

---

# A Methodological Approach for Data Collection and Geospatial Information of Healthy public Spaces in Peripheral Neighborhoods. Case Studies: La Bota and Toctiuco, Quito-Ecuador

---

[Ana Medina](#)\*, Diana Mosquera, [Francisco Gallegos](#)

Posted Date: 25 August 2023

doi: 10.20944/preprints202308.1792.v1

Keywords: peripheral neighborhoods; public space assessment tools; co-location network; open access data; geospatial data science; network analysis; cartography.



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

# A Methodological Approach for Data Collection and Geospatial Information of Healthy public Spaces in Peripheral Neighborhoods. Case Studies: La Bota and Toctiuco, Quito-Ecuador

Ana Medina <sup>1,\*</sup>, Diana Mosquera <sup>2</sup> and Francisco Gallegos <sup>2</sup>

<sup>1</sup> Universidad de Las Americas; anagabriela.medina@udla.edu.ec

<sup>2</sup> Diversa; diana@diversa.studio; francisco@diversa.studio

\* Correspondence: anagabriela.medina@udla.edu.ec.com; Tel.: +593 2 398 1000 .

**Abstract:** Adequate public spaces and urban green areas are key criteria for urban development and infrastructure implementation of healthy cities. Latterly, there has been an increasing number of research methods using artificial intelligence (AI) to monitor, quantify, and control the state of these spaces with an aim towards pioneering research in Urban Studies. However, in informal areas, open-data access tends to lack adequate and updated information, making it difficult to use AI methods. Hence, we propose a methodology for restricted open data collection and preparation for future use in machine learning or spatial data science models for similar areas. To that extent, we examine two peripheral and low-income neighborhoods in Quito, Ecuador, La Bota and Toctiuco, to analyze their public spaces, urban green areas, points of interest, and road networks, and how they address healthy cities criteria. We develop an original methodological approach that combines an index of proximity, accessibility, quantity, and quality of these spaces with geospatial and network analysis techniques. Results indicate that the connectivity and structure of these spaces are centralized and nodal, representing exclusion and segregation. This work provides insights on potential healthy spaces, and information to urban planners and policymakers in decision-making for healthy urban infrastructure.

**Keywords:** peripheral neighborhoods; public space assessment tools; co-location network; open access data; geospatial data science; network analysis; cartography

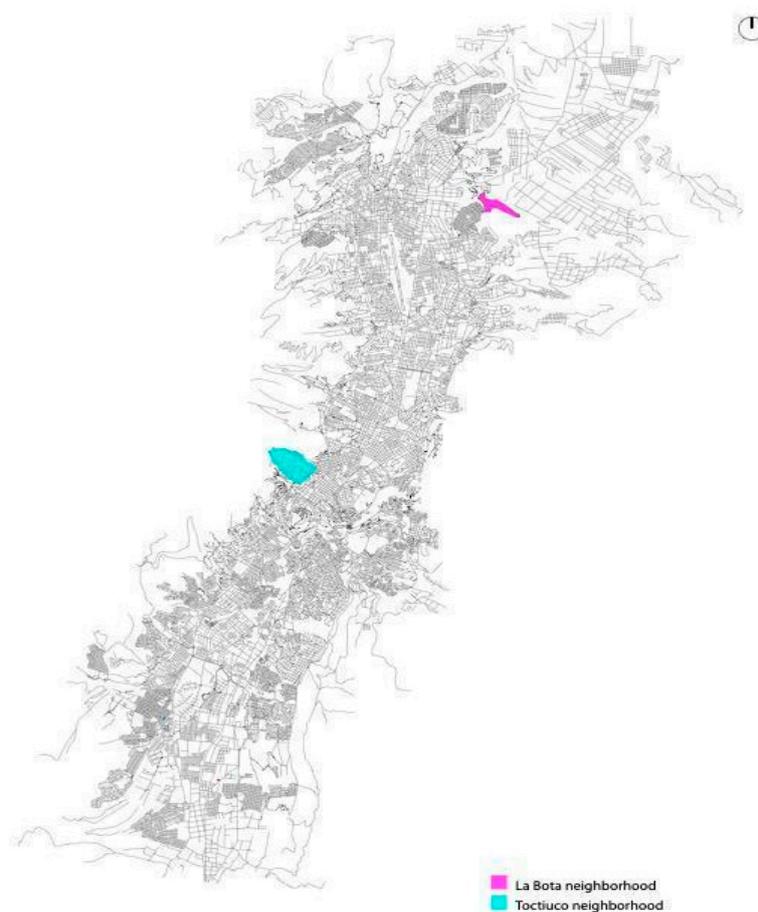
---

## 1. Introduction

In contemporary urban planning, the integration of digital technologies and artificial intelligence (AI)-driven engines has become pivotal for collecting, analyzing, and monitoring urban infrastructure data to foster healthier environments [1–3]. A critical aspect of measuring healthy urban parameters for inhabitants is the presence of urban green space (UGS) and natural outdoor environments [4,5]. These spaces have a substantial impact on the well-being of residents and AI technologies can play a crucial role in determining their proximity, accessibility, quantity, and quality. However, in places where access to open data is limited and urban infrastructure is insufficient or poor, as in informal settlements of the urban Global South, this process could be challenging and inaccurate.

In Latin America, disparities in UGS are closely linked to socioeconomic status [6], with low-income and peripheral neighborhoods often lacking in their quality [7,8]. The outbreak of the COVID-19 pandemic further exacerbated existing disparities in health, social, and economic outcomes, particularly within micro-level localities of major cities in this region [9,10]. Informal settlements in Latin America are predominantly vulnerable, experiencing higher mortality rates and increased instances of psychological and social disorders among their inhabitants [11]. Along with the lack of policies and effective urban design strategies in these often conflict-ridden areas, there has arisen an increasing debate over strategies for bringing basic healthy spaces [12] and equipment for mental and physical health to these localities [13].

To address these challenges and to offer a more comprehensive perspective on peripheral and low-income neighborhoods with limited open data access and insufficient infrastructure, this study provides a physical analysis of a set of existing-conditions factors — UGS, public spaces, roads, and points of interest (POI) — for two peripheral neighborhoods of Quito, Ecuador: Toctiuco and La Bota (Figure 1). These two neighborhoods have been the site of an ongoing public space and community-design-engagement research project that we, the authors have developed since 2019. For the purpose of examining existing-conditions factors, we present a novel methodological approach toward integrating AI methods in the examination of informal neighborhoods in the Global South that utilizes geospatial and network analysis techniques to develop an index of proximity, accessibility, quantity, and quality for these spaces. The broader purpose of this methodology is to make visible a network of potential healthy environment criteria.



**Figure 1.** Quito map highlighting La Bota and Toctiuco neighborhoods. Source: Authors.

We use a combination of neighborhood scale representation methods using Google Places API and Microsoft Planetary Computer as well as spatial clustering modeling. By employing cutting-edge technologies for data collection and analysis, our study aims to contribute to the location and enhancement of public spaces to reach healthy conditions and potentially inform policymakers, urban planners, and researchers in their efforts to create more inclusive and sustainable urban environments for peripheral and low-income neighborhoods.

In structure, this paper provides, first, a brief introduction to the two study neighborhoods. Second, in a discussion of methodology, we examine spatial data collection, geospatial layer information, and referent studies. Third, we identify, collect, and analyze data of spatial amenities and characteristics of the two case studies by using different open-source platforms. Fourth, we analyze the data collected and highlight key findings related to UGS, public spaces, roads and POI

and present the resulting values. Finally, we present conclusions of the methodological approach to spatial data collection.

#### *Brief approach to La Bota and Toctiuco*

La Bota neighborhood is located on the northeast side of Quito, named for its “boot” shape as seen from above. It has a low natural slope at its center but is surrounded by steep slopes along its perimeter. La Bota has comprises area of 58,95 ha. and, according to the most recent projections, a population of nearly 13,000 inhabitants whose main income is commerce and services [14]. During the 1970s and 1980s, an oil boom across the country and the concomitant intensification of industry and urban construction drove a wave of rural-to-urban migration into the Ecuador’s main cities [15]. Confronting this challenge, between 1974 and 1976, a team of professors and students of the Faculty of Architecture and Urbanism at Central University of Ecuador was to design an urban and architectural plan for two new neighborhoods, “Comité del Pueblo” and “La Bota”, out of what was formerly a private estate, “Hacienda Eloísa.” This study, “Towards a Popular Urbanism” [*Hacia un Urbanismo Popular*], configured the new neighborhoods as modular, with lots of 9x21 meters (189 m<sup>2</sup>) and blocks comprising either 12 or 16 lots. The design further organized these blocks into ‘cells’, corresponding to six aggregated blocks (84 lots), two of which were dedicated to essential neighborhood infrastructures. Moreover, the design organized these cells into ‘zones’ or ‘cell ensembles’, each of which gathered four cells — 336 lots each of which eight were dedicated to essential infrastructures [16]. Despite this intricate plan, the dedicated urban infrastructures never achieved its intended function owing to a lack of oversight in both management and construction from the local government.

