee. Linked data – design issues. https://www.w3.org/DesignIssues/LinkedData.html, 2006. Accessed: 2025-05-11.

S. E. Bibri and J. Krogstie. Smart sustainable cities of the future: An extensive interdisciplinary literature review. *Sustainable Cities and Society*, 31:183–212, 2017. doi: 10.1016/j.scs.2017.02.016.

Capgemini Invent. The economic impact of open data: Opportunities for value creation in europe, 2020. URL https://data.europa.eu/sites/default/files/the-economic-impact-of-open-data-2020.pdf. Accessed: 2025-05-12.

Christian Chiarcos, Philipp Cimiano, Thierry Declerck, and Jan P. M. Crae. Linguistic linked open data (llod) introduction and overview. In *Proceedings of the 12th International Conference on Computational Linguistics and Intelligent Text Processing*, Pisa, Italy, 2013. Association for Computational Linguistics.

L. Codina. ¿web 2.0, web 3.0 o web semántica?: El impacto en los sistemas de información de la web. In *I Congreso Internacional de Ciberperiodismo y Web 2.0*, Bilbao, 2009.

Comunità Montana di Valle Sabbia. Dotazione di servizi pubblici e privati e accessibilità in valle sabbia. Technical report, Ufficio di Statistica della Comunità Montana di Valle Sabbia, Brescia, Italy, 2003. Analisi territoriale dell'accessibilità ai servizi e autonomia familiare sulla base del XIV Censimento ISTAT.

OpenLayers Contributors. Openlayers documentation, 2023a. URL https://openlayers.org/en/latest/doc/. Accessed: 2025-05-15.

OpenTripPlanner Contributors. Opentripplanner documentation, 2023b. URL https://docs.opentripplanner.org/. Accessed: 2025-05-15.

Carey Curtis and Jan Scheurer. Planning for sustainable accessibility: Developing tools to aid discussion and decision-making. *Progress in Planning*, 74(2):53–106, 2010. doi: 10.1016/j.progress.2010.05.001.

Data/Maps Copyright 2018 Geofabrik GmbH and OpenStreetMap Contributors. Download openstreetmap data for this region: Centro. https://download.geofabrik.de/europe/italy/centro.html, 2025. Accessed: 2025-04-04.

Google Developers. Google maps platform documentation, 2023. URL https://developers.google.com/maps/documentation. Accessed: 2025-05-05.

Nicola Dragoni, Ivan Lanese, Stephan Thordal Larsen, Manuel Mazzara, Ruslan Mustafin, and Larisa Safina. Microservices: How to make your application scale. *arXiv preprint arXiv:1702.07149*, 2017.

ESPON. *PROFECY: Processes, features and cycles of inner peripheries in Europe*. 2017a.

ESPON. *PROFECY: Processes, features and cycles of inner peripheries in Europe: Annex 2 datasets and database*. 2017b.

ESPON. Shrinking rural regions in europe. https://www.espon.eu/shrinking-rural-regions, 2020. Accessed: 2025-02-15.

European Centre for the Development of Vocational Training (Cedefop). Skills forecast: Spain. https://www.cedefop.europa.eu/en/publications-and-resources/country-reports/spain-skills-forecast, 2023. Accessed: 2025-05-12.

European Union. Directive (eu) 2019/1024 of the european parliament and of the council of 20 june 2019 on open data and the re-use of public sector information. https://eur-lex.europa.eu/eli/dir/2019/1024/oj, 2019. OJ L 172, 26.6.2019, p. 56–83.

Eurostat. Average gross monthly earnings. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Average_gross_monthly_earnings_statistics, 2024. Accessed: 2025-05-05.

Emilio Ferrara, Pasquale De Meo, Giacomo Fiumara, and Robert Baumgartner. Web data extraction, applications and techniques: A survey. *arXiv preprint arXiv:1207.0246*, 2012.

A. Ferreira, E. Beukers, and M. Te Bremmelstroet. Accessibility is gold, mobility is not: A proposal for the improvement of dutch transport-related cost-benefit analysis. *Environment and Planning B, Planning & Design*, 39:683–697, 2012. doi: 10.1068/b38073.

