Master Thesis in Geographical Information Science nr 101

# Exploring the distribution of accessibility by public transport using spatial analysis.

A case study for retail concentrations and public hospitals in Athens.

**Evangelos Vafeiadis**

2019

Department of
Physical Geography and Ecosystem Science
Centre for Geographical Information Systems
Lund University
Sölvegatan 12
S-223 62 Lund

Evangelos Vafeiadis (2019). Exploring the distribution of accessibility by public transport using spatial analysis. A case study for retail concentrations and public hospitals in Athens.

Master degree thesis, 30/ credits in Master in Geographical Information Science

Department of Physical Geography and Ecosystem Science, Lund University

# Acknowledgements

I would like to thank Finn Hedefalk for the invaluable comments and guidance throughout the whole thesis. I would also like to thank Athens' Urban Transport Organization (OASA), the Hellenic Confederacy of Commerce and Entrepreneurship (ESEE) and the Ministry of Health for the kindly provision of datasets.

# Abstract

This thesis studies the distribution of accessibility by public transport to retail concentrations and public hospitals in Athens, using spatial autocorrelation analysis based on Local Indicators of Spatial Autocorrelation (LISA) method and Geographically Weighted Regression (GWR). For this reason, the study area was divided in a 300x300 meter square grid of points representing the location points of the study. For each point, the travel time to each of the retail concentration and public hospital location was estimated using Google Directions API. The accessibility of each point was then calculated as the percentage of reachable destinations for each land use group within 45 minutes. The results where then aggregated to zip code level, where along with data regarding the average annual zip code income, population density and destination from Athens' central business district formed the final dataset. The analysis that followed suggests that there is a cluster of high accessibility in the city center and a cluster of low accessibility in the outer suburbs. In addition, a middle zone of insignificant clustering is located between the two clusters. The chosen urban structure and socioeconomic variables that were used for the geographically weighted regression have significant effects except for the case of population density on public transport accessibility. More specifically, distance from the central business district (CBD) is found to be negatively correlated with accessibility to both retail concentrations and hospitals. On the other hand, annual average income seems to be positively correlated with accessibility to both destinations. Finally, population density has a positive correlation with only retail concentrations. In addition, the analysis indicated that there may be more unknown factors affecting accessibility to retail concentrations in the city center than the periphery.

# TABLE OF CONTENTS

## LIST OF FIGURES



## LIST OF TABLES

# 1 Introduction

Cities keep expanding and developing constantly. This expansion is coupled with the increase in human activity and social processes. Each day more people need to travel from home to other destinations within the city limits in order to reach their working places, meet with other people, do their daily shopping, go to a cinema, visit a hospital or relax in a park. In addition, because city expansion is more often horizontal than vertical, the distance between existing land uses and new parts of the city keeps increasing. As a result, this causes the establishment of new land uses closer to the new development location and away from the city center.

The land use destinations, whether already established or new, can be reached by the transport system. Where land uses can be conceptualized as areas in space, the transport system can be conceptualized as links and nodes. The links represents the space which enables movement between its nodes and they may be paths, residential roads, highways, rail tracks etc. The nodes represent the places of entrance or exit to the system and the more closely they are located to land use destinations the easier it is for the latter to be reached.

Since there are several transport modes, the availability of the whole system cannot be guaranteed for all of them individually. It is impossible to reach a certain cinema by rail if there is no rail station nearby, as it is almost impossible to reach a business park which is located in the junction of two highways on foot. For this reason the majority of the travels consist of a combination of the available transport modes with the most frequent element being walking as it provides access to both the car and public transportation. However, private motorized modes (automobile, motorbike etc) pose a problem as elements of a combined travel since they may not constitute an available option for individuals due to a variety of reasons (age, driving capabilities, cost etc). In this case travels are restricted to the choice of walking and public transportation.

Depending on the transportation system and whether it is fully available to an individual as well as the travel distance or travel time between the destination and the location, a destination can be reached from a specific location by a varying degree of ease. This degree of ease is called accessibility. Accessibility is defined by Geurs, Karst T. & van Wee (2004) as “the extent to which land-use and transport systems enable (groups of) individuals to reach activities or destinations by means of a (combination of) transport mode(s)." It follows that accessibility is a multi-dimensional concept, involving the land-use pattern, the transport network as well as the individuals. Its importance lies in the fact that accessibility “enable[s] (groups of) individuals to reach activities or destinations” (Geurs, Karst T. & van Wee (2004)), as activity participation is a fundamental function for the reproduction of societies. As accessibility increases so does the number of choices and alternatives that are available to a person.

Accessibility can be greatly impacted by the strategies and practices of land use and transport authorities. Both these aspects of urban structure can cause severe uneven distribution in accessibility. For instance, the construction of a metro station in a particular area can greatly increase its accessibility. At the same time, if there is inadequate transport connection of the surrounding areas to the metro station, their accessibility is significantly decreased in relation to the area with the metro station. This uneven distribution may cause social inequalities, such as social exclusion as the people do not have the same access to activities and/or displacement as a rapid increase in accessibility of an area may lead to increased rent.

As accessibility is a vast concept, this thesis focuses on just one transport system, that of public transportation, and two land-uses: retail markets and public hospitals. Public transportation was chosen because it is one of the major transport modes that people use to access and participate in various activities throughout the city. It is also the only mode that offers an alternative to car use for medium to long travel distances and it is widely inclusive for all social groups. In contrast to car use, which requires the users to be of certain age, possess certain abilities, hold a driving license and own a car, public transportation is addressed to everyone.

For the case of Athens, accessibility by public transport is of particular importance as during the last years over one million people chose to give up their car's plate in order to avoid the annual car use tax (Andris, 2016). It can be assumed that the majority of those people were driven in this decision by economic difficulties caused by the ongoing economic crisis. Thus, dependence on public transport did only increase in general, but it did so for a significant portion of the economic vulnerable population.

Finally, this thesis focuses on the impact of three variables in order to explain accessibility. The two of them are urban structure variables and namely, distance from the central business district (CBD) and population density. The third is a socioeconomic variable and consists of the annual average income. Distance from CBD impacts accessibility as Athens' Transport Network is to a large extend radial based and center focused. Most of the public transport routes pass through or close to the center of Athens. This creates the phenomenon that there is a wide variety of transport routes and stops as the distance to CBD decreases. In contrast, as the distance from CBD increases and public transport routes branch off there are less available options and connections for remote locations. Population density impacts accessibility because public transport operation is more feasible for locations with high population density. This is because a certain stop or line is expected to serve more people than it would in a low density location. The last considered variable, annual average income, affects public transport accessibility because from a planning perspective, locations with lower income are expected to be better served by public transport, since the option of the car may not be available. At the same time good public transport provision in a given location may result in increased rents so that lower income population cannot afford to live in such locations.

## 1.1 Research Questions

This thesis aims to answer the following questions:

1. How is public transport accessibility to a) retail concentrations and b) public hospitals distributed in Athens Metropolitan Area? Is it randomly distributed and/or are there clusters of high and/or low accessibility?
2. Is it possible to explain the spatial pattern of public transport accessibility to a) retail concentrations and b) public hospitals by the distance from the city center, population density and average income?

As far as the first question is concerned, the Null Hypothesis is:

- $H_{0\text{-}A1}$: Accessibility to retail concentrations is randomly distributed
- $H_{0\text{-}A2}$: Accessibility to retail concentrations is not significantly locally clustered.
- $H_{0\text{-}B}$: Accessibility to public hospitals is randomly distributed
- $H_{0\text{-}B2}$: Accessibility to public hospitals is not significantly locally clustered.

As far as the second question is concerned, the Null hypothesis can be broken down to three hypotheses, each for one of the three predicting variables:

- $H_{0\text{-}A1}$: In the presence of population density and average income, there will be no significant correlation between retail accessibility and distance from CBD.
- $H_{0\text{-}A2}$: In the presence of distance from CBD and average income, there will be no significant correlation between retail accessibility and population density.
- $H_{0\text{-}A3}$: In the presence of distance from CBD and population density, there will be no significant correlation between retail accessibility and average income.

And for public hospitals:

- $H_{0\text{-}B1}$: In the presence of population density and average income, there will be no significant correlation between public hospital accessibility and distance from CBD.
- $H_{0\text{-}B2}$: In the presence of distance from CBD and average income, there will be no significant correlation between public hospital accessibility and population density.
- $H_{0\text{-}B3}$: In the presence of distance from CBD and population density, there will be no significant correlation between public hospital accessibility and average income.

## 1.2 Thesis Outline

The thesis is structured in six additional sections. Section 2 contains the fundamental background information that is necessary in order to ground the research theoretically. Section 3 describes the study area as well as the data used for the analysis. In section 4 the main methods of the analysis are presented while section 5 contains the results of the analysis. Finally, in section 6 there is a discussion of the results which are put in a broader context and, section 7 contains some concluding remarks.

# 2 Background

This chapter presents the main theoretical background on which the thesis is built. Firstly, the concept of accessibility is presented and defined. Secondly, the public transport accessibility is being brought into focus. The third section describes the relation between accessibility and social equity. Finally, the analytical methods that were used in order to answer the research questions are described.

## 2.1 Defining Accessibility

Accessibility is a concept that has been used widely for social sciences. The first concise introduction of accessibility for land use research was introduced by Hansen (1959) in which he studies the effects of accessibility on land use patterns. A decade later, the work of Hägerstrand (1970) paved the way for the inclusion of agency in accessibility research. Geurs, Karst T. & van Wee (2004) define accessibility as "the extent to which land-use and transport systems enable (groups of) individuals to reach activities or destinations by means of a (combination of) transport mode(s)". The definition makes accessibility relevant for fields like urban and transport planning, economic geography, transport geography, market analysis, social exclusion, real estate and other disciplines that are not only concerned with the proximity of two location, but with their ease of access.

However, despite its wide use, accessibility still remains an elusive concept, with many researchers offering their own view of its definition and composition (Geurs, De Montis & Reggiani (2015)). For instance, Geurs, Karst T. & van Wee (2004) after an extensive review of accessibility studies, suggest that accessibility is composed of four elements: the land-use component (the spatial pattern and demand of land-uses), the transportation component (the transport network), the temporal component (temporal constraints such as rush hour or public transport operation hours) and the individual component (the needs and preferences of individuals). Ideally, each one has to be taken into consideration for accessibility evaluation. In another study Cass, Shove, & Urry (2005) suggest that accessibility has four dimensions: the financial (the monetary cost of travel), the physical (the physical ability to travel), the organizational (how systems and networks are organized) and the temporal (again, temporal constraints like rush hour or public transport schedule). As will be discussed later (section 2.2), this study uses the approach of Geurs, Karst T. & van Wee (2004).

Some other authors are concerned directly with how accessibility is measured rather than what are its main components. Although more specific, these accessibility metrics still wield information about how it is conceptualized. For example, Lei & Church (2010) categorize accessibility into six metrics. The first is *system accessibility*, which is concerned with the physical access to any transport system. The second is *system facilitated accessibility* and is concerned with the ease that one can reach a particular destination through the transport system. The third metric, *integral accessibility* is a general accessibility indicator of an area, taking into consideration multiple destinations and networks. The fourth metric is the time-space accessibility of the user as defined in Hägerstrand (1970). The fifth metric derives from utility theory, and assumes that agents are fully aware of alternatives as well as they are rational thinkers. Finally, *relative accessibility* is the difference in accessibility caused by different transport modes.

Moreover, although there are conceptual differences in accessibility, most models take into consideration in one way or another similar factors, such as travel time, travel cost, physical proximity, temporal aspects and user preferences/constraints. However as Bert van Wee (2016) highlights, there are still questions that researchers have to address in order to advance and fully mature the field.

## 2.2 Accessibility in the thesis

The accessibility indicator which is used in this thesis follows the approach of Geurs, Karst T. & van Wee (2004). However, it does not consider all four accessibility components, but only the land use and transportation system. The accessibility index is based on the percentage of destinations reachable within 45 minutes from a location (further described in section 4.1). The form was chosen in order to better represent the number of alternatives that are available to someone who wishes to carry out the activity in question. This was preferred over the minimum travel time destination, as the latter may be misleading in cases where the destination is very close but there are no other real alternatives available. Secondly, the temporal threshold was chosen after reviewing acceptable travel times (Milakis, Cervero, van Wee & Maat (2015), Mokhtarian & Salomon, (2001)). The first study focuses on commuting preferences and it concludes that the ideal travel time is 18.4 minutes while the acceptable travel time is 42.8 minutes. However, public transport users have reported the highest acceptable commute time (60 minutes). Similarly, for the second study, the most agreed ideal travel time for commute lies between 15-19 minutes with no respondent stating an ideal travel time higher than 50 minutes. However, the actual travel time is much longer for most of the sample.

Thus, the threshold of 45 minutes of this study was chosen in the basis that although it may not be an ideal travel time, this threshold seems to be acceptable which signifies that people would still consider it as an extreme case. In addition, the reported travel times of both studies were for commuting, which is different from the activities studied in this thesis. It is assumed that since shopping or visiting a hospital is not as frequent as commuting, ideal and acceptable travel times may be a bit longer than the ones reported.