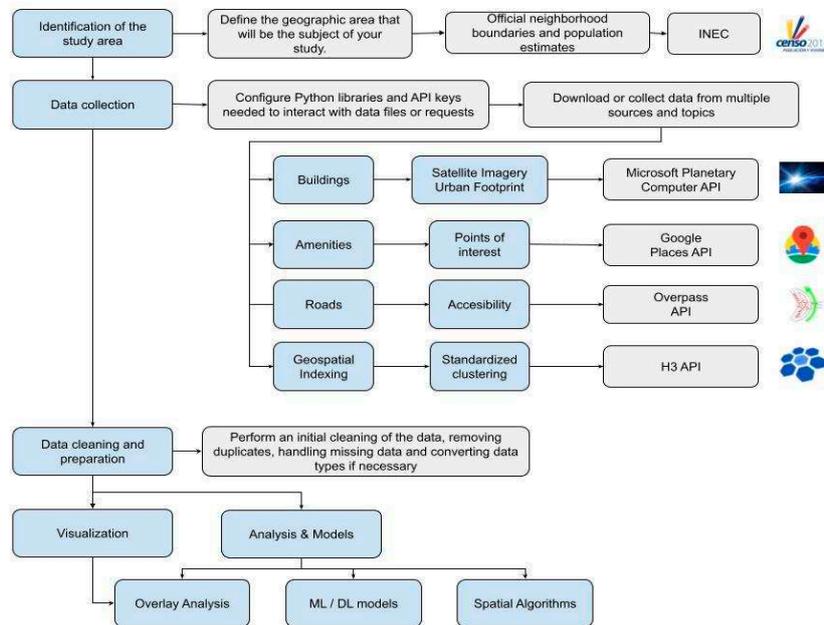
Toctiuco, whose name means “corn flower,” is a neighborhood traditionally inhabited by low-income socio-economic groups. It is located in the west side of Quito, on the slopes of the Pichincha mountain chain. The lowest point considered in our study has an altitude of 2,950 m.a.l.s.; the highest, meanwhile, rises to 3,300 m.a.l.s. The latter offers a wide view of the city and its valley but also maintains, on average, a temperature 4<sup>o</sup>(C) lower than the city center below [17]. The neighborhood has a population of 40,000 according to the last projection studies from 2022 [18]. Its origin dates from 1954 when, similar to La Bota albeit decades earlier, rural and largely Indigenous populations migrated from the countryside to the urban center of Quito. Given the severe lack of housing in the city at the time, especially for poor new arrivals, many of these migrants ‘invaded’ the officially privately-held land [15]. What became the present-day Toctiuco neighborhood was then a *hacienda* owned by the religious order “*Hermanos Mercedarios*” [19]. As in La Bota, the configuration of the neighborhood was influenced by rural-to-urban migration to Quito, especially by that of the later oil boom of the 1970s. This phenomenon directly changed the morphology of this and other urban centers across Ecuador as people established their dwellings little by little in an informal and unplanned manner much of which has remained this way [20,21].

## **2. Materials and Methods**

In order to make visible and provide spatial evidence about the current adverse conditions people endure everyday, due to the lack of adequate urban infrastructure in informal neighborhoods of the Global South, it is useful to analyze the open data network and the succeeding produced cartographies of such information. This paper draws on recent advances in the field of geospatial information of urban areas at a neighborhood scale and applies it to Toctiuco and La Bota.

In the face of the many urgent and critical changes taking place in Global South cities, there has grown an increasing interest in measuring and quantifying critical aspects of critical urban infrastructure. Toward this end, this article follows and measures the WHO key parameters that contribute to healthy urban environments: access to green spaces, hospitals, efficient mobility infrastructure, connectivity with neighboring communities, among the most important [13]. The intention is to develop a more comprehensive spatial understanding of these type of neighborhoods to reach minimum healthy cities parameters. To this end, we employ a combination of representation and spatial analysis techniques in the face of the limitation to open access information.

Data for this article includes a range of geographic information types: sociodemographic data, amenities of zones, road information, urban density and buildings, spatial indexing, and surfaces and elevations. In combining this range of geographic datasets, we present a methodology (Figure 2) for the study of healthy cities criteria in peripheral and deprived neighborhoods of Global South cities where more typical open-data access is frequently limited.



**Figure 2.** Methodological process.

The methodology begins by delimiting the study area, using the neighborhoods' boundaries and population data found in the shape file of the 2010 National Census (INEC) (The last official Census in Ecuador was carried out in 2010. However, there are official projections by INEC (National Institute of Statistics and the Censuses, from its Spanish initials) produced in these years. We use the 2022 population projection). With these boundaries so defined, we then correlate this study area with data from a set of available open-data sources. Four sources specifically help to understand the urban dynamics of the study area: Microsoft Planetary Computer API, Google Places API, Overpass API, and H3 API.

In addition to the use of available open-data sources, our approach draws upon open-source technologies, namely Python, to connect between Application Programming Interfaces (API) and to manage data download in each layer. It is necessary to have previously configured the program with respect to geospatial analysis libraries, data management, access credentials (for API's) and data visualization.

Further, with information from satellite images, we extract polygon data from the estimated footprint area of buildings within the study area. This layer of information comes from Microsoft Planetary Computer API, specifically the Urban Footprint dataset. This layer represents an Artificial Intelligence model previously trained to build polygons of building footprints from historical records of images on roofs. By utilizing the Google Places API, we can access a wide range of amenities and services available in a specific area which provides valuable information regarding geo-locational coordinates, use-type categories, and physical status. Moreover, with Overpass API, we retrieve information related to accessibility in internal circulation, differentiation of structures by use type, and the function and condition of roadways. Finally, in order to generate a homogeneous comparison and aggregation method, we have used H3 API, a global spatial indexing system.

Having collected the above data for the study area, we undertake a thorough cleaning and formatting process to ensure compatibility and uniformity among all information layers. With this prepared dataset, we are able to seamlessly create analyses, develop machine learning or deep

learning models, and apply geospatial algorithms effectively. This streamlined approach enhances reproducibility, empowering us to derive valuable insights on peripheral neighborhoods where accurate or current information is not available.

While this digital representation provides a spatial approach to the respective case studies, it is crucial to understand the underlying social realities that shape them on a daily basis. For that purpose, we first create a list of potential data sources that could be analyzed as part of further qualitative research (Table 1):

**Table 1.** Database and Source of digital images.

Layer	Source	Observations
Points of Interest (POIs)	Open Street Maps API	Data is 100% open source Some records are not up to date Relatively less data coverage
Points of Interest (POIs)	Google places API	These data have a quota (API usage limit) More frequently updated Increased data coverage
Buildings Footprints	Microsoft Planetary Computer	Data is 100% open source Requires knowledge in geo-parquets
Roads and paths	Open Street Maps API	Data is 100% open source Includes categorization of road types
Sociodemographic	Censo INEC	Only official sociodemographic data available in Ecuador. Data are not up to date; last census collected in 2010. Yet, we also use the 2022 INEC population projections.
Hexagonal Hierarchical Geospatial Index System	H3 API	Data is 100% open source Flexible to the size of the aggregation cells This allows us to have comparable areas in any case study
Topography	Municipio de Quito	Source: Authors

### 3. Results

#### 3.1. Open-source spatial analysis

##### 3.1.1. Amenities

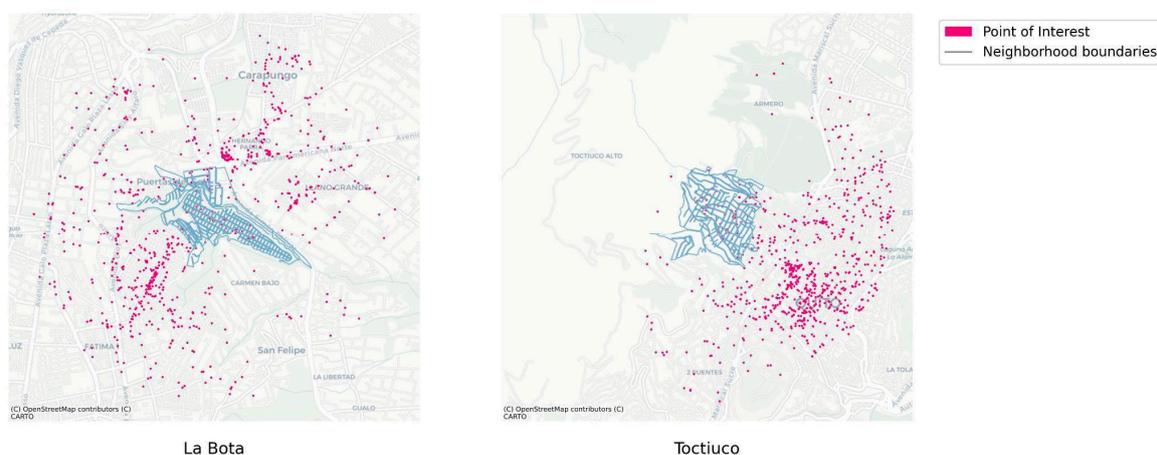
In order to obtain a more detailed context of La Bota and Toctiuco, we gather information related to services and POIs in throughout these two locations for their relevance. We initially use Open Street Maps (OSM), a collaborative open-source project that creates and provides free geospatial data. Anyone can contribute to OSM by adding, editing, and updating geographic information on the cartography [22]. The OSM community is responsible for maintaining and improving the data, resulting in a detailed and up-to-date geospatial dataset. OSM allows free access to its data and is based on the Open Database License (ODbL). This API provides data on points of interest (POIs) such

as parks, schools, green areas, stores, hospitals, and bus stops. However, we found that OSM had limited mapped data in these two neighborhoods.