David Flanagan. *JavaScript: The Definitive Guide: Master the World's Most-Used Programming Language*. O'Reilly Media, Inc., 2020. Accessed: 2025-05-15.

Gartner, Inc. It asset depreciation and lifecycle management best practices, 2021. Available at: https://www.gartner.com/en/documents/1234567, accessed 2025-05-26.

Karst T. Geurs and Jan R. Ritsema van Eck. Accessibility measures: Review and applications. evaluation of accessibility impacts of land-use transportation scenarios, and related social and economic impact. Technical Report 408505006, RIVM (National Institute of Public Health and the Environment), 2001.

K.T. Geurs and B. van Wee. Accessibility evaluation of land-use and transport strategies: Review and research directions. *Journal of Transport Geography*, 12(2):127–140, 2004.

Google and Transit Developers. General transit feed specification reference, 2023. URL https://gtfs.org/reference/static/. Accessed: 2025-05-15.

International Accounting Standards Board. Ias 16 - property, plant and equipment, 2024. Available at: https://www.ifrs.org/issued-standards/list-of-standards/ias-16-property-plant-and-equipment/, accessed 2025-05-26.

ISO - UNI Ente Italiano di Normazione. Sustainable cities and communities - indicators for resilient cities, a. https://store.uni.com/uni-iso-37123-2020.

ISO - UNI Ente Italiano di Normazione. Framework for integration and operation of smart community infrastructures, b. https://webstore.uni.com/en/iso-37155-1-2020.

ISTAT. Geometry of urban centres with population in 2021. https://www.istat.it/notizia/basi-territoriali-e-variabili-censuarie/, 2021. Accessed: 2025-01-28.

ISTAT. *Atlante Statistico dei Comuni*. Istituto Nazionale di Statistica, Rome, 2022.

ISTAT. Geometry of the municipalities of the year 2024. https://www.istat.it/notizia/confini-delle-unita-amministrative-a-fini-statistici-al-1-gennaio-2018-2/, 2024a. Accessed: 2025-01-28.

ISTAT. Population data of the municipalities in 2024, 2024b. Accessed: 2025-01-28.

Chris Jacobs-Crisioni, Filipe Batista e Silva, Carlo Lavalle, Claudio Baranzelli, Alexandros Barbosa, and Carlos P. Castillo. Accessibility and territorial cohesion in a case of transport infrastructure improvements with changing population distributions. *European Transport Research Review*, 8(9), 2016. doi: 10.1007/s12544-016-0197-5.

C. Jacobs-Crisioni, M. Kompil, and L. Dijkstra. Big in the neighbourhood: Identifying local and regional centres through their network position. *Papers in Regional Science*, 102(2):421–458, 2023.

Rajat Kasliwal and Navjot Singh. Web scraping using selenium for automated data collection. *International Journal of Computer Applications*, 180(46):20–23, 2018.

M. Kompil, C. Jacobs-Crisioni, L. Dijkstra, and C. Lavalle. Mapping accessibility to generic services in europe: A market-potential based approach. *Sustainable Cities and Society*, 47:101372, 2019.

P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind. *Geographic Information Systems and Science*. Wiley, Chichester, 3 edition, 2015.

Sergio López, Josep Silva, and David Insa. Using the dom tree for content extraction. *arXiv preprint arXiv:1210.6113*, 2012.

Microsoft. Backends for frontends pattern - azure architecture center, 2023. URL https://learn.microsoft.com/en-us/azure/architecture/patterns/backends-for-frontends. Accessed: 2025-05-15.

Ministerio de Universidades, Spain. Salaries in spanish higher education, 2022. Madrid, Spain.

Ryan Mitchell. *Web Scraping with Python: Collecting More Data from the Modern Web*. O'Reilly Media, 2nd edition, 2021.

MIUR. Addresses and typology of italian public schools in 2024. https://www.mim.gov.it/-/scuola-in-chiaro, 2024. Accessed: 2025-02-05.