## 2.3 Public Transport Accessibility

Public transport accessibility focuses on the role of the public transport network as an accessibility provider. Public transport may be the only alternative to car use for medium to long urban distances therefore its accessibility evaluation is important for environmental, equity and even economic reasons. Research on this field is conducted with multiple definitions of accessibility, methods as well as objectives, such as access to different land uses.

Albacete et al. (2017), following the accessibility typology suggested by Geurs, Karst T. & van Wee (2004),use two individual methods: Structure Accessibility Layer (SAL) and Public Transport Walking Access Index (PTWAI), to measure the overall accessibility of Helsinki. The first is origin based, meaning that it is concerned with the ease that facilities can be reached from the origin, whereas the second is destination-based and is concerned with the ease that this particular destination is accessible. Both methods do not consider individual preferences/constraints and focus on the land-use and transportation components. In addition PTWAI considers partially the temporal component.

In another study, Saghapour, Moridpour, & Thompson (2016) use public transport accessibility index (PTAI) to assess the accessibility levels of Melbourne. The model is destination-based and besides taking into consideration the walking time to the nearest stop and the frequency of the services, it also considers the population density in the origin points.

An interesting approach is taken in Fransen et al. (2015) who try to indentify the public transport gaps in the area of Flanders in Belgium. The gaps are identified as the difference between two indices, the index of public transportation needs (IPTN) and the index of public transport provision (IPTP).

As a conclusion, there is a wide selection of methodological concepts and measurement methods for accessibility. The choice of a particular method seems to depend on the requirements of each study. However, the common thread that connects all of them is the requirement of a land use element in order to assess the accessibility towards that land use and a measure of friction which is usually captured by the transport network in travel time from the location of the measurement to the desired land use.

## 2.4 Accessibility and social equity

Accessibility enables groups of people and individuals to reach and participate in activities. It is therefore a cornerstone for social equity since differences in accessibility result in differences in opportunities and social interaction (see Grengs (2012)). Researchers have neglected the spatial dimension of social equity/exclusion in the past, but during the past decade, new research tries to bring it in focus (Cass et al. (2005)). However, the concept of equity is diverse. As Wee & Geurs (2011) suggest, there are differences in how equity is understood depending on what ethical theory one follows. Examining three ethical theories (utilitarianism, egalitarianism and sufficientarianism) they conclude that utilitarianism, expressed through the Cost-Benefit Analysis Tool for accessibility evaluation may be the least desirable to use.

When examining equity of transport systems, Neutens et al. (2010) note that the more a model is geared towards individual preferences, "the more conservative it will be in terms of signifying a state of equity"'. This is particularly relevant with the discussion for residential location "self-selection". By self-selecting, people choose their residential location based on their transport preferences and therefore to some extent the area's accessibility levels. Taking personal preferences/constraints into consideration may obfuscate the actual levels of accessibility because of two main reasons. The first is that affluent individuals prefer to live in areas with low accessibility by public transport and the second because people have a subjective view of accessibility (c.f.Cheng & Chen (2015) and Lättman, Friman, & Olsson(2016)).

## 2.5 Spatial Autocorrelation

Spatial autocorrelation measures the degree to which neighboring objects share the same qualities. In other terms, it helps us to determine whether a particular phenomenon is clustered, dispersed or random in the geographical space. In our project, spatial autocorrelation indicates if there are any clusters of high or low accessibility or if accessibility is random across Athens.

One of the most widely used indicators of spatial autocorrelation is Moran's I statistic, which has two versions: the global and the local. In its global version, the statistic yields an overall (global) estimation about the distribution of a phenomenon. This practically means that there is only one value ranging between -1 and 1, with the former indicating complete negative spatial autocorrelation and the latter positive spatial. The value of 0 indicates that there is no autocorrelation; the phenomenon is randomly distributed in geographical space. The global Moran's I statistic formula according to Rogerson (2001, pp: 167) is:

$$I = \frac{n \sum_i^n \sum_j^n w_{ij} (y_i - \overline{y})(y_j - \overline{y})}{\left(\sum_i^n \sum_j^n w_{ij}\right) \sum_i^n (y_i - \overline{y})^2} \quad [1]$$

Where *n* is the number of spatial entities, $w_{ij}$ is a weight of spatial proximity between entities *i* and *j*, $y_i$ is the measured value of the variable in question at spatial entity *y* and $\overline{y}$ the mean value.

On the other hand, local statistics such as local Moran's I statistic, are more sensitive to local clustering. Local statistics may pick up pockets of spatial autocorrelation that would have passed under the global statistic's radar. The formula of the local Moran's I statistic according to Rogerson (2001, pp: 167) is:

$$I_i = n(y_i - \bar{y}) \sum_{i \neq j} w_{ij} (y_j - \bar{y}) \qquad [2]$$

### 2.5.1 Local Indicators of Spatial Association

Local Indicators of Spatial Association (LISA) is a technique proposed in Anselin (1995), which:

- for every observation "gives an indication of the extent of significant spatial clustering of similar values around that observation" and
- "the sum of LISAs for all observations is proportional to a global indicator of spatial associations"

The first aspect of LISA is the identification of local clustering, in what is often referred as hot spots. As Anselin (1995), notes "Similar to the rationale behind the significance tests for the Gi and Gi* statistics of Getis and Ord (1992), the general LISA can be used as the basis for a test on the null hypothesis of no local spatial association. However, in contrast to what holds for the Gi and Gi* statistics, general results on the distribution of a generic LISA may be hard to obtain".

The second aspect of LISA allows the link of local indicators to a "global measure of spatial association, […], thus enabling the assessment of influential observations and outliers."

Finally, LISA can be constructed from a variety of spatial autocorrelation statistics, one of them being Moran's I.

## 2.6 Multiple Linear Regression

Multiple Linear Regression (MLR) is used for modeling the effects of one or more variables on another. The former are called independent variables and the latter dependent variable. This is because the values of the dependent variable are influenced by, or depend on the values of the independent ones.

MLR as the name suggests, estimates a linear relationship between the dependent and independent variables. In order to estimate the line, it uses Ordinary Least Squares method, which is a curve fitting method based on the principle that the sum of the squared errors of the independent variables to the fitted line should be the minimum possible. The general formula is:

$$\hat{y} = a + b_1x_1 + b_2x_2 + b_3x_3 + \ldots.. b_nx_n \qquad [3]$$

Where $\hat{y}$ is the predicted value, a represents a constant and each set of $b_1x_1$ represent the coefficient (b) and the value $(x)$ of each independent variable. MLR will be used in the thesis as a first overall indicator for the global tendencies of the sample. However,The main analysis and conclusions will be drawn by applying Geographic Weighted Regression.

## 2.7 Geographically Weighted Regression

Geographically Weighted Regression (GWR) is a special form of spatial regression put forth by Fotheringham et al. (1998), where not only the residuals are a function of neighboring residuals but also the regression function has a local character rather than a global. This is achieved by assigning weights to all other observations. These weights are a function of distance and their value decreases as the distance increases. So, the dependent variable at location *i* is modeled as:

$$y_i = b_{i0} + \sum_{j=1}^{p} b_{ij} x_{ij} + \varepsilon_i \qquad [4]$$

Where *p* indicates the number of independent variables, $b_{ij}$and the$x_{ij}$ represent the coefficient and value for variable *j* and location *i* respectively (note that $b_{i0}$ represents the constant for location *i*) and $\varepsilon_i$ represents the random error for the dependent variable in location *i*. Since accessibility is a highly spatially correlated phenomenon, GWR along with spatial autocorrelation are used in order to tackle the research questions. GWR was chosen because of its sensitivity to local variability rather than global as it is the case with MLR.

# 3 Study Area and Data

## 3.1 Study Area

The study area is the current Urban Transport Service Area which includes most of the Athens Metropolitan area (Figure 1). The main responsible organization for it is Athens Urban Transport Organization (OASA). The metropolitan area consists of 84 settlements, which are organized in 52 municipalities. The transport network consists of 3 metro lines, 3 tram lines, 2 rail lines and 243 bus lines and the total length of each subsequent public transport network as well as the number of stops can be seen in the Table 1 (Table 1).

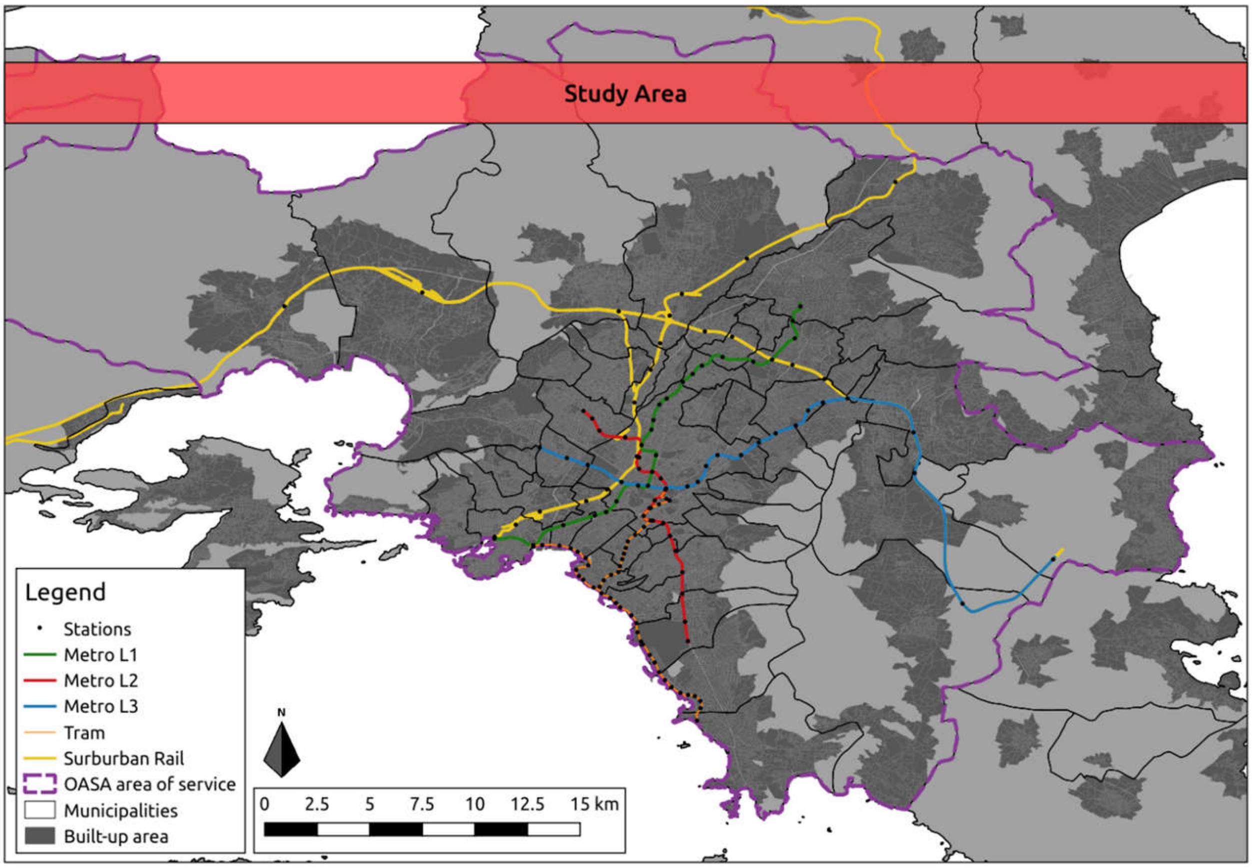


*Figure 1: Location of Athens in Greece*

*Table 1: Length and stops by mode for Athens' Public Transport Network*

| | Metro | Tram | Rail | Bus | total |
|---|---|---|---|---|---|
| Total Length (km) | 166 | 104 | 190 | 6011 | 6471 |
| Number of Stops | 61 | 49 | 26 | 7787 | 7923 |

## 3.2 Data

This study used data from multiple sources in order to study the research questions. All data where aggregated to zip code spatial level. The final dataset that was used for analysis consisted of the zip codes and their geometry, retail and public hospital access scores for each zip code as well as the distance from the city center, the population density and average annual income of each zip code (Table 2).