Given the limited data available via OSM, we resorted to using Google Places API to supplement further data. This is a paid service provided by Google through its API which allows developers to access information about places and POIs, albeit with a limited number of free queries. To obtain POI data using Python we first acquire an API key from Google Cloud Console, and then we send requests to the API endpoint using the Python library "requests" which allows HTTP requests to be sent to the web API. These requests include the API key and the desired location, either by latitude/longitude coordinates or by the name of the place. The API responds with a JSON object containing information about the POIs at the requested location. This information is then parsed using a "json" library to extract the necessary data, such as the name, address, and type of each POI. The collected data provides an idea of the available services in La Bota and Toctiuco. This series of API requests was carried out using 1-km and 2-km radius perimeters. It was found that the 2 km radius permitted the identification of a greater number and typologies of POIs.

We collect data on a wide range of amenities including: restaurants, ATMs, banks, bars, beauty salons, bicycle stores, book stores, bus stations, cafes, car repairs, car washes, cemeteries, churches, city halls, cycling routes, dentists, doctors, farmland, fire stations, forests, funeral homes, gardens, gas stations, gyms, graveyards, hospitals, green spaces, laundries, libraries, liquor stores, local government offices, night clubs, parks, parking lots, playgrounds, pet stores, pitches, pharmacies, physiotherapists, police stations, post offices, schools, shopping malls, ramps, spas, stairs, stadiums, stores, supermarkets, taxi stands, transit stations, universities, and veterinarians.

Together, these amenities deliver an informatic portrait of the services available, their total number, and the proximity they have in our case studies (Figure 3). In Toctiuco we identify 745 amenities, and in La Bota 786 amenities with the following variables: business status, geometry, icon, icon background color, URL, name, opening hours, place ID, rating, reference, scope, types, user ratings total, vicinity, photos, and plus code.



**Figure 3.** Magenta dots represent the Points of Interest, 2km around in La Bota and Toctiuco.

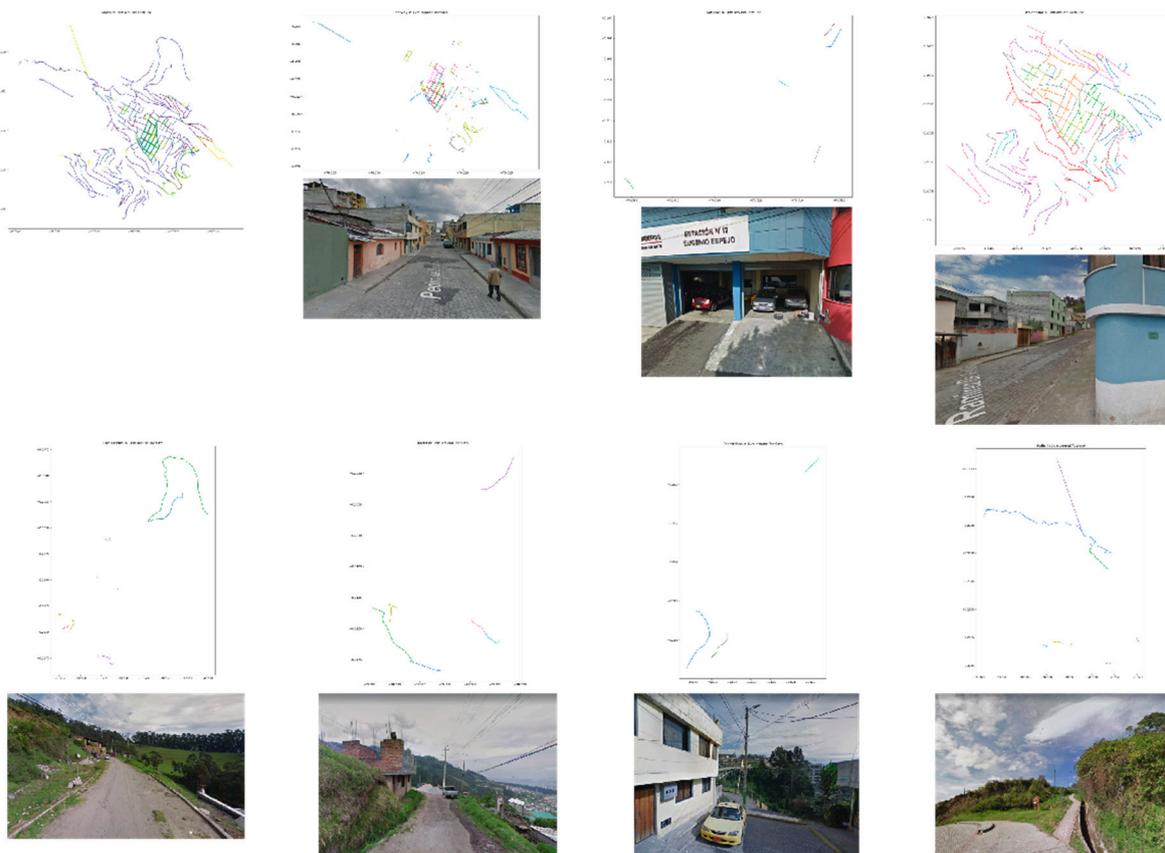
### 3.1.2. Roads, connectivity, and accessibility

The condition of primary and secondary streets in La Bota and Toctiuco, makes visible the accessibility and connectivity of these areas both within the respective neighborhoods and in their connection to surrounding areas, with a special focus on the quality to reach UGS within each neighborhood. This analysis reveals information on the availability of public transportation, pedestrian and cycling infrastructure, and road conditions for the local population. To access this data, we use Open Street Maps API, which contains the following variables: type, geometry, alt name, highway, name (Figure 4). The variable 'highway' provides information on the type of road, which can be categorized as residential, service, footway, unclassified, track, path, or pedestrian.



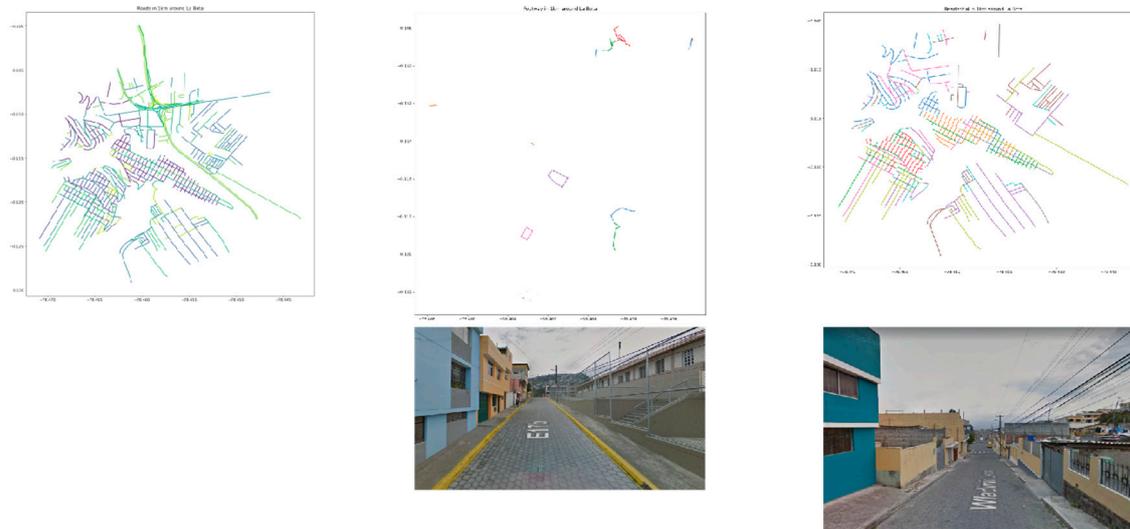
**Figure 4.** Neighborhood roads 1 km around in La Bota and Toctiuco.