A. Mizen, M. King, and J. Pearce. Comparing network and euclidean-based measures of travel distance in urban accessibility analysis. *Applied Geography*, 60:36–45, 2015. doi: 10.1016/j.apgeog.2015.02.007.

S. S. Mohri, S. Mortazavi, and N. Nassir. A clustering method for measuring accessibility and equity in public transportation service: Case study of melbourne. *Sustainable Cities and Society*, 74:103241, 2021.

OECD. *Territorial Review of Italy*. Organisation for Economic Co-operation and Development, Paris, 2020.

OECD. *Rural Well-being: Geography of Opportunities*. Organisation for Economic Co-operation and Development, Paris, 2022.

Openpolis and Con i Bambini. Educational opportunities also depend on school accessibility. *Openpolis*, 2023. URL https://www.openpolis.it/le-opportunita-educative-dipendono-anche-dallaccessibilita-delle-scuole/. Accessed: 2025-05-15.

J. Parra-Domínguez, R. López-Blanco, and F. Pinto-Santos. Approach to the technical processes of incorporating sustainability information - the case of a smart city and the monitoring of the sustainable development goals. *Processes*, 10(8):1651, 2022.

Peter Pelzer. Usefulness of planning support systems: A conceptual framework and an empirical illustration. *Computers, Environment and Urban Systems*, 54:229–239, 2015.

PNE - Esplora Strutture. Agenzia nazionale per i servizi sanitari regionali. https://pne.agenas.it/ospedaliera/strutture?reg=lazio, 2025. Accessed: 2025-05-22.

A. Páez, D.M. Scott, and C. Morency. Measuring accessibility: Positive and normative implementations of various accessibility indicators. *Journal of Transport Geography*, 25:141–153, 2012.

Development Team QGIS. Qgis user guide, 2023. URL https://docs.qgis.org/. Accessed: 2025-05-15.

Leonard Richardson. Beautiful soup documentation. Online, 2007. Available at: https://www.crummy.com/software/BeautifulSoup/.

ScopicSoftware. Web application architecture: Types, components & best practices, 2023. URL https://scopicsoftware.com/blog/web-application-architecture/. Accessed: 2025-05-20.

Rizwan Shahid, Stefania Bertazzon, Merril Knudtson, and William A. Ghali. Comparison of distance measures in spatial analytical modeling for health service planning. *BMC Health Services Research*, 9(1):200, 2009. doi: 10.1186/1472-6963-9-200.

Carlos Silva, Luca Bertolini, and Marco te Brömmelstroet. Towards an understanding of the use of accessibility instruments in planning practice. *Transport Policy*, 61:87–95, 2017.

Klaus Spiekermann, Michael Wegener, Vit Kveton, Miroslav Marada, Carsten Schürmann, Olga Biosca, Artur Ulied Segui, Hannu Antikainen, Ossi Kotavaara, Jarmo Rusanen, Dominika Bielanska, Davide Fiorello, Tomasz Komornicki, and Piotr Rosik. TRACC: Transport Accessibility at Regional/Local Scale and Patterns in Europe – Final Report, Volume 2: TRACC Scientific Report. Technical report, ESPON & Spiekermann & Wegener Urban and Regional Research (S&W), 2015. Version 06/02/2015.

B. Tahmasbi, M. H. Mansourianfar, H. Haghshenas, and I. Kim. Multimodal accessibility-based equity assessment of urban public facilities distribution. *Sustainable Cities and Society*, 49: 101633, 2019.

Marco te Brömmelstroet. Designing planning support systems to facilitate meaningful collaboration: A framework based on dimensions of interaction. *Planning Practice & Research*, 30(3): 295–309, 2015.

Marco te Brömmelstroet. Characteristics of the usefulness of planning support systems: Findings from a review of dutch planning practice. *Environment and Planning B: Urban Analytics and City Science*, 44(4):570–591, 2017.

Computer Engineering Department UC3M. Lodcoremadrid project, 2024. URL https://lodcoremadrid.es/. An initiative for publishing territorial indicators as Linked Open Data.