*Table 2: Final dataset of the study*

| Field | Units | Type | Description |
|---|---|---|---|
| zip_code | - | string | Primary key of the table |
| Geom | - | Geometry(Point, 2100) | The geometry of the point in Greek Grid |
| Retail access | Percentage value (%) | Double precision | The accessibility score of the zip code |
| Hospital Access | Percentage value (%) | Double precision | The accessibility score of the zip code |
| Average income | € (euro) | Double precision | Average income of the zip code |
| Population density | population/Ha | Double precision | Residential density of the zip code |
| CBD distance | m (meters) | Double precision | Euclidean distance from Athens' Central Business District (also, the city's center) |

## 3.3 Data processing

This section describes the processing of the acquired data in order to reach their final state as described in the previous section (3.2). The dataset presented in Table 3 is a combination of several data sources and processes that are further described in this section. All collected data were imported in a PostGIS database, a spatial extension of PostgreSQL relational database management system. The acquired data include:

- Transport Direction data
- Population data
- Income data
- Retail location data
- Hospital location data

*Table 3: Dataset Characteristics*

| Dataset | Format | Spatial Level | CRS | Provider |
|---|---|---|---|---|
| Transport Direction Data | JSON | - | WGS'84 | Google |
| Population | spreadsheet | zip code | | ELSTAT |
| Income | text | zip code | | GSIS |
| Retail business locations | shapefile | | GGRS'87 | ESEE |
| Hospital locations | spreadsheet | | | Ministry of Health |

The overall processing of the data can be seen in Figure 2. The key parts of the data consist of the Origins, the Destinations, Google Maps Directions API and the Zip codes. The origins and the destinations were used as input to Google Maps Directions API in order to calculate the travel time needed from each Origin point to each Destination. With the calculated travel times between every Origin-Destination pair, the retail and hospital access of each Origin point was calculated as the percentage of reachable destinations within 45’ to the total number of destinations. Next, the accessibility values of each Origin were aggregated to their respective zip codes in order to calculate the mean accessibility of the zip code. This was done in order to bring all relevant data the same spatial level as well as to have a more realistic accessibility value for each zip code than having to rely on its centroid accessibility value alone. Finally, as far as the three remaining variables are concerned, population density was calculated using the population of the zip code. The calculation of distance from CBD is a straightforward Euclidean distance and the average annual income was provided as is.

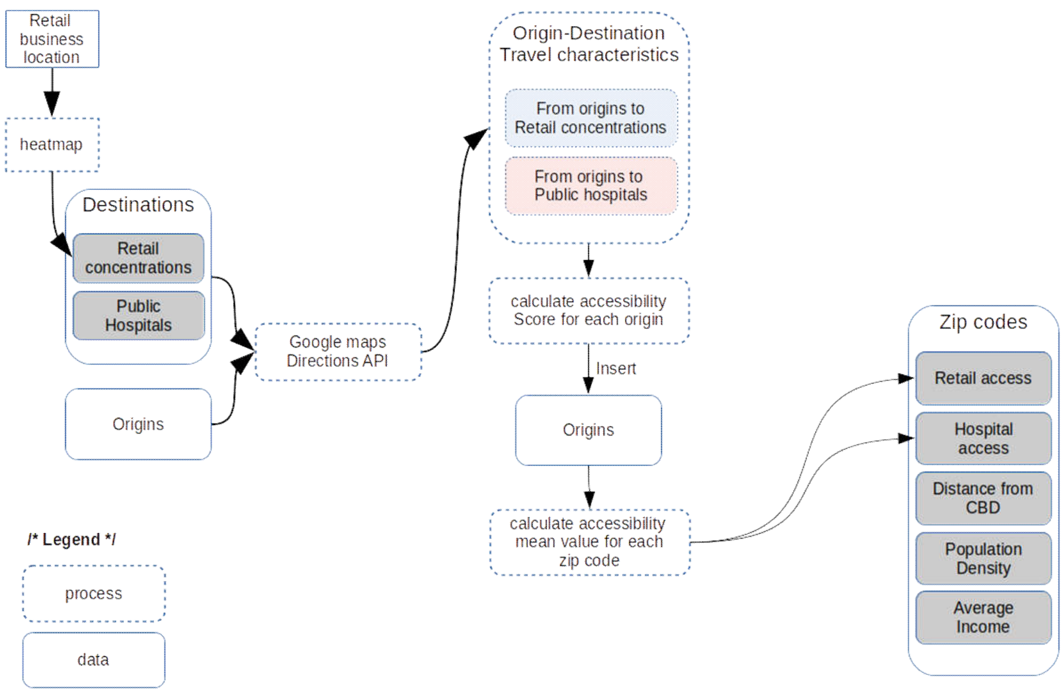


*Figure 2: Overall data processing workflow*

### 3.3.1 Construction of Origin points

More specifically, the Origin points (Figure ) represent the locations whose accessibility was assessed. They were constructed by creating a point grid of 300 meters inside Athens' Urban Transport Organization (OASA) area of service. Because OASA's area of service is described by administrative boundaries, points that fell outside of built up area and to some extend inside special urban land uses such as airports/parks were removed. The final product was a grid consisting of 6526 points. The 300m step is considered a good compromise between spatial detail and calculation efficiency.

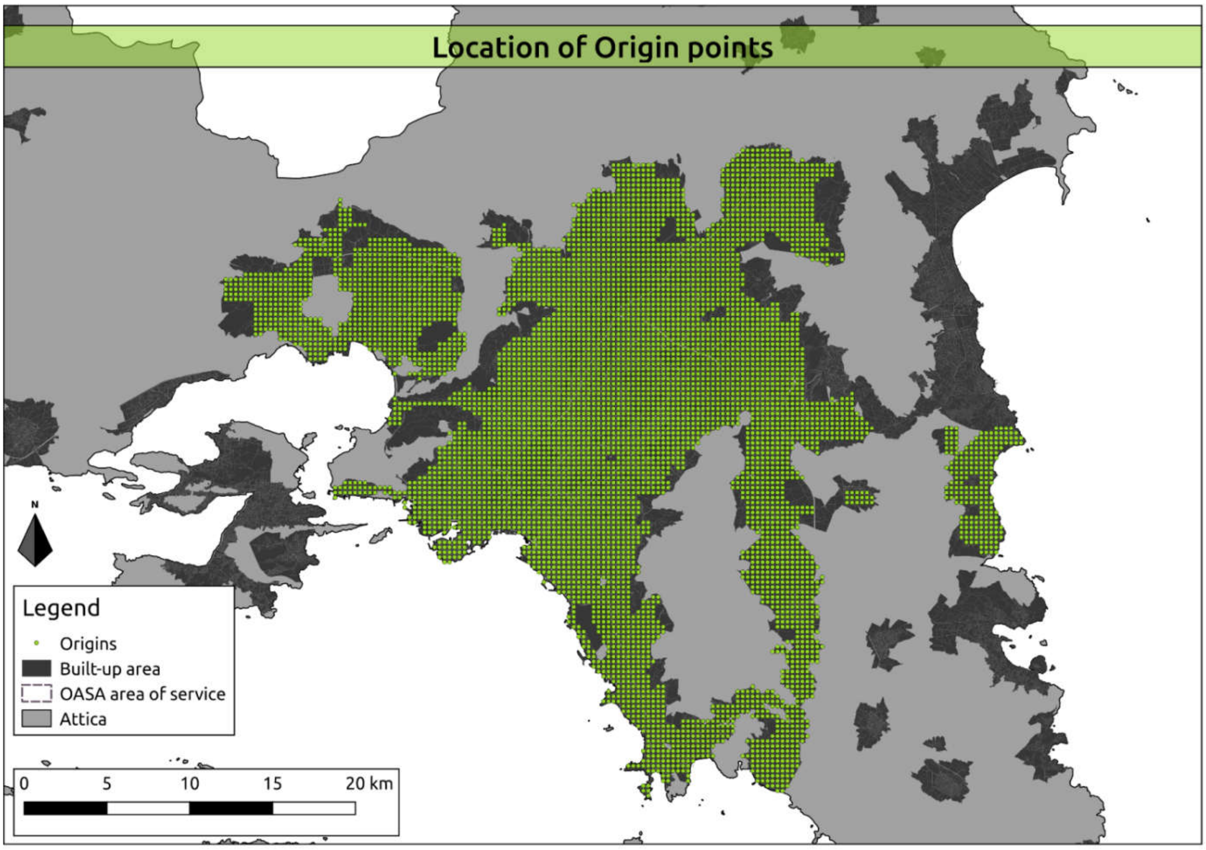


*Figure 3: The location of origin points.*

### 3.3.2 Construction of Destination points

The destinations represent the access destinations of the study. They consist of retail concentrations and public hospitals. The retail dataset is composed of 21 urban retail concentrations and shopping malls (Figure ). The shopping malls' locations were geocoded using their addresses. In cases where more than one shopping mall was in close proximity, a single point was geocoded. The urban retail concentrations derived after creating a heatmap (Figure 5), which shows the concentration of retail businesses in number. The map was created by applying a kernel density function over the business locations. The kernel itself is a Gaussian kernel while the search radius was set to 150m. The datataset for the heatmap was provided by ESEE. The dataset is a geocoded version of the data that is being updated by G.E.MI (General Commercial Registry) and includes all retail business. The hospitals dataset is compiled by a list of 30 public hospitals provided by the Ministry of Health. The dataset was processed so that hospitals that are in close proximity are represented with only one Destination point (Figure ).