In Figure 5, Toctiuco contains a greater variety of road types, yet they do not connect within the neighborhood although they do make connections with the surrounding area. Footways, used by pedestrians, are well connected in the most defined blocks (rectangular shapes). Within the natural and built border, the shape of blocks is more irregular as they follow the geography of gorges and slopes. In this area, connection and accessibility are low and difficult, showing the presence of many unpaved roads which also cause a physical segregation between the inhabitants and the urban equipment.



**Figure 5.** Toctiuco roads 1 km around, with their respective photos on site.

In La Bota (Figure 6), the majority of the study area is covered by paved roads used by cars. Sidewalks are narrow or nonexistent in many streets, making it very difficult for inhabitants to walk within the neighborhood. Here, connectivity is well covered in the neighborhood, but accessibility is low or null.



**Figure 6.** La Bota roads 1 km around, with their respective photos on site.

### 3.1.3. Buildings Footprint

An important part of this analysis is to understand how La Bota and Toctiuco are shaped by the urban fabric and population density by contrasting the open data access with reality. For that purpose, we use the last official census in Ecuador from 2010, which shows specific data about buildings footprint in neighborhoods, but also, the 2022 population projection, from which we can extract an approximation of density, housing, and population. It must be noted that this data is not precise, and indeed the unavailability of accurate and up-to-date census data is a recurrent problem in Ecuadorian cities.

For that purpose, we use Microsoft Planetary Computer API, which retrieves precise buildings geometrics. The building footprints are 2D contours, and we examine them at 1-km scale in both neighborhoods. These footprints can be used to create detailed cartographies of urban areas, to estimate population density, and to track changes in urbanization over time. First, the connection to Microsoft Planetary Computer API is established by using the corresponding connection string. Then, a request is made to access the layer of building footprints in Ecuador. We verify that the date with better data quality is "Ecuador\_2022-07-06", as opposed to the latest version of API. Once the data is obtained, the geo parquet file is ready for processing, and the analysis corroborated the number of partitions and columns in the dataset, which is important to access the bounding box of the continental territory of Ecuador. To complete the preparation of the necessary data, the shapefiles corresponding to the neighborhoods are prepared for a spatial union that is made between the obtained GeoDataFrame and the mentioned shapefiles, allowing analysis and extraction of the relevant information. The results for La Bota estimate 1,234 buildings and for Toctiuco, 1,007 (Figure 7).

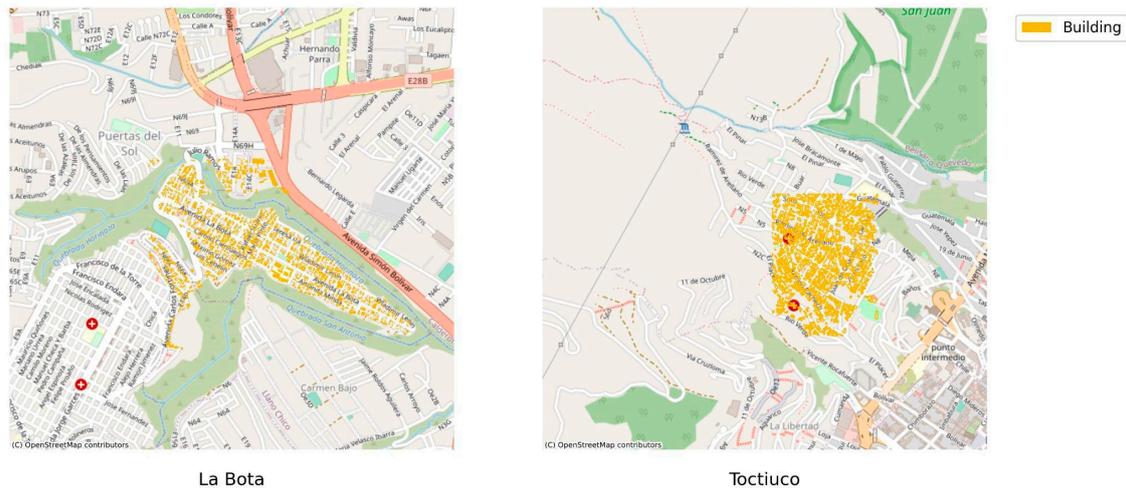


Figure 7. Buildings footprint.

### 3.1.4. Hexagonal Hierarchical Geospatial Indexing System (H3)

The Hexagonal Hierarchical Geospatial Indexing System (HHGIS) is a method of indexing and organizing geographic data based on a hexagonal grid. This indexing system is designed to efficiently store and query geospatial data at various scales and resolutions. It also organizes data into nested levels of increasing granularity. Each level of the hierarchy is presented by a different hexagonal grid, with finer grids nested within coarser grids. Through this organization, HHGIS can efficiently query data at different levels of detail, from a global view to individual city blocks. The ability to represent geographic data at multiple scales without losing precision or detail helps to indicate more specific geographic data of an area. For instance, a larger hexagon can represent a city block while a smaller one might represent a building or a tree.

For our purpose, we use HHGIS in various applications, such as the analysis of geospatial data. It is important to choose the area of the coordinate system in both neighborhoods (see Figure 8), for what we selected a resolution of 10, that means 15,047.502 m<sup>2</sup>, or about 1.5 urban blocks.

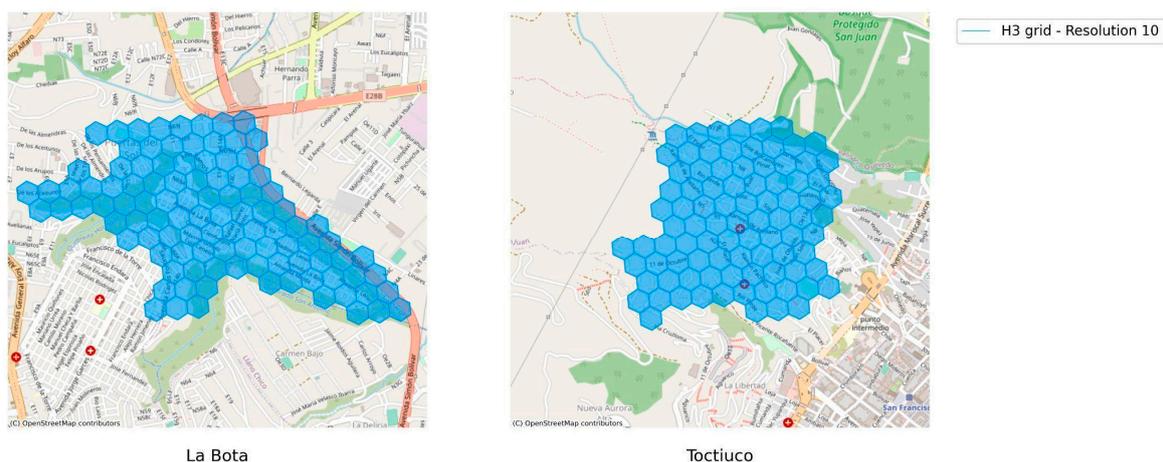


Figure 8. Hexagonal Hierarchical Spatial Index.

### 3.2. Network analysis

In this section, we present an analysis of the two case studies and highlight key findings related to the importance of public spaces and UGS in creating healthy neighborhoods based on Setha Low's "Why Public Space Matters" [23]. Low identifies six areas that demonstrate how public space contributes to a healthful neighborhood:

- Social justice and democratic practices
- Health and well-being
- Play and leisure
- Informal economy and social capital
- Environmental and ecological sustainability
- Cultural identity and attachment to place

In our analysis, we define public spaces as including sidewalks, streets, playgrounds, plazas, parks, beaches, libraries, art museums, historical monuments, and other elements that constitute the urban fabric of the built environment. These spaces are vital for socialization, learning, and play, and serve as infrastructure for inclusivity [24].

To gain a comprehensive understanding of the physical space of Toctiuco and La Bota, we collect data on various amenities and services (that is, POIs) as well as streets, sidewalks, pedestrian paths, buildings in 2D, Census data, and H3 data. These datasets enabled us to visualize and analyze the digital information, providing a closer representation of the physical space in these neighborhoods.