Imran Ullah, Shakir Khusro, Azhar Ullah, and Muhammad Naeem. An overview of the current state of linked and open data in cataloging. Technical report, University of Peshawar, 2018.

Bert Van Wee, Karst Geurs, and Caspar Chorus. Information, communication, travel behavior and accessibility. *Journal of Transport and Land Use*, 6(3):1–16, 2013.

E. Verachtert, I. Mayeres, K. Vermeiren, M. Van der Meulen, M. Vanhulsel, G. Vanderstraeten, and L. Poelmans. Mapping regional accessibility of public transport and services in support of spatial planning: A case study in flanders. *Land Use Policy*, 133:106873, 2023.

W3C. DCAT - Data Catalog Vocabulary (Version 3). https://www.w3.org/TR/vocab-dcat-3/, 2022. W3C Recommendation.

Y. Xia, H. Chen, C. Zuo, and N. Zhang. The impact of traffic on equality of urban healthcare service accessibility: A case study in wuhan, china. *Sustainable Cities and Society*, 86:104130, 2022.

Evan You and Vite Contributors. Vite - next generation frontend tooling, 2023. URL https://vitejs.dev/. Accessed: 2025-05-15.

A Work plan and budget

The objective of this part is to quantify the various phases of this work in terms of time and cost. Specifically, it describes the planning and allocation of resources that supported the development of the project. It is divided into two main sections: project timeline and analysis of the cost.

1.1 Project timeline

First, a detailed work plan is presented in the form of a Gantt chart 40, illustrating the timing of activities, milestones and dependencies throughout the project lifecycle. This visual representation helps to highlight the organisation, scheduling and progress monitoring strategies adopted to ensure timely and structured execution.

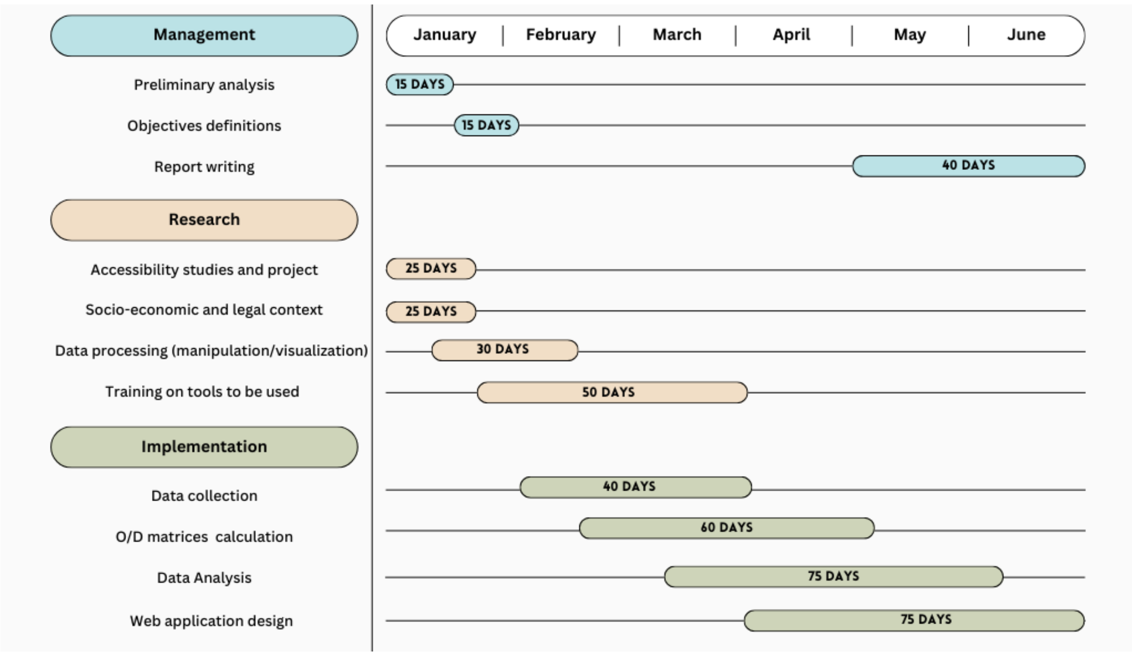


Figure 40: *Gantt chart - project timeline*

After completing the various bureaucratic procedures, the initial introductory meetings with the tutor, defining the area of work and part of the preliminary research, which took place in October 2024, the actual implementation of the project began in the first half of January. The various stages of the project have been summarised into three core areas: general management, research and implementation. Each area includes a series of tasks, distributed over the months from January to June 2025, with overlapping activities to ensure continuity and efficient time management. The whole thesis project was finally completed with the completion of the report and its presentation to the Carlos III University committee at the end of June.