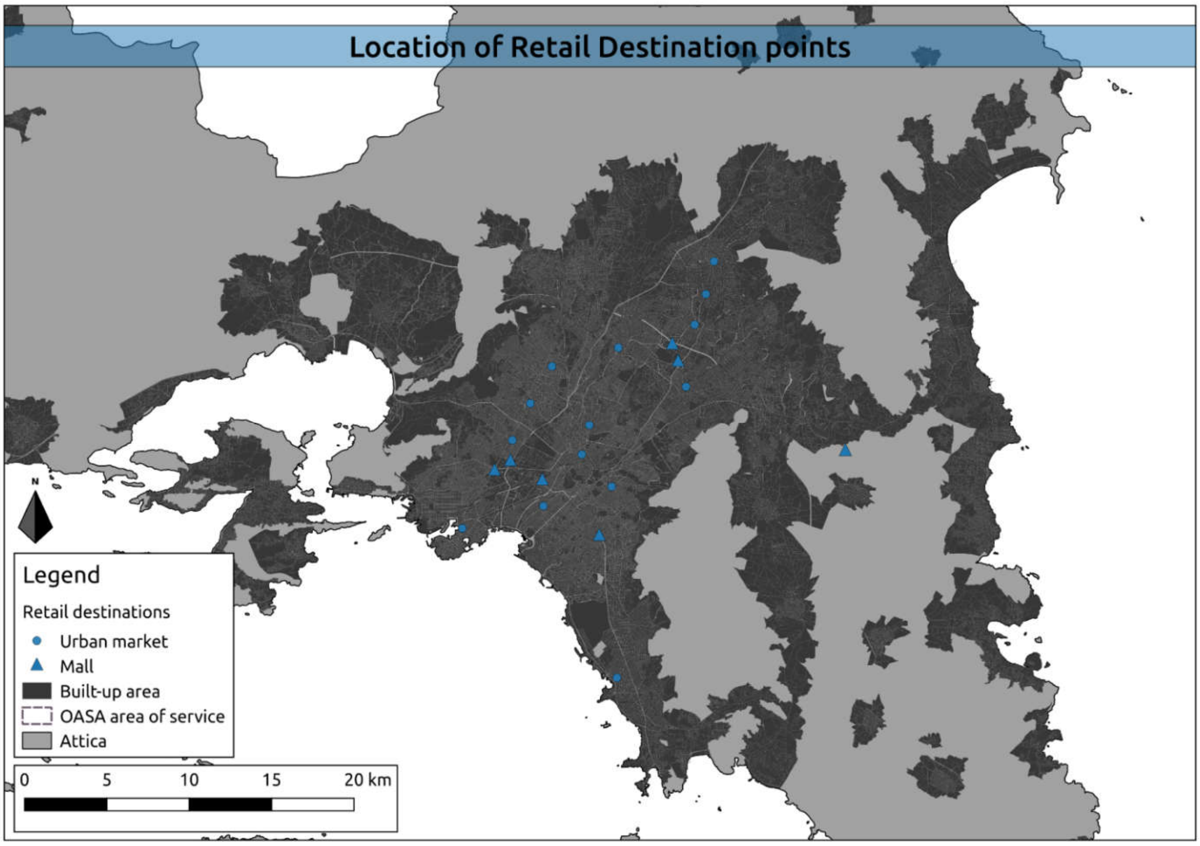


*Figure 4: The location of retail concentrations*

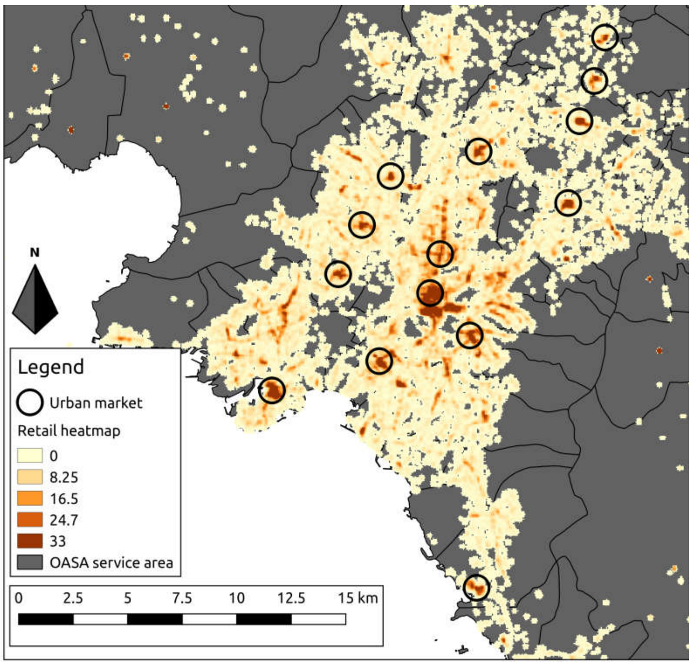


*Figure 5: Concentration of retail business*

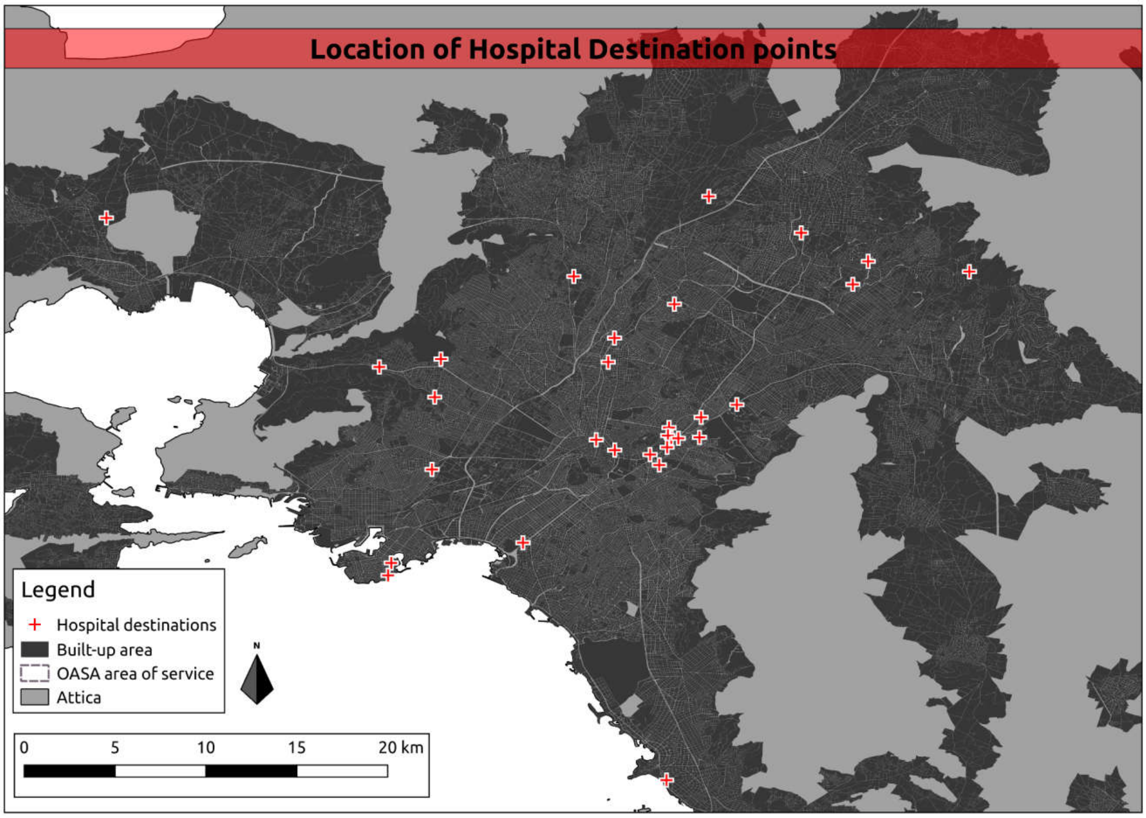


*Figure 6: The location of public hospitals*

### 3.3.3 Calculation of Origin-Destination travel characteristics

The transport connection between Origins and Destinations was calculated using Google Map Direction API. For this reason, a Python script was developed in order to pass the origin and destination locations and handle the response (see Appendix). The original response from Google Map Direction API is a *JSON* file with directions from the origin location to destination. This response was parsed and processed by the Python script in order to create a more compact file that would be easier to import to a database. Each file contains information about the origin point, destination point, total travel distance, total travel time, number of transport modes used and the details of each one (travel distance, travel time, type, line number and stops). An example can be seen in the following box:

9705|6|2772|23.633|3|[[1168, 13.767, 'WALKING'], [1529, 9.0, 'BUS', '560', 7], [75, 0.883, 'WALKING']]

For both datasets the departure time was set for 9 am of a weekday morning with no special traffic regulations, which means that the results and conclusions of this study are limited to this specific time of day.

### 3.3.4 Calculation of Retail and Hospital Access

Retail and Hospital access was calculated as the number of Destinations which are accessible from each Origin within 45 minutes divided by the total number of destinations of this particular land use. The accessibility of each Origin point was then aggregated to zip code level, by calculating its mean value, in order to estimate the average accessibility score of the zip code.

### 3.3.5 Calculation of Population Density

Residential density was calculated by dividing the zip code population by its built-up area, which is different from the zip code area. This is because the boundaries of some zip codes include other land-uses as well like farmlands, forests etc. Including the area of these land-uses dramatically decreases residential population density and may lead to false assumptions.

# 4 Methods

This thesis attempts to answer the research questions by applying two main methods. Firstly spatial autocorrelation statistics, both global and local, are used in order to determine and identify if public transport accessibility to retail and hospitals is a random, clustered or dispersed phenomenon. Secondly, geographic weighted regression is applied to examine if there is a relationship between the public transport and urban structure or socioeconomic factors.

## 4.1 Measuring Accessibility

Public transport accessibility was measured in zip code level by averaging the accessibilities of each of the origins points inside a given zip code. The indicator satisfies two out of the four components according to the typology of Geurs, Karst T. & van Wee (2004), namely the land-use and transportation. This does not mean that the other two components (the temporal dimension and user preferences) are not important. Rather, the reasons for their omission are practical. Inclusion of the temporal component requires multiple runs of the same dataset for different times of the day. This would require more time and, more importantly, higher costs. Similarly, the individual component requires an additional questionnaire survey to assess the individual needs and characteristics of individuals. Since the aim of the study was to cover the whole area of OASA's service it would almost be impossible to gather that many questionnaires, let alone achieve a homogeneous spatial distribution.

## 4.2 Data Collection

In this study the accessibility of an area is seen as relative to a specific social activity. In order to calculate it, a three step approach was taken to compile the dataset. Firstly, because it is impossible (and unnecessary) to measure the accessibility of every point in a continuous surface, the study area was chosen to be represented by points which form a 300m grid in space. These points mentioned from now on as *Origins*.

Secondly, as far as specific types of activities are concerned, the study focuses on retail marketplaces and public health-care. The location of those activities had to be identified in order to calculate the transport relation between them and the Origins. Activity locations are going to be mentioned as *Destinations* for the rest of the document. Thirdly, the transport relation between Origins and Destinations was calculated by Google Maps Directions API. The algorithm constructed for this purpose executed requests from all Origins to all Destinations and stored the total travel distance, total travel time, number of transport mode changes and specific details for each transport mode used (travel distance, travel time, transport type, number of stops).

Additionally, some other datasets were collected for analytical purposes. OASA's public transport network and General Transit Specification Feed (GTFS) for the year 2016 was used in order to calculate the proximity of Origins to public transportation stops and lines. Furthermore, the average annual income by zip code compiled by General Secretary of Information Systems for the year of 2011 was used to assess the overall income level of an area and the population by building block of Hellenic Statistical Authority for the year of 2011 was used to calculate residential densities. All of the above datasets were used as independent variables and are described in the following chapter.

### 4.2.1 Independent Variables

The study makes use of three explanatory variables for accessibility variability. The first is the distance from central business district (CBD) which is also Athens' city center and is represented as a single point located in the Omonoia Square. This can affect accessibility in two ways. Firstly, given that the Destinations are spread throughout Athens, the shortest origin location to reach all of them would be the center of Athens. Therefore, as the distance from the city center increases so does the distance to some of the Destinations (while the distance to other destinations decreases). Secondly, the road network of Athens has a radial pattern with Athens' center as its center. Therefore, it is expected that there are more transport options in its center compared to its periphery.

The second variable is residential density. Residential density is expected to have a positive effect on accessibility. This is because as density rises, the operation of public transport becomes more sustainable because there it can serve more people. Moreover, when taking into account the land use, it is expected that more retail stores and hospitals will be in the vicinity of a high density neighborhoods than in the vicinity of low density ones. It is worth noting that the study uses a simple approach for Hospital access, in which the number of hospitals that are accessible to a location are measured instead of the number of hospital beds. This is because hospitals are still being used in Greece as a first-level medical provider, meaning that people visit hospitals without having first visited a doctor. In addition, if an incident can't be treated in a particular hospital the patient will be transferred to another hospital with an ambulance. Finally, in medical emergencies it is hypothesized that there will be an ambulance call instead of using public transportation in order to reach the hospital.

The third variable is average annual income. The hypothesis behind the effect of this variable is that public transport accessibility will decrease as the income increases. This is because, even though there is a preference of affluent people to live in the outskirts, where there is low public transportation service, their number is far smaller than those who live in under-serviced areas because they have no other choice.

## 4.3 Spatial statistics

The study hypothesizes that the spatial distribution of variability is not random and that it can be explained by urban and socio-economical characteristics. This means that variability in accessibility will be neither homogeneous nor random, but instead it will rise and decline in specific areas according to some underlying causal mechanisms.

### 4.3.1 Spatial Autocorrelation

This study uses both the global as well as the local version of Moran's I statistic in combination with Local Indicators of Spatial Association (LISA) technique in order to examine whether there is spatial autocorrelation and clustering in accessibility. The method was proposed by Anselin (1995). The global Moran's I statistic will yield an overall estimation about the spatial autocorrelation of accessibility while the LISA technique using local Moran's I will identify potential clusters of accessibility.

The basis of clustering estimation is each zip code and the estimation is calculated by taking into consideration the queen contiguity; that is all the neighboring zip codes that share an edge or a corner with the zip code in question. The results will be two maps for each destination land use. The first will depict the clustering results for accessibility. There are five possible categories:

- high-high, indicating that the high accessibility value of the zip code is surrounded by high accessibility values of the neighboring zip codes
- low-low, indicating that the low accessibility of value of the zip code is surrounded by low accessibility values of the neighboring zip codes
- high-low, indicating that the high accessibility value of the zip code is surrounded by low accessibility values of the neighboring zip codes
- low-high, indicating that the low accessibility of value of the zip code is surrounded by high accessibility values of the neighboring zip codes
- insignificant clustering

The second map depicts the *p-values* of the zip codes for the estimation of LISA indicators.

The above to maps will allow for the confirmation or rejection of the null hypothesis as well as shed light on the spatial clustering of accessibility.

### 4.3.2 Ordinary Least Squares Regression

An Ordinary Least Square (OLS) regression will be executed before the Geographically Weighted Regression (GWR), in order to get an overall view of the dataset as well as its global tendencies. The OLS well reveal whether all independent variables (population density, distance from CBD and average annual income) are significant and it will act as a guide for the participating variables in the GWR.

In order to test for multicollinearity of the independent variables the Variance Inflation Factor (VIF) test will be used. The test is the ratio of the variance of the model including all three independent variables divided by the variance of the model which is specified by only one independent variable. The closer its result is to 1 the less are the effects of multicollinearity. Although there is no agreeable value, as Myers suggests (in Field (2009, p. 224) VIF values higher than 10 suggest that there are multicollinearity issues.

Additionally, Koenker (BP) statistic will be used in order to test for model heteroscedasticity. Significant results will mean that the variation in the relationship between dependent values and each independent variable change with changes in independent variable magnitudes. According to ESRI (2018) models with significant heteroscedasticity are considered non-stationary and they are good candidates for Geographical Weighted Regression.

### 4.3.3 Geographically Weighted Regression

Geographically Weighted Regressions are used as the main method of inquiry regarding the association between the dependent variables on retail and Hospital access and the independent variables. The regression was executed two times, one for retail concentrations and one for public hospitals. The dependent variable is the accessibility score of the zip code while the independent variables are the ones discussed in section 4.2.1.

The output of the analysis is a shapefile which contains information for every local model. More specifically, it contains the Local $R^2$ which takes values between 0 and 1 and describes the goodness of fit for the model. The closer the number is to 1 the better is the dependent variable explained by the independent ones. In addition, the coefficients of the independent variables are also contained. The coefficients indicate the magnitude of the effects of the independent variable to the dependent one.

# 5 Results

This chapter contains the results from the methods used in order to answer the research questions. The first section contains an overall overview of the sample with some descriptive statics for the involved variables as well as figures illustrating the spatial dimension of the variables. In the second section, the results of the autocorrelation statistics are presented for the global and the local Moran's I statistic. Finally, the results of GWR are presented in the third section.