**Table 2.** Amenities by typologies qualitative.

Category	Subcategory	Types (examples)
Built Environment	Transportation	Bus stations Transit station Parking
Built Environment	Public Space	Park Stadium Plaza
Built Environment	Entertainment	Shopping mall Tourist attraction
Education	Schools	Primary school Secondary school University
Government	Local Government office	City hall
Government	Public services	Police Fire station
Local Economy	Food and Drinks	Restaurant Meal delivery Meal takeaway Bakery Bar Nightclub
Local Economy	Finance	ATM Bank
Local Economy	Entertainment	Museum Movie theater
Local Economy	Services	Beauty salon

		Hair care Car repair Car wash
Health	Health Services	Dentist Doctor Pharmacy Hospital Physiotherapist
Natural Environment	Entertainment	Camp Ground

To better organize and represent the POI data, we adopt categories from the World Health Organization's Healthy Cities Effective Approach to a Rapidly Changing World [13]. These categories include: structuring categories, subcategories, and types in this manner, we establish a more organized framework for organizing and representing the POI data (Figure 9) (Table 3).

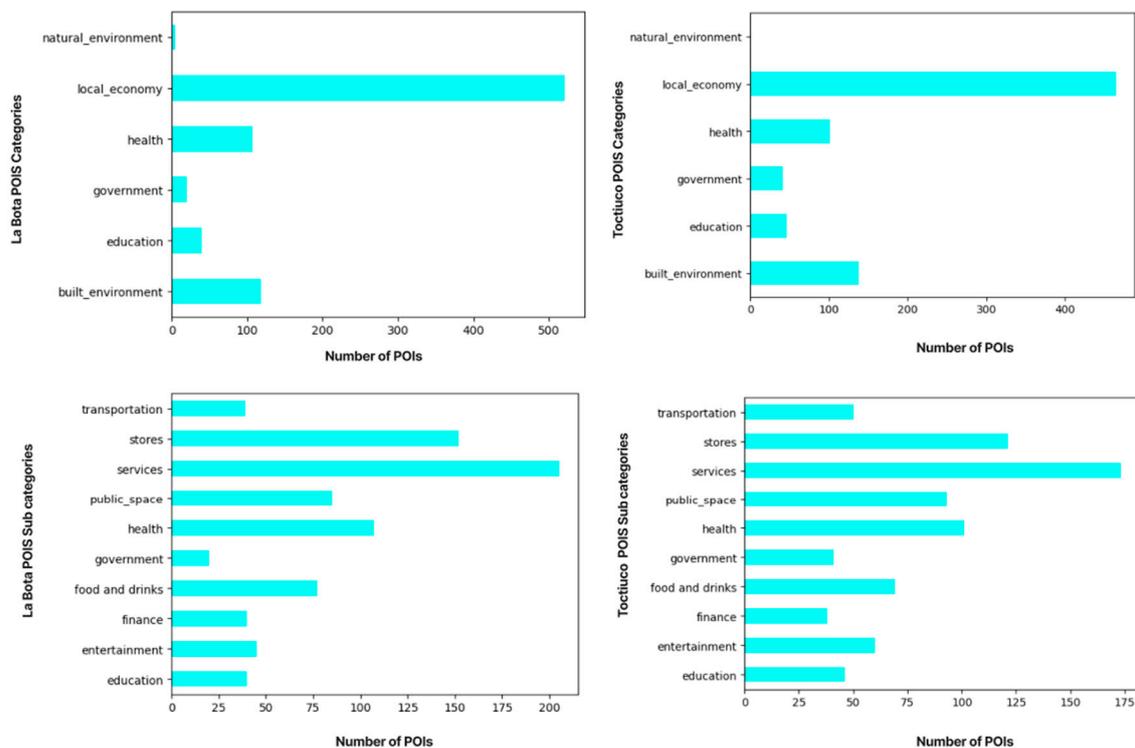


Figure 9. Exploratory Data Analysis (EDA) Categories and Subcategories.

Table 3. Amenities by study area – Quantitative.

Study Area	Category	Subcategory	Number of POIs
LA BOTA	Built environment	Public Space	84
		transportation	33
	education	education	39
	government	government	20
	health	health	106
	Local economy	entertainment	41

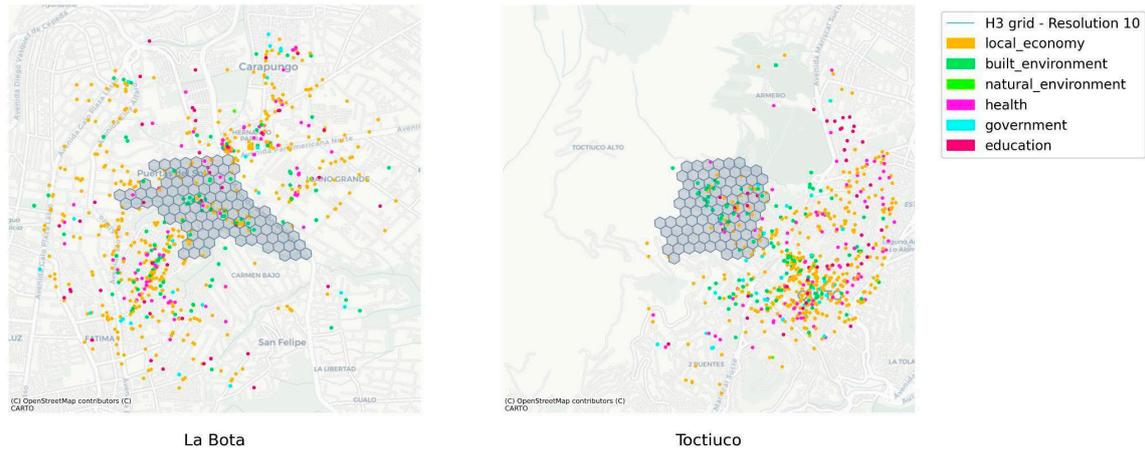
		finance	40
		food and drinks	76
		services	203
		stores	149
		transportation	6
	Natural environment	entertainment	4
TOCTIUCO	Built environment	entertainment	4
		Public space	92
		transportation	41
	education	education	44
	government	government	41
	health	health	101
	localeconomy	entertainment	55
		finance	38
		food and drinks	68
		services	171
		stores	115
		transportation	9
	naturalenvironment	entertainment	1

In both neighborhoods, the predominant category of POIs is *Local Economy*, with approximately 500 POIs in La Bota and 400 POIs in Toctiuco. This finding indicates the presence of numerous small businesses such as stores, hairdressers, and restaurants. These data reveal the significant presence of local businesses and enterprises owned by residents, meaning that the local economy relies heavily on small establishments owned by the local population.

### 3.3. Principal uses

#### 3.3.1. Hexagonal Hierarchical Geospatial Indexing System (H3) with Points of Interest

In our study, we conduct a comprehensive analysis by leveraging the layer crossing technique and employing H3 geo-indexing with points of interest (POIs). The primary objective of this analysis is to quantify the number of POIs within each hexagonal region, and to categorize them based on their respective categories or subcategories. By visualizing this data on a cartography (Figure 10) using contrasting colors to better understand the connections and spatial representations, we identify areas with the highest concentration of amenities and gain insights into the types of facilities present.



**Figure 10.** H3 and POIs. The colors represent each POI category.

Through the application of layer crossing, we merge multiple layers of geographic information, including POIs and H3 hexagons, to establish relationships and derive meaningful insights. This approach allows more effective assessment of the spatial distribution and density of amenities across the study area. Using H3 geo-indexing offers a hierarchical hexagonal grid system; we organize and index the POIs based on their precise geographic locations within each hexagon.

By conducting this analysis and visualizing the results on a cartography, we identify hotspots with high concentrations of specific types of amenities. The color-coded cartography allows one to discern areas where particular categories or subcategories of POIs are more prevalent. For example, Figure 11 represents areas with higher concentration of amenities within the corresponding hexagons. This information offers insights into the spatial patterns of amenity distribution and highlights areas with an abundance of specific facilities.