1. The management phase played a crucial role in laying the groundwork for the entire project. It involved conducting a preliminary analysis to define the problem space, aligning expectations with the academic supervisor, and formulating clear objectives. This phase ensured that the project was strategically oriented and supported by a well-structured plan. The final step of this phase focused on report writing and presentation creation, allowing for an organized documentation of the results and findings, aligned with the project timeline.

- **Preliminary analysis**
  - Duration: 15 days — January 07 to January 21
  - Initial exploration of the topic, objectives, and available data sources.
- **Objectives definition**
  - Duration: 15 days — January 21 to February 04
  - When the research was almost complete. Specification of the research goals and expected outcomes in collaboration with the academic supervisor.
- **Report writing**
  - Duration: 40 days — April 30 to June 9
  - Drafting, revision, and finalization of the thesis document. This phase coincides with the conclusion of all technical tasks.

2. The research phase was dedicated to acquiring the theoretical, legal, and contextual knowledge necessary to address the project objectives. This included exploring existing studies on accessibility, understanding the socio-economic and legal frameworks influencing public service distribution, and developing familiarity with the tools and platforms used in the subsequent phases. The data manipulation and visualization sub-phase also allowed for a structured understanding of how raw data could be transformed into meaningful insights.

   - **Accessibility studies and project framing**
     - Duration: 25 days — January 07 to January 31
     - In-depth analysis of accessibility theories, studies and methodologies applied to the selected area.
   - **Socio-economic and legal context**
     - Duration: 25 days — January 07 to January 31
     - Exploration of legislative frameworks and socio-economic factors influencing accessibility in Europe and in Italy.
   - **Data processing (manipulation/visualization)**
     - Duration: 30 days — January 14 to February 14
     - Research about data cleaning, and visualization method, also for spatial .shp files using Python and GIS tools.
   - **Training on tools to be used**
     - Duration: 50 days — January 21 to March 11
     - Self-learning and tutorials on software platforms such as QGIS, OpenTripPlanner technology, API requests, JavaScript & Open-layers, or custom-built dashboards.

3. The implementation phase represented the operational core of the project. It involved the systematic collection and processing of data, the calculation of origin-destination matrices, and the analysis of accessibility indicators. This phase was characterized by the integration of technical skills, such as spatial analysis and scripting, with a focus on generating actionable outputs. Finally, a dedicated effort was made to design a web-based application capable of visualizing results in an intuitive and interactive way, thus maximizing the project's practical relevance for planners and stakeholders.

- **Data collection**
  - Duration: 40 days — February 01 to March 12
  - Gathering raw datasets (e.g., geometrical municipality and urban core data, GTFS, census, health/education locations), formatting and preparing them for distances calculation.
- **O/D matrices calculation**
  - Duration: 60 days — February 10 to April 11
  - At the same time as data collection, when data from a certain category has been collected (e.g. education). Generation of origin-destination matrices based on public transport or road network accessibility.
- **Data analysis**
  - Duration: 75 days — March 1 to May 15
  - Every time the distances data were calculated, evaluation of accessibility metrics, spatial disparities, and service coverage.
- **Web application design**
  - Duration: 75 days — March 20 to June 9
  - At the same time, after the data format has been defined and the data has been validated. Development of the interactive platform to visualize the results and make them usable by stakeholders.

It is important to note that the total number of working days indicated in the timeline, when multiplied by the average number of working hours per week, results in a higher number of total hours than initially estimated. This discrepancy arises because the calendar days reported for each phase also include weekends and public holidays, during which no work was carried out. Consequently, the effective workload should be interpreted in terms of net working days, which are fewer than the gross calendar days indicated in the timeline.