## 5.1 Sample overview

This section presents some descriptive statistics of the sample. Table 4 contains the basic statistics for the five variables, in zip code level, used in the regressions. The variables of Retail access and Hospital access measure the accessibility to retail and hospitals. They represent the percentage of the destinations which are accessible from a zip code within 45 minutes. According to it, the mean retail accessibility by public transport in the whole Athens Metropolitan Area is 47.46% while the median is 44.44% which indicates that half the study's zip codes reach up to 44.44% of the retail concentrations. The range, on the other hand, indicates that there are zip codes with absolutely no access to retail concentrations and zip codes which have access to all retail concentrations. Similarly, for public hospital access, the mean accessibility value for the whole sample is 39.35% while the median value is 43.44%. The range for hospital access indicates that there are zip codes with no access to public hospitals and also the highest public hospital accessibility that can be achieved for a zip code is 80%

*Table 4: descriptive statistics of zip codes for participating variables*

| variable | N | mean | sd | median | range |
|---|---|---|---|---|---|
| Retail access (%) | 268 | 47.46 | 27.76 | 44.44 | 0-100.0 |
| Hospital access (%) | 268 | 39.35 | 22.55 | 43.33 | 0-80.0 |
| CBD distance (m) | 268 | 5603.67 | 4585.05 | 4430.56 | 0.1-22709.91 |
| Population density (population/ha) | 268 | 122.88 | 99.96 | 103.75 | 0.0-416.10 |
| Average income (€) | 267 | 26201.00 | 11128.27 | 22547.00 | 16081-122879 |

In **Error! Reference source not found.**, the correlation between each pair of variables is shown. The bottom diagonal contains a scatterplot for each pair of the variables, whereas the top diagonal contains a measure of their correlation. The degree to which each circle is filled with color indicates the value of the correlation coefficient (ρ). The hue indicates whether a correlation is positive (blue) or negative (red). Finally, the tint indicates a qualitative description of the correlation's strength (high, moderate, low). The strong positive correlation between the two accessibility variables indicates that the accessibility to the land uses in question is fairly similar for a given location. A moderate positive correlation is observed between the population density and the retail and Hospital access and small positive correlation between the average income and the distance to CBD. Furthermore, there is a strong negative correlation between both hospital and Retail access to the distance from the CBD. In addition, there is a moderate negative correlation between the population density with both average income and the distance from CBD. Finally, there is a small negative correlation between average income and both accessibility variables.

The relation of the three explanatory variables of the study is further illustrated in **Error! Reference source not found.**. The figure shows a scatterplot of population density and distance from CBD. Additionally, average income is incorporated into the figure by adjusting the color and size of the plot points. More specifically, the radius of each point is an indicator of the average income, while its colors an indicator of its categorization within the variable's standard deviation. It can be observed that low income zip codes tend to have higher population density and closer to the city center than their more affluent counterparts. This is especially true for the high income zip codes which although have a varying degree of distance from the center of Athens, their majority is found on the low side of population density.

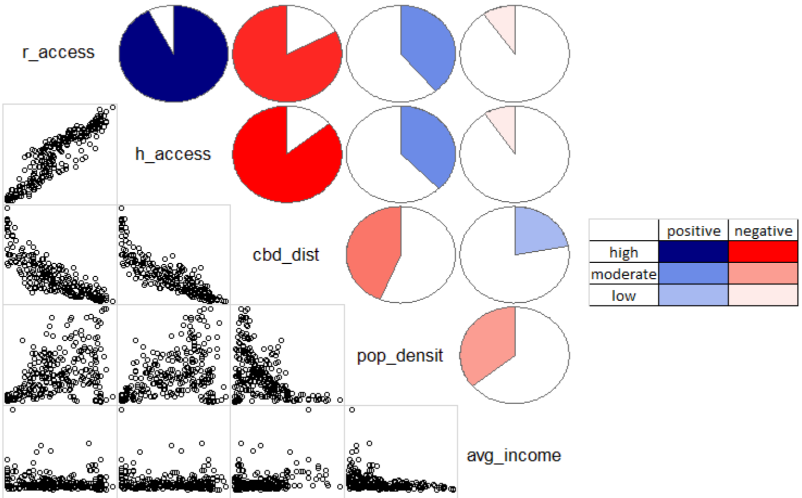


*Figure 7: correlation matrix between dataset variables*

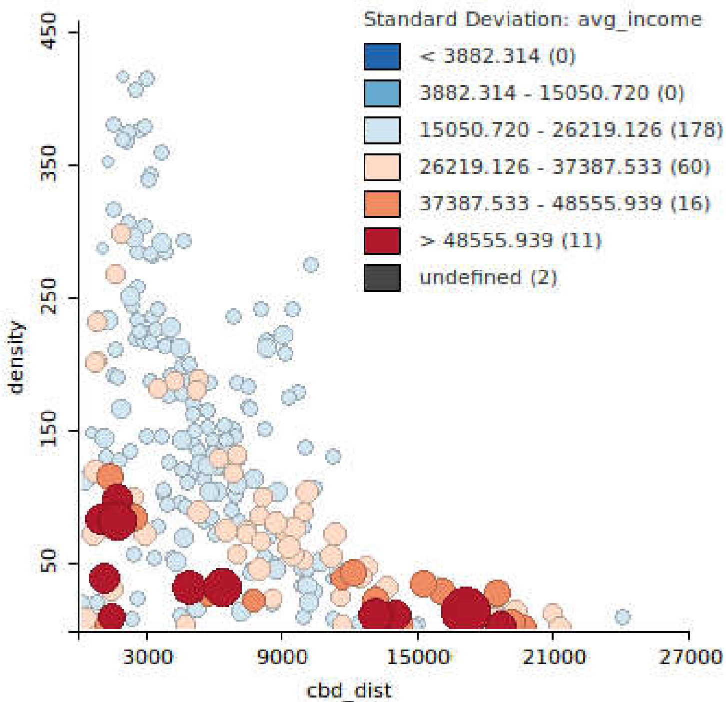


*Figure 8: Scatterplot of CBD distance, Average income and Population density*

The spatial dimension of the variables is presented in Figures 9-12. As can be observed, for both land uses, the accessibility is higher in the city center and decreases in a rather radial pattern. The "corridors" of accessibility that are formed, are caused by the metro lines. The spatial dimension of population density can be seen in Figure 11. According to the figure, the high population density zip codes are clustered around the historical and business center of Athens, where the residential population is rather low. Finally, Figure 12 depicts the average annual income distribution. As can be clearly seen the most affluent are located in the north-east and south-east of Athens as well as some zip codes in the city center.