**Figure 11.** H3 + POIs Density.

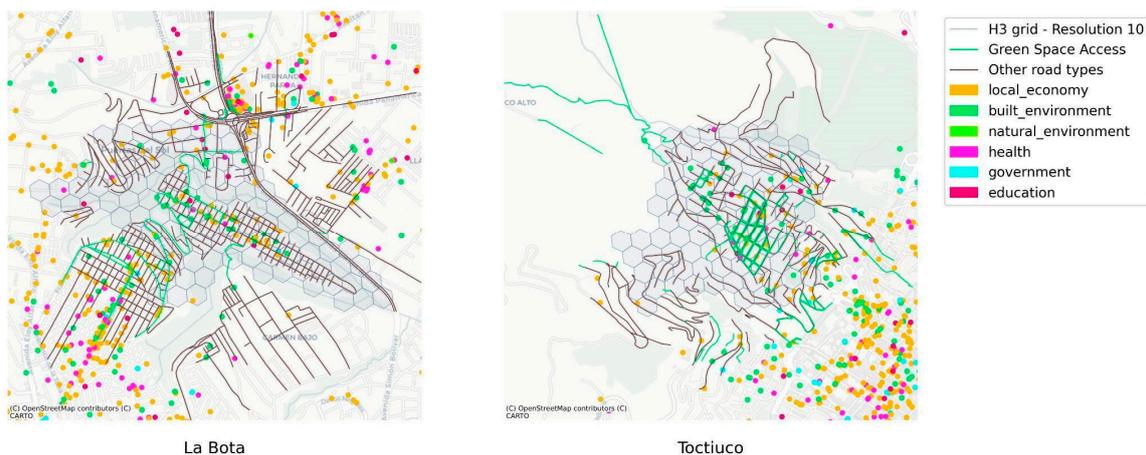
Furthermore, by expanding the scope of our analysis to encompass neighboring areas, we assess the relative deficiencies in amenities. By comparing the concentration of POIs in different hexagons and their surrounding regions, we are able to identify areas with relatively lower numbers of amenities. As a result, we uncover potential deficiencies in access to specific facilities like UGS in the given case study neighborhoods. This information can be of great significance for urban planners, policymakers, designers, and local authorities as they work towards enhancing the quality of urban conditions and ensuring equitable access to essential amenities for residents.

### 3.3.2. Hexagonal Hierarchical Geospatial Indexing System (H3) with OSM

In addition, we conduct a cross-layer analysis by incorporating OpenStreetMap (OSM) roads data into our study. The primary objective of this analysis is to assess the quality of road infrastructure and evaluate the accessibility of the study areas. The OSM roads data provides information related to road categories, such as high-speed conditions, pedestrian paths, and one-way streets, among others.

By integrating the H3 hexagonal grid system with the OSM roads data, we explore the relationship between road infrastructure within each hexagonal region and the accessibility based on the road network characteristics. The road categories derived from the OSM roads data delivers insights into the nature and functionality of the road network. High-speed roads, for example, indicate major transportation arteries facilitating rapid movement between areas while pedestrian paths indicate routes that prioritize foot traffic and promote walkability. One-way streets, meanwhile, suggest unidirectional flow of traffic, usually low-speed, which may influence access and connectivity within and outside the studied areas.

By examining the intersection of the H3 hexagons and the OSM road categories, we gain understandings into the quality and accessibility of each hexagonal region (Figure 12). For example, hexagons intersecting with high-speed roads may indicate areas with efficient transportation links and easy access to various amenities. Hexagons intersecting with pedestrian paths may signify neighborhoods that prioritize walking and have enhanced pedestrian infrastructure, potentially leading to improved accessibility to nearby amenities.



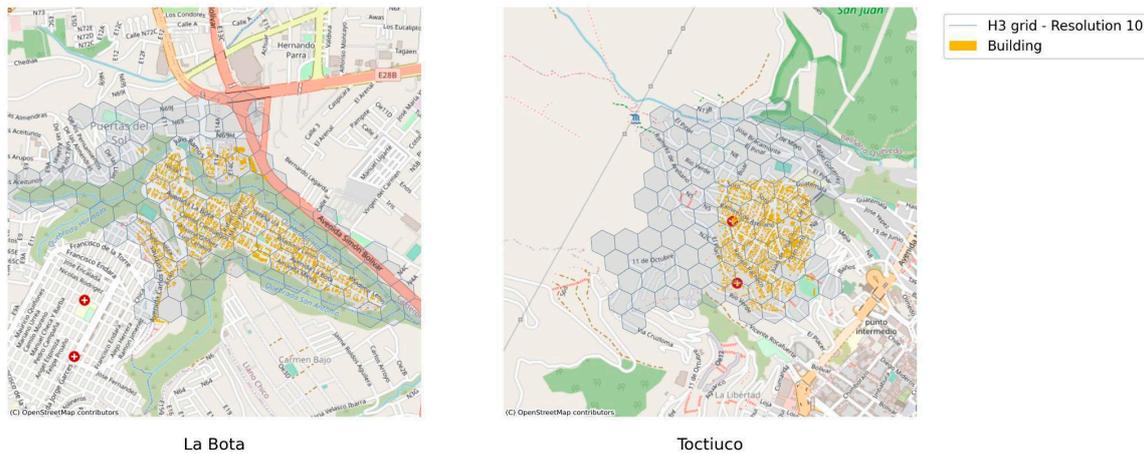
**Figure 12.** H3, POIs and Roads. The green line represents the types of roads used to access green spaces.

Despite these useful insights, we face challenges in identifying green areas and assessing their accessibility using open data sources. The availability and accessibility of green spaces play a vital role in promoting the health and well-being of urban residents, but mapping and understanding access to these areas solely through open data sources is complex.

We observe that in both case study neighborhoods only a limited number of pathways provide access to nearby green spaces due to surrounding gorges. Specifically, pedestrian routes are the most common type of access, while vehicle access was less prevalent. This finding suggests a prioritization of walkability and a focus on promoting pedestrian-friendly environments. Despite the challenges in identifying precisely green areas with open data sources, our cross-layer analysis incorporating modern map visualization, allows us to gain insights into the relationship between road infrastructure and green areas. This information can guide decision-making processes related to urban planning, transportation improvements, bus stops as public space, and the promotion of inclusive and accessible communal zones.

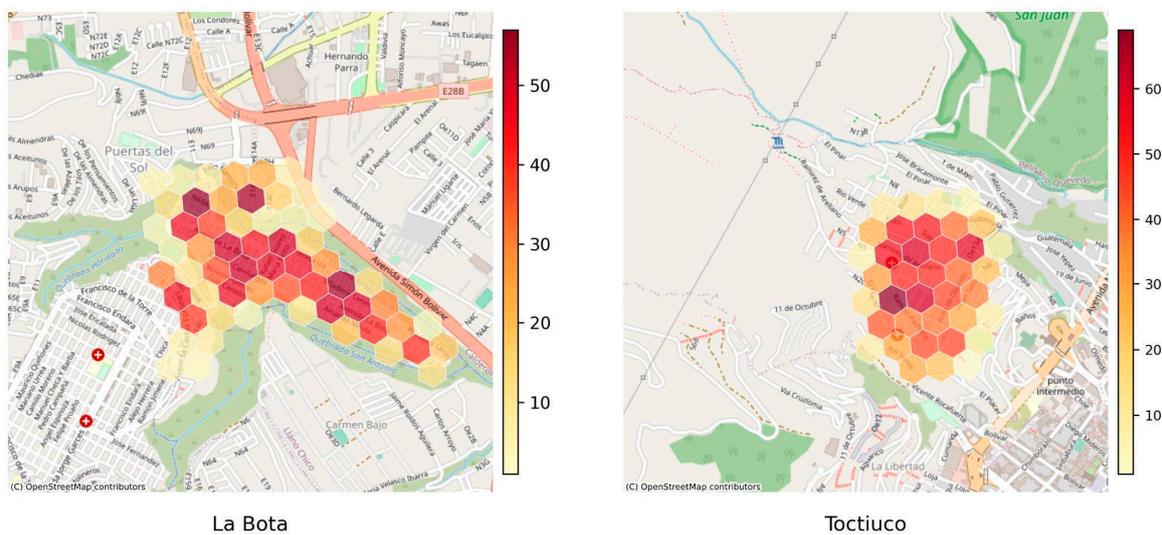
### 3.3.3. Hexagonal Hierarchical Geospatial Indexing System (H3) with MSPC Buildings Footprints

By combining H3 geo-indexing with the Microsoft Planetary Computer Building Footprint Layer, we are able to assess the density and overcrowding in these peripheral zones. Likewise, by intersecting the H3 hexagons with the detailed building polygons generated from the Microsoft Planetary Computer Building Footprint Layer, we quantify the number of buildings within each hexabin (Figure 13).



**Figure 13.** H3 and Buildings.

By identifying areas with a higher concentration of buildings within a limited space, we identify areas that may experience higher population density and potential challenges related to overcrowding (Figure 14).



**Figure 14.** H3 heat cartography and Buildings.

### 3.3.4. POI — Network Analysis

To explore the potential impact of POIs on analytical model development, we conduct a network analysis to understand the spatial context of these points [25]. Our primary objective is to visually represent the density network [26] of amenities in the study areas of the POIs categorized by their types. The network analysis permits assessing the connectivity and proximity of different POIs within specific categories, further enabling us to evaluate the availability and accessibility of amenities.

Understanding the spatial relationships between health-related POIs is vital for examining possible health infrastructure and urban design strategies in these neighborhoods.