1.2 Project cost analysis

Following the work plan, the chapter provides a budget analysis, which summarises the financial resources involved in the project. This includes direct costs, such as equipment, software and potential service subscriptions, as well as indirect costs, where applicable, and also the cost of the needed human resources to carry the investigation. The aim is to assess the overall economic impact of the project and evaluate its cost-effectiveness in relation to the results achieved.

In parallel, a budget analysis is conducted not only to quantify the actual resources used but also to provide an economic perspective in case the project is to be further developed. The aim is to offer a clear understanding of the financial implications and requirements for scaling or replicating the solution, also in similar contexts.

Human resources costs

As outlined in the work plan, the project spanned approximately six months, from January to June 2025. Throughout this period, the total working hours, including preliminary meetings, research activities, and report writing, amounted to roughly 750 hours. To estimate the personnel costs associated with the project, it is essential to consider the hourly wage rates of the involved human resources.

For the primary investigator, who holds a Bachelor's degree in engineering and is responsible for research and development activities, the hourly wage has been estimated based on the average salary data for research engineers in Spain. According to the European Centre for the Development of Vocational Training (European Centre for the Development of Vocational Training (Cedefop), 2023) and Eurostat labour statistics (Eurostat, 2024), the average gross monthly salary for an early-career engineer currently in Spain is approximately €2,000 to €2,500. Assuming a 37.5-hour work week, this corresponds to an hourly wage between €13.33 and €16.67. For the purposes of this analysis, the conservative value of €15 per hour is used.

The university tutor, who contributed through supervision and project review activities, typically earns a higher hourly wage reflecting academic ranks. Based on data from the Spanish Ministry of Universities (Ministerio de Universidades, Spain, 2022), an assistant professor or lecturer's gross monthly salary ranges between €3,000 and €4,000. Considering the same weekly working hours, the approximate hourly rate is €20 to €27. For this report, an hourly wage of €25 per hour has been adopted.

Taking these figures and the estimated hours worked into account, the human resource costs for the project are calculated as follows:

Table 41: Human resources cost summary

Human Resources	Position	Hourly Wage	Hours Worked	Total Cost
Kevin Gumina	Principal Investigator	15 (€)	750	11,250 (€)
Javier Garcia G.	Tutor	25 (€)	20	500 (€)
Total				**11,750 (€)**

These cost estimates provide an overview of the human capital investment made in the project. It is important to note that actual costs may vary based on contract terms, benefits, and institutional overheads. Nonetheless, this analysis offers a useful benchmark for understanding the economic resources allocated to human resources during the project's lifecycle.

Equipment costs

For the successful execution of the project, the use of dedicated equipment was necessary. In particular, a Lenovo DESKTOP-FIVDI9N was employed as the main computing device to carry out research, data processing, and software development tasks.

To estimate the cost related to this equipment for the project duration, a depreciation (amortization) approach is adopted. Typically, the useful life of a desktop computer is considered to be approximately 3 to 5 years according to guidelines from the International Accounting Standards (International Accounting Standards Board, 2024) and industry best practices. (Gartner, Inc., 2021) For this analysis, a conservative useful life of 4 years (48 months) is assumed.

Given that the project spanned 6 months, the amortized equipment cost attributable to the project is calculated as the fraction of the equipment's initial cost proportional to the project duration. The initial purchase price of the Lenovo desktop was €1,200 (estimated market value). The amortization cost for the 6-month period is calculated as:

$$\text{Amortized Cost} = \text{Purchase Price} \times \frac{\text{Project Duration (months)}}{\text{Useful Life (months)}}$$

$$\text{Amortized Cost} = 1200 \times \frac{6}{48} = 150$$

Table 42: Amortized equipment cost summary

Equipment	Lenovo DESKTOP-FIVDI9N
Purchase price	1,200 (€)
Estimated useful life (months)	48
Project duration (months)	6
Amortized cost for project	**150 (€)**

This amortized value reflects the proportionate equipment cost consumed during the project and is included as part of the overall project budget.