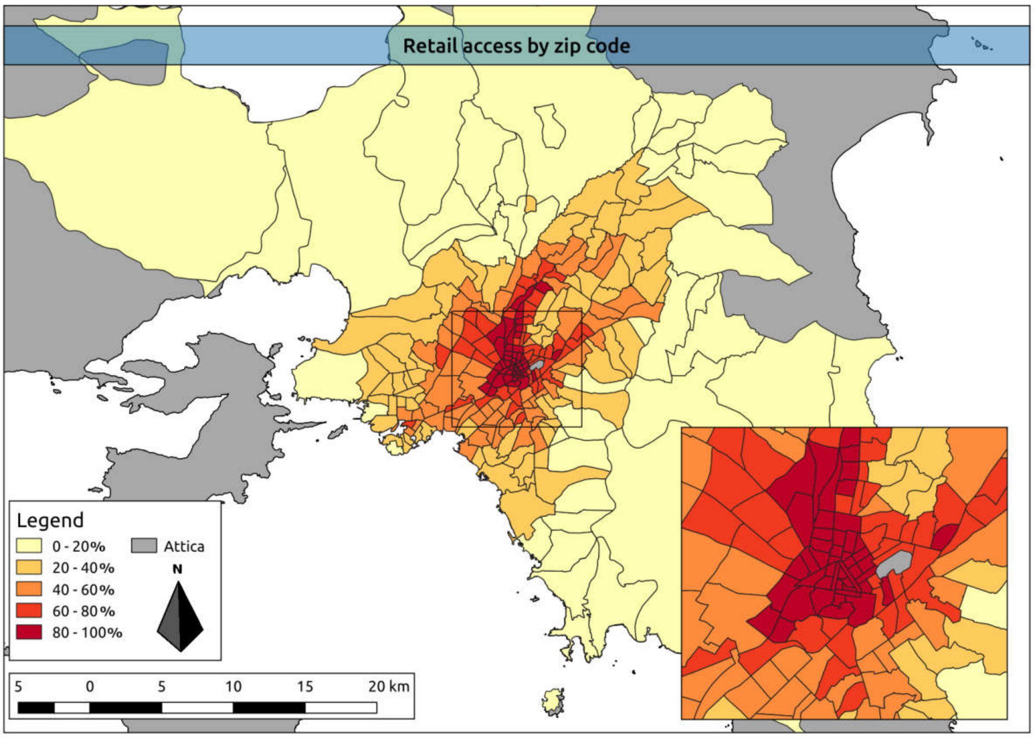


*Figure 9: Accessibility to retail by zip code*

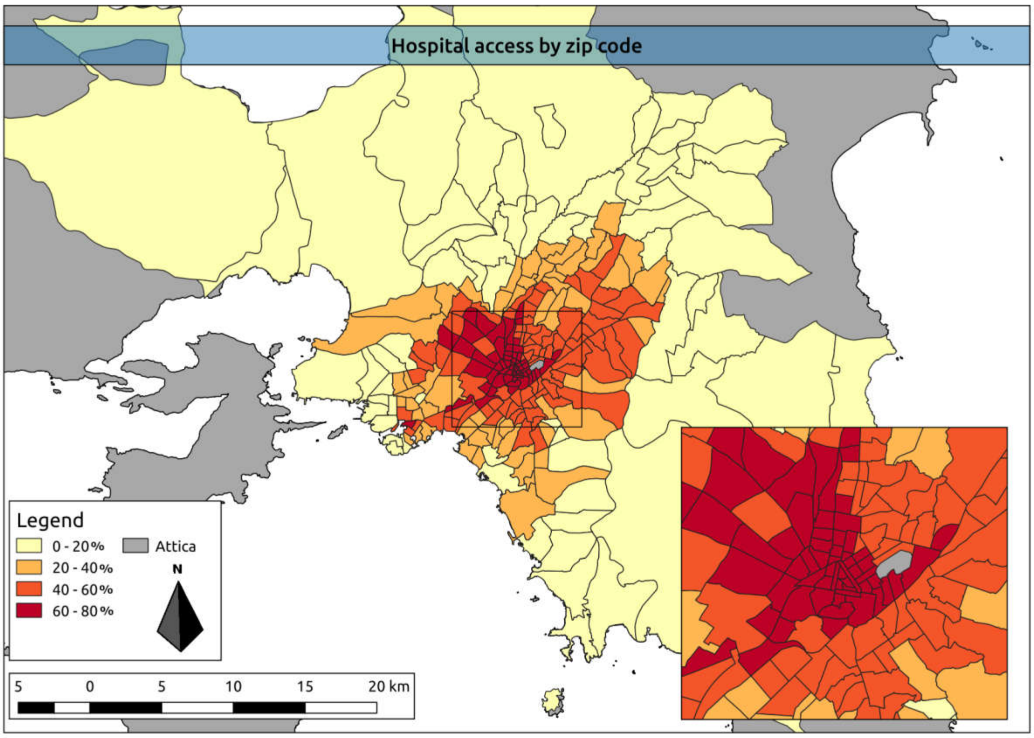


*Figure 10: Accessibility to hospitals by zip code*

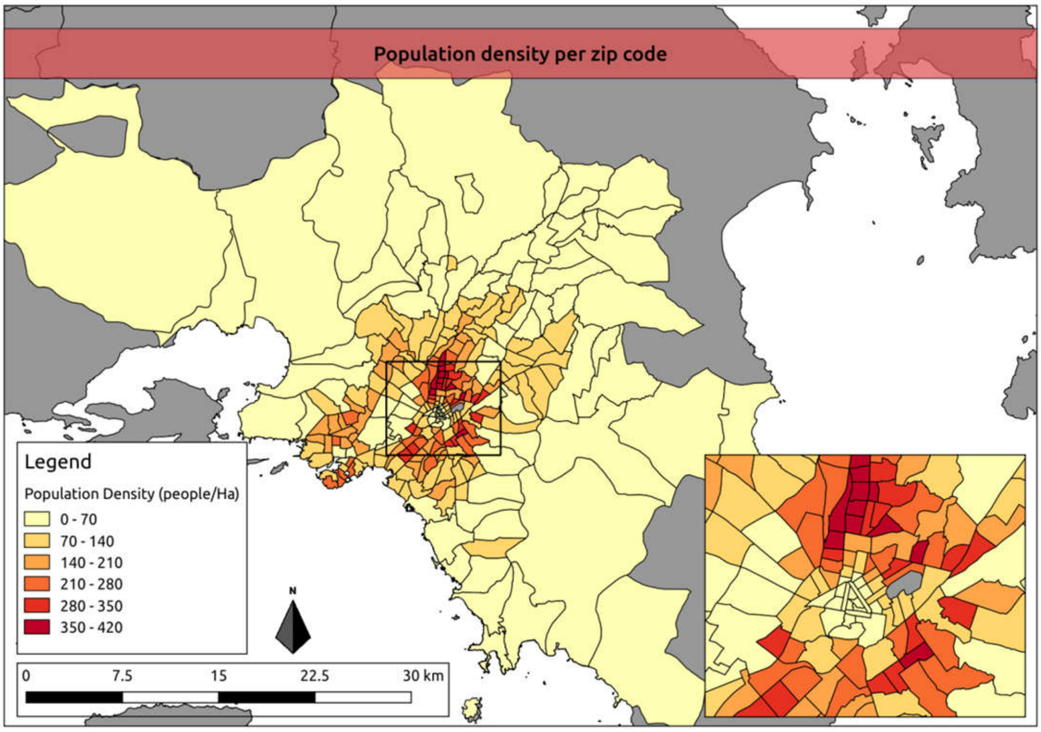


*Figure 11: Spatial distribution of population density*

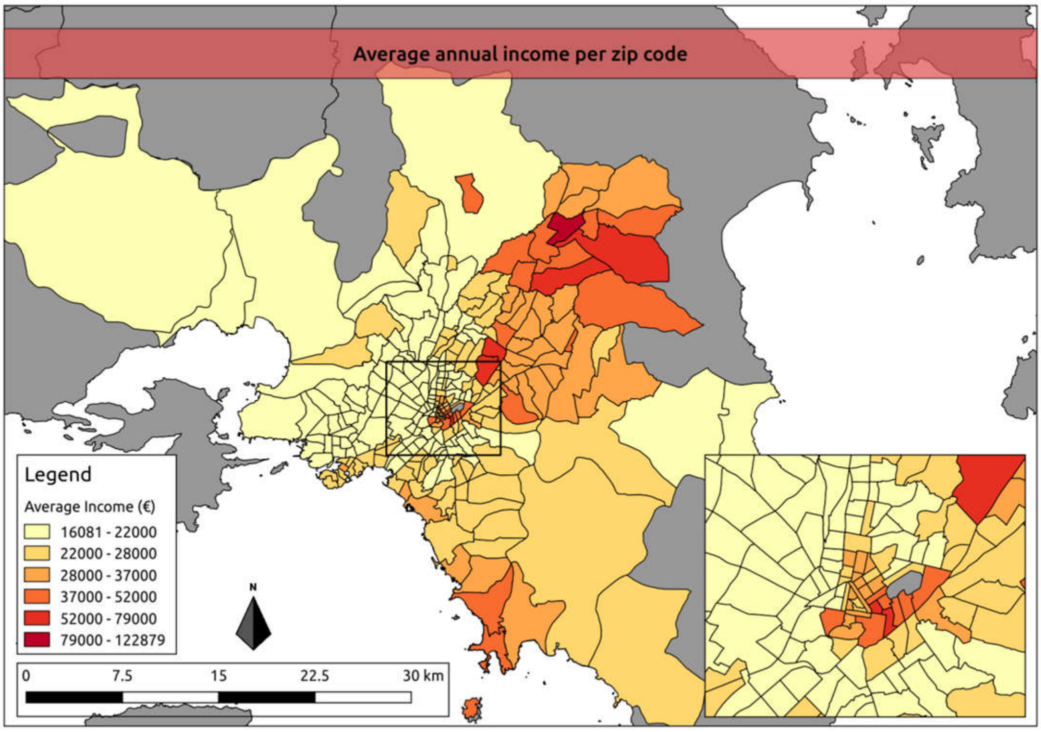


*Figure 12: Spatial distribution of annual average income*

## 5.2 Spatial autocorrelation in accessibility

### 5.2.1 Global Moran's I statistic

The results for the global Moran's I statistic are presented in Table 5. The high positive value of the index in both cases suggests that there is a significant, high spatial autocorrelation between the accessibility values of the zip codes. Thus the null hypothesis for both destination land uses can be rejected.

*Table 5: Global Moran's I summary*

| | Moran's Index | z-score | p-value |
|---|---|---|---|
| Retail access | 0.83 | 23.72 | 0.0000 |
| Hospital access | 0.81 | 23.05 | 0.0000 |

### 5.2.2 Local Indicators of Spatial Association (LISA)

The LISA, using local Moran's Index, shed more light on how the accessibility is clustered spatially. Figure 13, presents the result for the local clustering of Retail access. A cluster of high accessibility is observed in the city center (CBD) followed by a middle ring of insignificant autocorrelation and finally an outer ring of low Retail access clustering. The significance of LISA for retail concentration is depicted in Figure 14. As can be observed, the High-High and Low-Low accessibility clustering for most of the zip codes of Figure 13 is significant.

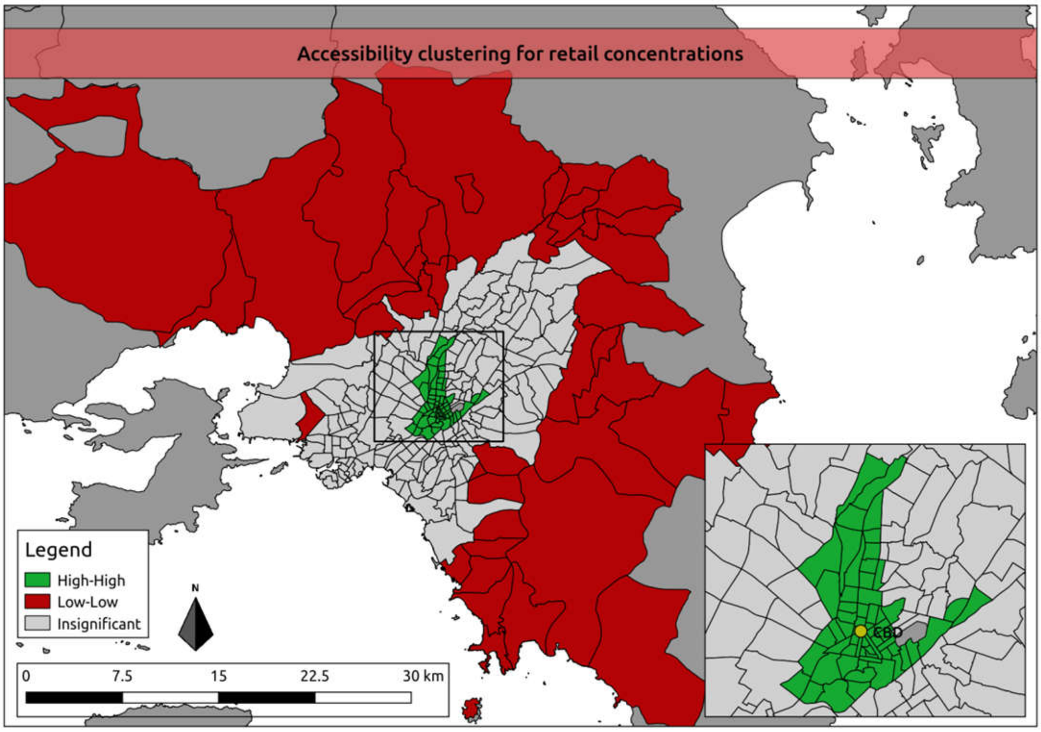


*Figure 13: Accessibility clustering for retail concentrations*

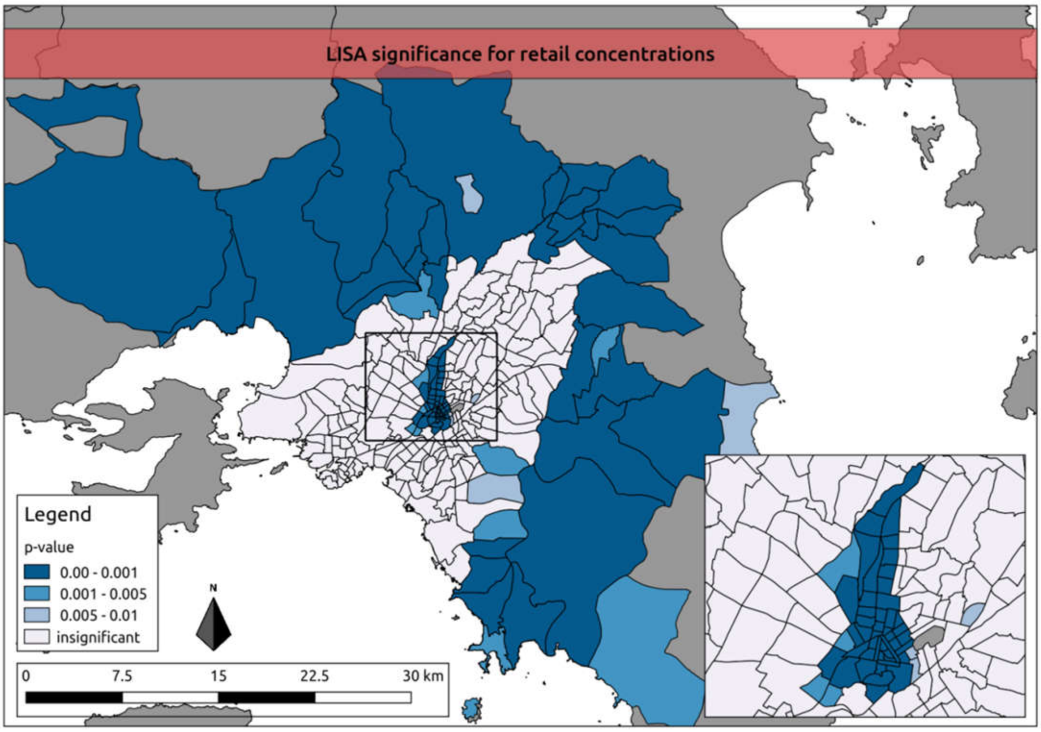


*Figure 14: LISA significance for retail concentrations*

The accessibility clustering for hospitals has more or less the same characteristics (Figure 15). As for the Retail access, an inner but slightly different cluster of high accessibility is observed, followed by a similar middle ring of insignificant clustering and finally the ring of low accessibility clustering. The significance of the respective LISA indicators can be seen in Figure 16. Again, there is significant clustering for most of the High-High and Low-Low zip codes.

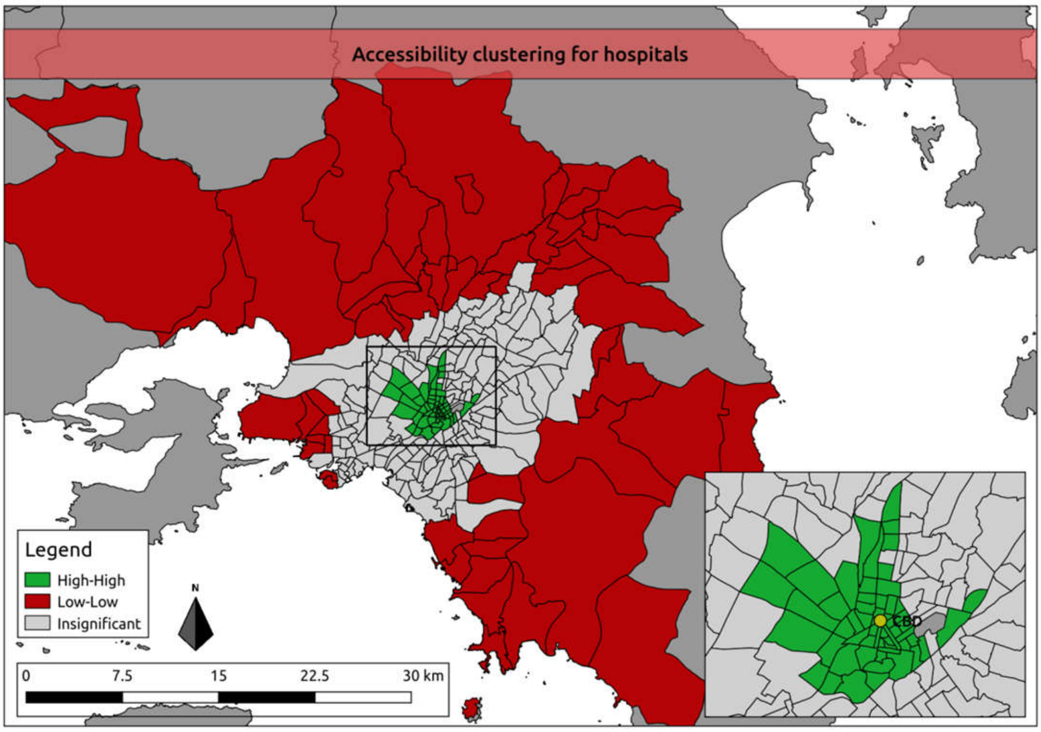


*Figure 15: Accessibility clustering for hospitals*

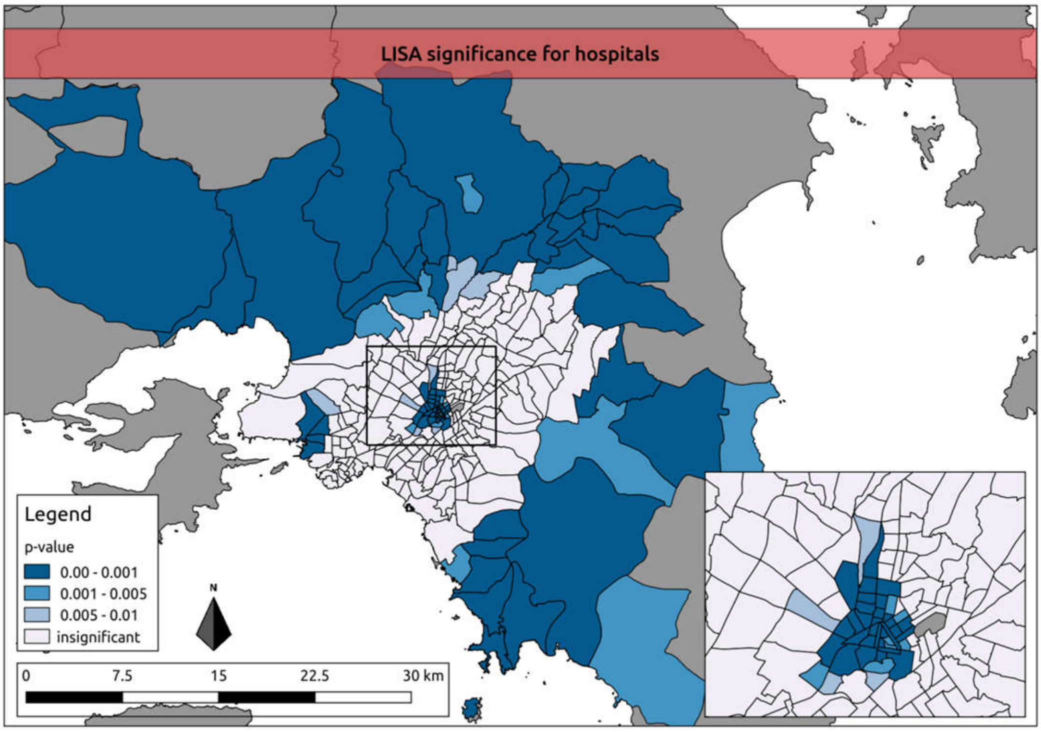


*Figure 16: LISA significance for hospitals*

## 5.3 Ordinary Least Squares regression

Taking a closer look on Retail access, an initial OLS regression reveals the overall tendencies of the model (Table 6). According to the regression, the explanatory power of the model is slightly less than 70% ($R^2$and adj.$R^2$). All our independent variables are significant at $p<0.1$. Therefore we can reject the null hypothesis and assert that all independent variables have a significant effect on Retail Access. The effect of CBD distance is negative and Population density is positive, indicating that the accessibility increases as distance to CBD decreases and as population density increases. In addition, there is a positive relation between Retail access and income, indicating that as income increases so does the accessibility to retail concentrations.