To build the network, we calculate distances between all pairs of POIs using the geographic coordinates of each location. The distance calculation utilizes the geodesic distance formula, based on the law of cosines for spherical triangles, as given by:

$$a = \sin^2(\Delta\text{lat}/2) + \cos(\text{lat1}) * \cos(\text{lat2}) * \sin^2(\Delta\text{lon}/2)$$

$$c = 2 * \text{atan2}(\text{sqrt}(a), \text{sqrt}(1-a))$$

$$d = R * c$$

Where:

- **lat1** and **lat2** are the latitudes of the two points in radians.
- **lon1** and **lon2** are the longitudes of the two points in radians.
- **Δlat** is the difference between the latitudes (**lat2 - lat1**).
- **Δlon** is the difference between the longitudes (**lon2 - lon1**).
- **R** is the radius of the Earth (e.g., in meters or kilometers).
- **d** is the geodesic distance between the two points in the unit of measurement used for the Earth's radius.

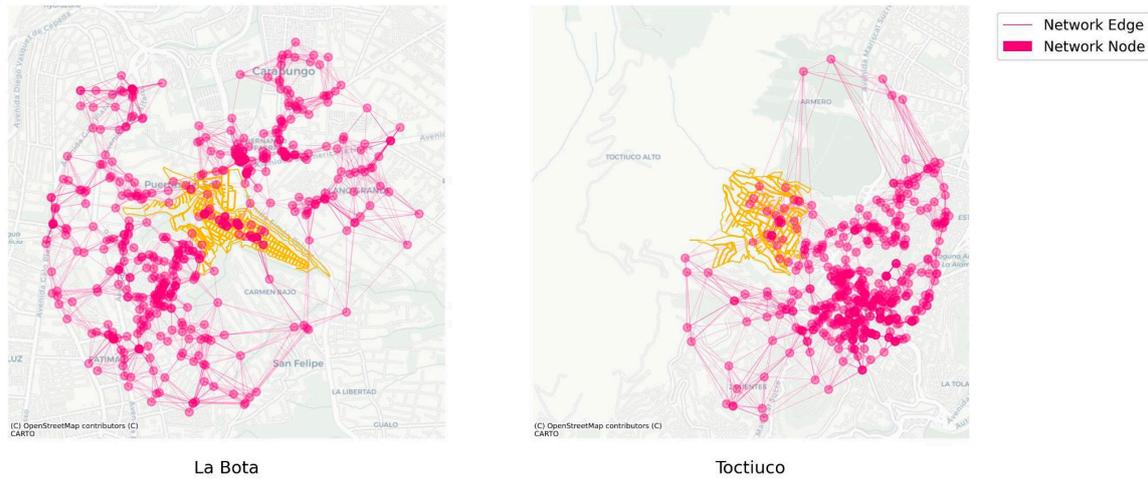
By leveraging the distance information, we identify the ten next closest POIs for each individual POI. The distances between each pair of POIs in each category subset are then calculated and the indices of the POIs obtained in ascending order based on their respective distances. This step returns an array of indices that functions to sort the distances in ascending order. For example, if the distances array is [5, 2, 10, 3, 1], the resulting indices array would be [4, 1, 3, 0, 2]. This means that the POI at index 4 is the closest to the POI sampled, followed by the POI at index 1, and so on.

Next, the first 10 indices (excluding the POI itself) from the sorted indices array are selected. Finally, the selected indices representing the 10 closest POIs for each individual POI in each category are stored in the "indices\_cercanos" matrix. This step establishes proximity-based connections among the health-related amenities.

Afterwards, we create an adjacency matrix — a mathematical representation of a network that shows the connections between nodes. In this case, the nodes represent POIs and this matrix is constructed by assigning a value of 1 to the corresponding positions in the matrix to indicate the presence of a connection between POIs. The matrix represents the relationships and connections between the nodes of the network in a binary fashion. After creating the adjacency matrix representing the connections between the POIs, we build a graph from that matrix composed of nodes (also known as vertices) and edges (also known as connections or links). In this case, each node of the graph represents a POI and each edge represents a connection between two POIs. The adjacency matrix that we have previously created provides the information needed to construct this network. This process brings information for a deeper network analysis, by examining key characteristics such as number of nodes, number of edges and the average degree of nodes.

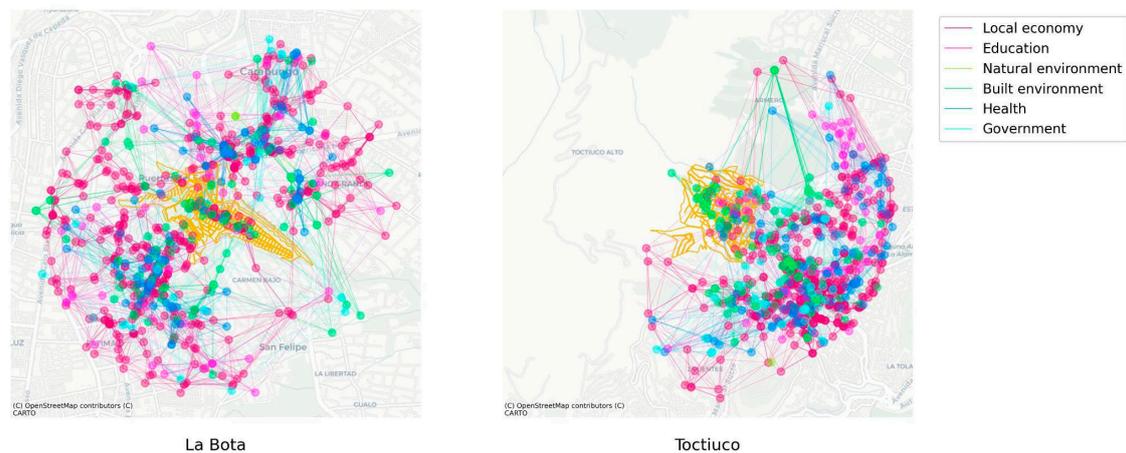
The network is then visualized using a graph visualization technique. The nodes (POIs) are plotted on a graph, with each node represented by a purple dot. The edges (connections) between nodes are depicted as curved lines, illustrating the relationships and proximity between the health-related amenities.

From this analysis, the examination of the health-related amenities network highlights areas that present a higher concentration of health services and the conditions of accessibility to these places. The network analysis lets us understand the availability, proximity, and accessibility of "local economy" amenities within the neighborhoods (see Figure 15). It is possible to see that the levels and concentration of these four parameters are high in the densest areas while in the periphery it is absent, excluding those communities from reaching these kinds of POIs. As a final analysis, the average degree of a node — the number of edges that connect it to other nodes — is a measure that indicates the average number of connections that each node has in a graph.



**Figure 15.** Local Economy Network.

The result of Average degree of local economy in Toctiuco, 12.56, indicates that each node in the neighborhood is connected to approximately 12 other nodes in the same category. This metric helps us to understand the connectivity and structure of the constructed graphs, revealing information about the density of connections in each network (Figure 16), the centrality of certain nodes, and the interaction between the different elements represented by the nodes.



**Figure 16.** All Network in La Bota and Toctiuco.

#### 4. Discussion

In this paper, we address the challenge of data acquisition in two peripheral neighborhoods of Quito, Ecuador: La Bota and Toctiuco. Both neighborhoods share similar physical and demographic conditions: peripheral, low-income, and inadequate and restricted open green areas. Despite having a large green area surrounding the built environment, their accessibility and quality are poor. We create models to make visible physical information where access to open data is restricted, limited in availability, and/or not up to date. This research proposes a methodology for the collection of open data and its preparation for future use in machine learning or spatial data science models.

We localize and analyze networks of UGS, public space, roads, and POIs as an urban infrastructure that gives information for reaching healthy city parameters. To this end, we developed a multi-source methodology to obtain as much information as possible from a combination of geographic data sources which are then correlated with on-the-ground fieldwork. Our methodology utilizes open spatial data sources, including H3 geo-indexing, OpenStreetMap (OSM), Microsoft Planetary Computer Building Footprint Layer, and official sources like census data, to propose a

comprehensive approach for identifying places that could serve as healthy areas for these types of neighborhoods.

Importantly, we conduct a quality assurance work, *in-situ*, to validate the accuracy, reality, and coherence of the Google Places API and OSM road data. During these visits, we engage with residents to gather firsthand insights on the neighborhood's reality. We observe the current conditions of accessibility, how people interact in public spaces, and economic activities developed in shops and sidewalks. These interactions enriched our understanding of the studied areas.

By integrating diverse datasets and employing cross-layer analysis and data visualization, we obtain valuable insights into patterns of amenity distribution, road infrastructure, and building density. Our analysis demonstrates the effectiveness of this methodology for highlighting the neighborhoods infrastructure on roads' network, public spaces, UGS, and POIs, and the potential deficiencies in essential services. Additionally, network analysis emerged as a promising tool for understanding the relationships and connectivity between amenities, offering new possibilities for urban planning and healthier city development.