Google Maps API costs

During the project, several APIs were used to obtain necessary data, particularly for calculating distances, travel times, and other relevant metrics. For example, the Google Maps Distance Matrix API and Goole Geocoding API were employed to calculate driving distances and travel times between health or educational facilities, considering factors like departure time. Or for geo-localising the public services. The cost of using each API service depends on the number of elements processed and the unit cost per element.

For each service used in the project, the total API cost is calculated by multiplying the number of elements processed by the cost per element, as specified in the pricing documentation for each service. The total amount represents the cost incurred for using the respective services throughout the project duration.

Below is a summary (table 43) that outlines the API services used, including details on the type of service, its purpose, the number of elements, price per element, and the resulting total cost for each service. In particular, the acronym DR refers to the use of the Google Maps Distance Matrix API for calculating distances by car with traffic, while PT refers to distances calculated using public transport.

Table 43: API Services cost summary

API Service	**Used For**	**Elements**	**Element price**	**Total**
Geocoding	Schools	3261	0.0040 ($)	13.00 ($)
Geocoding	Doctor ambulatory	669	0.0045 ($)	03.00 ($)
Geocoding	Emergency room	51	0.0045 ($)	00.22 ($)
Distance Matrix	Education DR	6489	0.001 ($)	64.89 ($)
Distance Matrix	Education PT	2991	0.001 ($)	29.91 ($)
Distance Matrix	Hospitals DR	2638	0.001 ($)	26.38 ($)
Distance Matrix	Hospitals PT	473	0.001 ($)	04.73 ($)
Distance Matrix	Doctor ambulatory DR	5494	0.001 ($)	54.94 ($)
Distance Matrix	Doctor ambulatory PT	1092	0.001 ($)	10.92 ($)
Distance Matrix	Emergency room DR	7534	0.001 ($)	75.34 ($)
Total				**283.33 ($)**

Google for the API services, charges in USD (United States Dollars), and the total amount spent on API usage was approximately 283.33 ($). Given that the cost of using the API in April/May 2025 was calculated in USD, the corresponding amount in EUR (Euros) needs to be determined. Using the historical exchange rate of 1 USD = 0.92 EUR during that period, the total cost in EUR can be calculated as:

$$\text{Total cost in EUR} = 283.33\,(\$) \times 0.92 = \mathbf{260.66\,(€)}$$

Indirect costs

In addition to direct costs, such as human resources, equipment and API services, there are also indirect costs associated with running a project. These costs, although not always immediately obvious, are crucial for the overall operation and should be accounted for in any comprehensive budget analysis. Indirect costs include a wide range of expenditures, such as:

- Utilities: Costs for electricity, water, and heating, which are necessary to support the daily operations of the workspace;
- Internet and Wi-Fi: Essential for research activities, communication, and data transfer;

- Office supplies and administrative support: Items such as stationery, printing materials, and other general office supplies used throughout the project;

According to the guidelines provided by University Carlos III of Madrid, indirect costs are typically estimated at 15% of the total project cost. This percentage is an approximation based on standard accounting practices and is useful for projects where the exact breakdown of indirect costs may be difficult to quantify precisely.

The total cost of the project includes both direct costs (such as personnel and equipment) and indirect costs (such as utilities, internet, and office supplies). Table 44 outlines the various costs involved in the project and calculates the total project cost by factoring in the indirect costs at a rate of 15% of the direct costs.

Table 44: Total project cost breakdown (VAT included, 21% standard rate in Spain)

Cost Category	**Cost**
Direct Costs	
Human Resources (Personnel)	11,750 (€)
Equipment (PC)	150 (€)
Google Maps API services	260.66 (€)
Total Direct Costs	12,160.66 (€)
Indirect Costs (15% of Direct Costs)	1,824.099 (€)
Total project cost	**13,984.759 (€)**

It is important to note that all costs presented in this section are considered inclusive of Value Added Tax (VAT). In Spain, the standard VAT rate currently stands at 21%, and this has been implicitly accounted for in all listed amounts.