The preliminary OLS regression for Hospital access suggests that the models explanatory power is slightly greater than the one for Retail access. More specifically, the value of $R^2$ and adj.$R^2$ is 75%. However, in this model, population density seems to have an insignificant impact on accessibility. On the other hand, CBD distance and Average income are both significant at $p<0.01$, with CBD distance and Average income having a negative and positive impact, respectively, on accessibility. Therefore the null hypothesis can only be rejected for the two variables (distance from CBD and annual income).

For both models the coefficient values indicate the increase of accessibility score relative to the increase of the independent variables. For instance, in the retail access model, an increase of the distance by 1 meter is associated with a decrease of retail access by 0.005. Using more appropriate values, accessibility decreases by 1% for every 200 meters away from the city center. Similarly, for the hospital model an increase by 1 euro in the annual average income is associated with an increase of accessibility by 0.002%. Again, using more appropriate measures, for every 500 euro increase in annual average income the accessibility is increased by 1%. Finally, for the population density coefficient, a population density increase by 50 population/Ha is associated with an increase in retail access by 1%.

*Table 6: OLS regression summary*

| | Retail access model | Hospital access model |
|---|---|---|
| CBD distance coefficient | -0.005*** | -0.004*** |
| Average income coefficient | 0.0003*** | .0002*** |
| Population density coefficient | 0.020* | 0.008 |
| Constant | 66.031*** | 56.760*** |
| Observations | 267 | 267 |
| R2 | 0.695 | 0.753 |
| Adjusted R2 | 0.691 | 0.750 |
| Residual Std. Error | 15.375 (df = 263) | 11.235 (df = 263) |
| F Statistic | 199.302*** (df = 3; 263) | 266.839*** (df = 3; 263) |

Note: *p<0.1; **p<0.05; ***p<0.01

The models do not seem to have any particular multicollinearity issues as the Variance Inflation Factor (VIF) suggest (Table 7). In addition, the Koenker (BP) statistic for stationarity (Table 8) suggests that Hospital Access model is not stationary across geographical space and therefore, a GWR would be a better method to construct the model.

*Table 7: Variance Inflation Factor for independent variables*

| CBD distance | Average income | Population density |
|---|---|---|
| **1.243714** | 1.154339 | 1.356413 |

*Table 8: Koenker (BP) Satistic*

| Model | BP | Df | P value |
|---|---|---|---|
| **Retail access** | 1.04 | 3 | 0.7916 |
| **Hospital access** | 31.219 | 3 | 7.643e-07 |

## 5.4 Geographically weighted regression

In order to determine whether there is any effect of urban structure or socioeconomic factors on public transport accessibility, two separate GWR models were specified. The dependent variables were retail and Hospital access and the independent variables were distance from CBD (CBD distance), population density (Population density) and average annual income (Average income). The explanatory power of the two models can be seen in Figures 15 & 16. Overall, the models demonstrate a high $R^2$ which suggests that the selected variables adequately explain retail and hospital access. Locally, the power of both models seems to gradually increase as the distance from the CBD increases, but even in the central parts of Athens $R^2$ is greater than 0.5.

### 5.4.1 Accessibility to retail concentrations

The model's explanatory power ($R^2$) for accessibility to retail concentrations ranges roughly from 0.6 to 0.8 (Figure 17a). However, for most zip codes this value is somewhere between 0.6 and 0.7. In addition, there is a clear spatial pattern on its power, as it increases along with the distance from the city center. The summary of the model can be seen in Table 9.

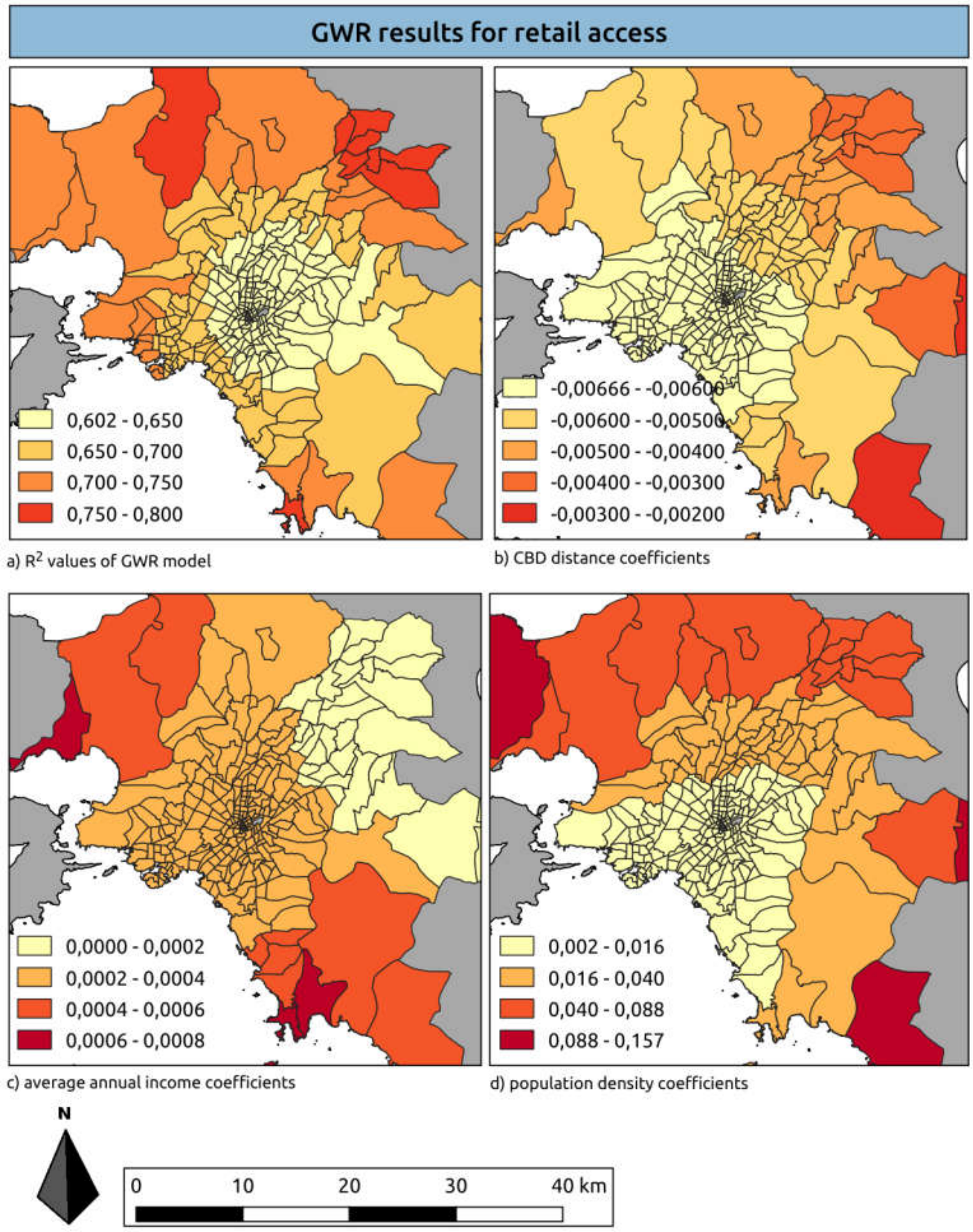


*Figure 17: GWR results for retail access*
*a) $R^2$ values of GWR model b) CBD distance coefficients*
*c) average annual income coefficients d) population density coefficients*

The effect of the independent variables on Retail access can be seen in Figures 15b-d in the form of their estimated coefficients, one for each coefficient. The purpose of the maps is to showcase the effects of each coefficient and by no means to suggest an isolated operation from the others. According to it, distance from CBD (CBD distance, Figure 17b) has a negative effect on Retail access, which is highest for the zip codes to the south-west part of the city, as in this area the regression coefficients have the lowest values. For this area, there is a 0.06% decrease in Retail access for every meter away from the city center. On the contrary, the zip codes which seem to be less affected by the distance to the city center are those in the edges of the area of research.

Average income (Average income, Figure 17c) seems to have a homogenous, positive effect for the most zip codes of the research area. Exceptions to this are the northern suburbs where income has the lowest effect and again the zip codes located to the edges of the research area which seem to be affected the most by it.

Finally, the effects of population density (Population density, Figure 17d) seem to increase in a radial pattern, especially to the north and eastern directions. The largest effect of population density to the accessibility of retail concentrations seem to be for the zip codes at the edge of the research area. For this cases, there is an up to 8.8% increase in Retail access for every 1 person/HA increase in population density.

*Table 9: GWR summary for Retail access*

| **R2** | **R2 adj.** | **AICc** | **Bandwidth** | **Residuals** |
|---|---|---|---|---|
| **0.758** | 0.745 | 2177.34 | 10793.75 | 49235.14 |

### 5.4.2 Accessibility to Hospitals

The model's explanatory power ($R^2$) for accessibility to hospitals ranges roughly from 0.52 to 0.82 (Figure 18a), while for the vast majority of zip codes this value is somewhere between 0.68 and 0.82. However, in this case the model's explanatory power seems to be more homogenous throughout the study area. The model's summary statistics can be seen in Table 8. It has to be noted that since population density was found to be insignificant during the OLS specification, it has been excluded.

Respectively, the coefficients of the independent variables, and hence their effect on accessibility to hospitals can be seen in Figures 18b-c. It is worth mentioning again that the purpose of the maps is to showcase the effects of each coefficient and by no means to suggest an isolated operation from the others. As far as distance from CDB (CBD distance, Figure 18b) is concerned, its effects seem to match the ones of the retail concentration model, at least for the south west part of the research area. However, in this case, the relevance of CBD distance seems to decrease for the north and east zip codes as they get further away from the city center.

Average annual income (Average income, Figure 18c) seems to have a positive effect for all cases. The coefficient values seem to be quite homogeneous as almost all zip codes have values in the same category (0.0001-0.0003). The only exception to that are the two westernmost zip codes of the study as well as the single most south zip code. Other than these, zip codes located to the northeast seem to be less affected by income as they all belong in the lowest category (0-0.0001).