This work presents a robust framework for studying peripheral and low-income neighborhoods in Latin American cities, within and beyond Quito, in relation to healthy cities criteria by leveraging open spatial data and incorporating on-site quality assurance visits. When considering the unique challenges faced in these contested urban contexts, we seek to develop more informed urban design strategies to address the specific needs of these communities, by improving the existent infrastructure and network of UGS, roads, POIs, and public space, promoting accessibility, well-being, and social equity. However, it is important to keep in mind that reality can change rapidly and affect the data provided. It should be noted that this analysis is based on information collected at the time of the study and it is possible that current data and conditions may differ due to possible changes and updates.

This research establishes a methodology that can be further developed in several ways related to advanced models for Urban Studies in low-income and peripheral neighborhoods of Global South cities. Utilizing the comprehensive datasets from this study and spatial analysis with machine learning techniques, this methodology may provide deeper insights on the physical urban infrastructure of public space. This research can generate interactive visualizations could democratize data access, and an open data platform could broaden the impact of this work by enabling independent analyses and applications. Lastly, comparative studies could be conducted to identify factors contributing to urban health across different cities or regions, providing valuable insights for urban planning, design, and management.

**Author Contributions:** The authors developed the different contributions: Ana Medina. Conceptualization; methodology; investigation; resources; visualization; writing – original draft preparation; writing – review and editing; supervision; project administration; funding acquisition. Diana Mosquera. Methodology, investigation; software, validation, formal analysis, data curation, writing – original draft preparation; visualization. Francisco Gallegos. Methodology, investigation; software, validation, formal analysis, data curation, writing – original draft preparation; visualization.

**Funding:** This research was funded by Universidad de Las Americas, Quito, Ecuador, grant number ARQ.AMG.20.02.

**Data Availability Statement:** The repository regarding data and scripts that supports reported results, can be found in the following link to publicly archived datasets analyzed and generated during this study: [https://github.com/DiversaStudio/Peripheral\\_Neighborhoods\\_Methodology](https://github.com/DiversaStudio/Peripheral_Neighborhoods_Methodology)

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. S. E. Bibri, "Data-Driven Smart Eco-Cities of the Future: An Empirically Informed Integrated Model for Strategic Sustainable Urban Development," <https://doi.org/10.1080/02604027.2021.1969877>, 2021, doi: 10.1080/02604027.2021.1969877.
2. A. H. Hassan, R. bin Sulaiman, M. A. Abdulgaber, and H. Kahtan, "Balancing technological advances with user needs: User-centered principles for AI-driven smart city healthcare monitoring," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 3, 2023.
3. T. H. Son, Z. Weedon, T. Yigitcanlar, T. Sanchez, J. M. Corchado, and R. Mehmood, "Algorithmic urban planning for smart and sustainable development: Systematic review of the literature," *Sustain Cities Soc*, vol. 94, p. 104562, Jul. 2023, doi: 10.1016/J.SCS.2023.104562.
4. E. P. Barboza *et al.*, "Green space and mortality in European cities: a health impact assessment study," *Lancet Planet Health*, vol. 5, no. 10, pp. e718–e730, Oct. 2021, doi: 10.1016/S2542-5196(21)00229-1.
5. M. J. Nieuwenhuijsen, H. Khreis, M. Triguero-Mas, M. Gascon, and P. Dadvand, "Fifty Shades of Green: Pathway to Healthy Urban Living," *Epidemiology*, vol. 28, no. 1, 2017, [Online]. Available: [https://journals.lww.com/epidem/Fulltext/2017/01000/Fifty\\_Shades\\_of\\_Green\\_Pathway\\_to\\_Healthy\\_Urban.11.aspx](https://journals.lww.com/epidem/Fulltext/2017/01000/Fifty_Shades_of_Green_Pathway_to_Healthy_Urban.11.aspx)
6. Y. Ju *et al.*, "Latin American cities with higher socioeconomic status are greening from a lower baseline: evidence from the SALURBAL project," *Environmental Research Letters*, vol. 16, no. 10, p. 104052, Oct. 2021, doi: 10.1088/1748-9326/ac2a63.
7. J. Jacobs *et al.*, "Variation in the physical activity environment according to area-level socio-economic position—A systematic review," *Obesity Reviews*, vol. 20, no. 5, pp. 686–700, May 2019, doi: 10.1111/obr.12818.
8. A. Rigolon, M. Browning, K. Lee, and S. Shin, "Access to Urban Green Space in Cities of the Global South: A Systematic Literature Review," *Urban Science*, vol. 2, no. 3, p. 67, Aug. 2018, doi: 10.3390/urbansci2030067.
9. L. D. Acosta, "Capacidad de respuesta frente a la pandemia de COVID-19 en América Latina y el Caribe," *Revista Panamericana de Salud Pública*, vol. 44, p. 1, Sep. 2020, doi: 10.26633/RPSP.2020.109.
10. R. Ahrend, B. Cournède, P. Veneri, and V. Ziemann, "Has COVID-19 Triggered an Urban Exodus?," *OCDE Forum Network*, May 12, 2022. <https://www.oecd-forum.org/posts/more-room-with-a-view-has-covid-19-triggered-an-urban-exodus> (accessed Feb. 09, 2023).
11. T. Burki, "COVID-19 in Latin America," *Lancet Infect Dis*, vol. 20, no. 5, pp. 547–548, May 2020, doi: 10.1016/S1473-3099(20)30303-0.
12. J. Corburn *et al.*, "Slum Health: Arresting COVID-19 and Improving Well-Being in Urban Informal Settlements," *Journal of Urban Health*, vol. 97, no. 3, pp. 348–357, 2020, doi: 10.1007/s11524-020-00438-6.
13. World Health Organization, "Healthy Cities Effective Approach to a Changing World," May 2020. Accessed: Mar. 05, 2023. [Online]. Available: <https://www.who.int/publications/i/item/9789240004825>
14. INEC, "Encuesta Nacional de Empleo, Desempleo y Subempleo (ENEMDU), enero 2022," Feb. 2022. Accessed: Mar. 01, 2022. [Online]. Available: [https://www.ecuadorencifras.gob.ec/documentos/web-inec/EMPLEO/2022/Enero-2022/202201\\_Boletin\\_empleo.pdf](https://www.ecuadorencifras.gob.ec/documentos/web-inec/EMPLEO/2022/Enero-2022/202201_Boletin_empleo.pdf)
15. F. Carrión and J. Erazo Espinosa, "La forma urbana de Quito: una historia de centros y periferias," <http://journals.openedition.org/bifea>, no. 41 (3), pp. 503–522, Dec. 2012, doi: 10.4000/BIFEA.361.
16. W. Herdoiza, "Hacia un Urbanismo Popular," Quito, 1976.
17. P. Cubillo, D. Orellana, and P. Maigua, "Análisis De La Distribución Espacial De La Temperatura Superficial Y La Valoración De Criterios Que Influyen En La Isla De Calor Urbano (Icu) En El Distrito Metropolitano De Quito," Quito, 2020. Accessed: Mar. 28, 2023. [Online]. Available: <https://www.ciuq.ec/INFORME%20FINAL%20ISLAS%20DE%20CALOR.pdf>
18. INEC, "Proyecciones Poblacionales Ecuador," *Proyecciones Poblacionales*, 2022.
19. E. Peralta and R. Moya Tasquer, *Quito Patrimonio Cultural de la Humanidad*. Quito: Trama, 2007.
20. M. E. Acosta, "Políticas de vivienda en Ecuador desde la década de los 70: análisis, balance y aprendizajes," FLACSO Sede Ecuador, Quito, 2009. Accessed: Mar. 31, 2022. [Online]. Available: <https://repositorio.flacsoandes.edu.ec/handle/10469/892>
21. F. Carrión, "Existen los barrios periféricos en Quito?," *Trama*, vol. 24, 1981.
22. P. Mooney and M. Minghini, "A review of OpenStreetMap data," *Mapping and the citizen sensor*, pp. 37–59, 2017.
23. S. Low, *Why public space matters*. Oxford University Press, 2022. doi: 10.1093/oso/9780197543733.001.0001.
24. L. Kurgan, *Close up at a distance: Mapping, technology, and politics*. MIT Press, 2013.
25. Y. Xu *et al.*, "A framework for urban land use classification by integrating the spatial context of points of interest and graph convolutional neural network method," *Comput Environ Urban Syst*, vol. 95, p. 101807, Jul. 2022, doi: 10.1016/J.COMPENVURBSYS.2022.101807.
26. C. Huang, C. Xiao, and L. Rong, "Integrating Point-of-Interest Density and Spatial Heterogeneity to Identify Urban Functional Areas," *Remote Sensing 2022, Vol. 14, Page 4201*, vol. 14, no. 17, p. 4201, Aug. 2022, doi: 10.3390/RS14174201.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.