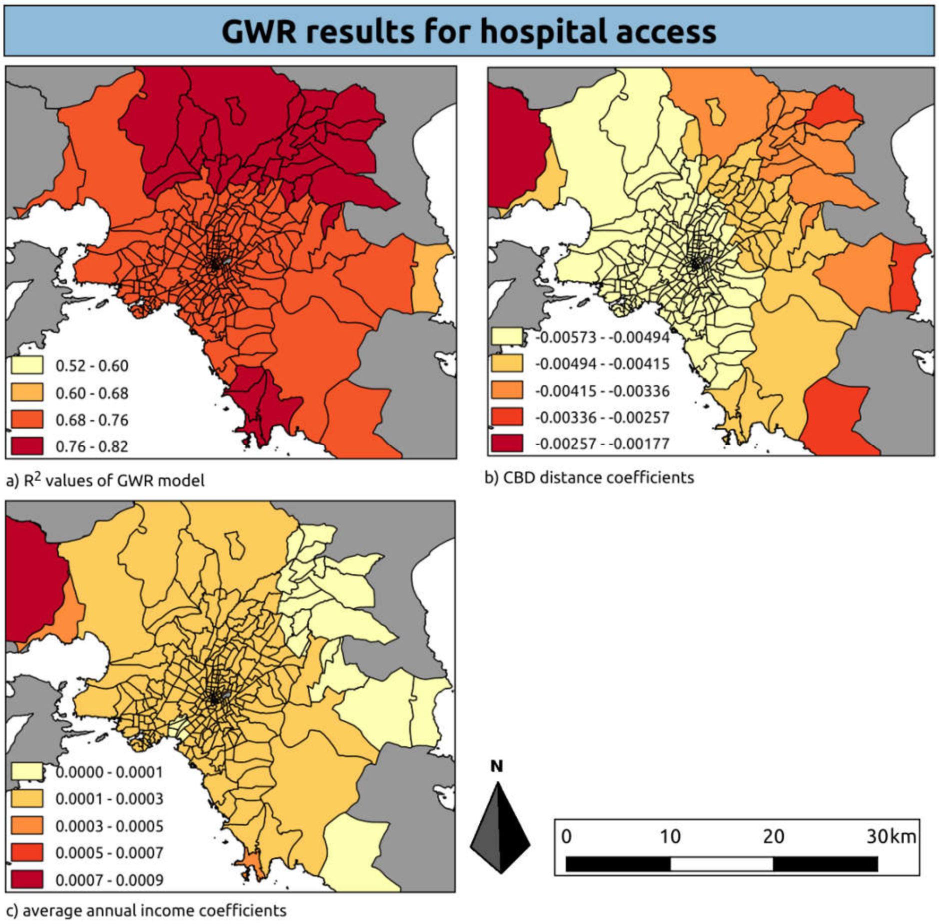


*Figure 18: GWR model performance for hospital accessibility*
*a) $R^2$ values of GWR model b) CBD distance coefficients*
*c) average annual income coefficients*

*Table 10: GWR summary for Hospital access*

| R2 | R2 adj. | AICc | Bandwidth | Residuals |
|---|---|---|---|---|
| 0.807 | 0.798 | 2001.84 | 10793.75 | 25849.14 |

# 6 Discussion

This thesis has set out to answer two questions. Firstly, whether there was any pattern in the spatial distribution of accessibility to retail concentrations and public hospitals by public transport and secondly, whether accessibility can be explained by basic urban structure and socioeconomic variables. Although some patterns and relations seem to be intuitive, what is not intuitive is the magnitude of the effect of the factors on accessibility when they are examined in combination. Before proceeding, it has to be highlighted that the following findings are solely based on the research's results and are limited to the specific parameters of the study, including the travel time threshold (45 min), the date and time of the estimation (weekday 9am), the selection of the destinations (retail concentration and public hospitals), the definition of accessibility (percentage of destinations reachable within 45 min) and the overall presentation of figures and maps.

For the first part, the global Moran's I suggests that there is significant spatial autocorrelation of accessibility and the LISA analysis based on Moran's I indicates that the distribution of accessibility by public transport to retail concentrations and public hospitals has two significant clusters: one of High to High and one Low to Low accessibility. More specifically, there is a high accessibility cluster in the zip codes located in the core of Athens' Metropolitan network and a low accessibility cluster in the zip codes of the periphery. There is no evidence that there are other clusters of high or low accessibility within the research area.

In the middle ring, which is located between the high accessibility and low accessibility clusters, the analyses suggests that there is no spatial clustering for accessibility, suggesting that the accessibility on those zip codes is circumstantial. A reason for this may be the location of metro stations inside this middle ring. Metro stations dramatically increase the accessibility to other land uses in general, creating islands of high accessibility amid an ocean of inaccessibility. However, this is not enough to create a local cluster of accessibility or a cluster of inaccessibility as a high accessibility zip code due a metro station is surrounded by low accessibility zip codes with no metro station. The same effect is created by bus lines running on major roads.

As far as explaining accessibility, the OGR and GWR models confirm that there is indeed significant prediction of accessibility by the independent variables as all of the Null Hypotheses are rejected except from the effects of population density on accessibility to public hospitals. However, it has to be noted that the difference in explanatory power of the model between the city center and periphery for the case of retail accessibility, suggests that there might be other factors in operation which affect accessibility. As expected, the distance from the CBD has a negative impact on accessibility for both land uses. The fact that the effects of the distance (as they are perceived through the model's CBD coefficients) become less important for the zip codes located in the edge of our research area, far from the city center, may imply that beyond a particular distance from the city center, the effects of distance on accessibility decrease as the accessibility has stabilized in low values and it is no further decreasing.

The results confirm that the association between population density and accessibility by public transport is significant as hypothesized only for the case of retail concentrations. The importance of the population density increases along with the distance from the city center, suggesting that population density is a critical factor for public transport in areas away from the city center.

Finally, the output of both models suggests that there is a positive relation between income and accessibility. The only exception is the local negative effect of income to Hospital access for the zip codes to the north edge of the study. Despite of this, the positive link between the two variables suggests that there is an underlying social equity concern. Accessibility by public transport should be higher in low-income zip codes where car use may either be not an option or a very expensive one.

At this point and from a planning perspective, it has to be noted that accessibility as was defined by Geurs, Karst T. & van Wee (2004), is not a characteristic of the transportation network alone, but a combination of the transport network with the land use pattern. Under this light, inequalities in accessibility are not only caused by an unjust transportation network, but also by an unjust land use pattern. This is especially relevant for Hospital access, since the location of public hospitals is a result of a very specific planning process.

Although this thesis brings new knowledge on the distribution of accessibility and its association with urban structure and socioeconomic factors, it is not free of limitations. Firstly, in order to investigate the effects of accessibility on vulnerable social groups, one has to use a questionnaire survey, something which was not possible within the framework of this thesis. This is mainly related with the access to the transportation network itself. So for instance, the access to a metro station or bus stop can differ for certain social groups with disabilities. The analysis that was conducted by the study did not take into consideration such factors. The reason for this is not that they are insignificant (on the contrary) but because the inclusion of such factors requires additional data collection techniques, something that was not possible under the scope of this Thesis. In addition, the initial intention of the thesis was to study accessibility in a more detailed spatial level than that of the zip code. The goal was to use each point in the grid as an input for the spatial autocorrelation analysis and GWR. This was not possible due to the fact that the available population and income data were aggregated to zip code level. Sequentially, this caused a problem for local coefficient estimation in GWR due to the fact that different accessibility values had the same set of values for their independent variables because they were inside a given zip code.

Second, the definition of accessibility is something which can greatly affect the results of the research. This thesis used a rather simple but straightforward definition, lacking the resources for a more complex one which would require additional data and probably disaggregate in nature. This was also the reason why the all travel times within the 45 min threshold have the same weight. Although it is clear that having one hospital within 5 minutes is more preferably than having a hospital within 45 minutes, the actual weighing of the two choices (and all the possible alternatives) requires disaggregate stated preference data.

Third, due to area differences of the zip codes the accessibility scores of each zip code are calculated using different number of points. Small zip codes in the city center may include just a couple of origin points while large zip codes of the periphery can include significant more origin points. This may cause small zip codes to have more extreme accessibility values while large zip codes to smooth accessibility differences. Essentially, this issue is linked with the issue discussed in the previous paragraph, that of the availability of population and income data for more detailed spatial levels that forced the analysis to be carried out at the zip code level.

Under this light, although the Thesis answers the questions it set in the beginning, it also opens new questions for further research. One of them is the inclusion of the other two components of Geurs, Karst T. & van Wee (2004) concept of accessibility, namely the temporal and the individual. The same process could be carried for different time of the day as well as the weekends, in order to have better overall picture of accessibility. In addition, the individual component requires questionnaire surveys and interviews in order to uncover new factors affecting accessibility as well as understand what the concept means for different people. Finally, there are other land uses worth investigating such as green/open areas, cultural places and education.

# 7 Conclusions

This thesis studied the distribution of accessibility by public transport to retail concentrations and public hospitals by the use of spatial analysis methods namely spatial autocorrelation and geographically weighted regression. It has to be highlighted that the following conclusions are solely based on the research's results and are limited to the specific parameters of the study, including the travel time threshold (45 min), the date and time of the estimation (weekday 9am), the selection of the destinations (retail concentration and public hospitals), the definition of accessibility (percentage of destinations reachable withn 45 min) and the overall presentation of figures and maps.

In order to achieve this, the travel time to 21 retail and 30 hospital destinations was measured from a grid of 6500 points. The accessibility score of each point to retail and hospital destination was calculated as the percentage of accessible destination within 45 minutes. These scores were then aggregated to zip code level and produced the average accessibility for each zip code. Along with the zip code distance to CBD, population density and annual average income, they form the final dataset for the analysis.

The global Moran's I and LISA of local Moran's I indicated that there is a cluster of high accessibility for both land used located in the city center, followed by a ring of insignificant autocorrelation and by an outer ring of low accessibility clustering. The findings found some support for the initial hypothesis, though more clusters were expected to form in the middle ring.

The hypothesized associations between accessibility and urban structure and socioeconomic factors were confirmed by both OLS regression and GWR, except for the case of population density in the public hospitals model. The model, although simple, manages to explain a significant portion of the variability in accessibility. However, the GWR models provide a better insight regarding the local effects of the variables. Distance from CBD was found to be negatively related to accessibility, meaning that as distance from CBD increases the accessibility decreases. On the other hand, there is a positive relation of population density and annual income to accessibility.

# 8 References


Andris, G. (2016, December 12). Καταθέτουν τις πινακίδες τους για να γλιτώσουν από Τέλη Κυκλοφορίας, τεκμήρια και χαράτσια. *Enikonomia*. Retrieved from http://www.enikonomia.gr/auto/132552,katathetoun-pinakides-kykloforias-gia-na-glitosoun-apo-teli-kyklo.html

Anselin, L. (1995). Local indicators of spatial association – LISA. *Geographical Analysis*, 27, 93-115

Albacete, X., Olaru, D., Paül, V., & Biermann, S. (2017). Measuring the Accessibility of Public Transport: A Critical Comparison Between Methods in Helsinki. *Applied Spatial Analysis and Policy*, *10*(2), 161–188.

Anselin, L., &Bera, A. (1998). Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. In A. Ullah & D. E. Giles (Eds.), *Handbook of Applied Economic Statistics* (pp. 237–289). New York: Marcel Dekker.

Anselin, L., & Rey, S. J. (2014). *Modern Spatial Econometrics in Practice: A Guide to GeoDa, GeoDaSpace and PySAL*. Chicago, IL: GeoDa Press LLC.

Cass, N., Shove, E., & Urry, J. (2005). Social exclusion, mobility and access. *Sociological Review*, *53*(3), 539–555.

Cheng, Y. H., & Chen, S. Y. (2015). Perceived accessibility, mobility, and connectivity of public transportation systems. *Transportation Research Part A: Policy and Practice*, *77*, 386–403.

ESRI. (2018). How OLS works. Retrieved from. http://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/how-ols-regression-works.htm

Field, A. (2009). Discovering Statistics Using SPSS. London: SAGE

Fotheringham, A. S., & Charlton, M. E. (1998). Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*, *30*, 1905–1927.

Fransen, K., Neutens, T., Farber, S., De Maeyer, P., Deruyter, G., &Witlox, F. (2015). Identifying public transport gaps using time-dependent accessibility levels. *Journal of Transport Geography*, *48*, 176–187.

Geurs, Karst T. & van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography*, *12*(2), 127–140.

Geurs, K. T., De Montis, A., &Reggiani, A. (2015). Recent advances and applications in accessibility modelling. *Computers, Environment and Urban Systems*, *49*, 82–85.

Goodchild, M. F. (1992). Geographical Information Science. *International Journal of Geographical Information Science*, *6*(1), 31–45.

Grengs, J. (2012). Equity and the social distribution of job accessibility in Detroit. *Environment and Planning B: Planning and Design*, *39*(5), 785–800. http://doi.org/10.1068/b36097

Hägerstraand, T. (1970). WHAT ABOUT PEOPLE IN REGIONAL SCIENCE? *Papers in Regional Science*, *24*(1), 7–24.

Hansen, W. G. (1959). How Accessibility Shapes Land Use. *Journal of the American Institute of Planners*, *25*(2), 73–76.

Lättman, K., Friman, M., & Olsson, L. E. (2016). Perceived Accessibility of Public Transport as a Potential Indicator of Social Inclusion. *Social Inclusion*, *4*(3), 36.

Lei, T. L., & Church, R. L. (2010). Mapping transit-based access: Integrating GIS, routes and schedules. *International Journal of Geographical Information Science*, *24*(2), 283–304.

Milakis, D., Cervero, R., van Wee, B., & Maat, K. (2015). Do people consider an acceptable travel time? Evidence from Berkeley, CA. *Journal of Transport Geography*, *44*, 76–86.

Mokhtarian, P., & Salomon, I. (2001). How Derived is the Demand for Travel? Some Conceptual and Measurement Considerations. *Transportation Research Part A*, *35*, 695–719.

Neutens, T., Schwanen, T., Witlox, F., & de Maeyer, P. (2010). Equity of urban service delivery: A comparison of different accessibility measures. *Environment and Planning A*, *42*(7), 1613–1635.

Rogerson, A. P. (2001). *Statistical Methods for Geography*. London, Thousand Oaks and New Delhi: SAGE.

Saghapour, T., Moridpour, S., & Thompson, R. G. (2016). Public transport accessibility in metropolitan areas: A new approach incorporating population density. *Journal of Transport Geography*, *54*, 273–285.

van Wee, B. (2016). Accessible accessibility research challenges. *Journal of Transport Geography*, *51*, 9–16.

Wee, B., &Geurs, K. (2011). Discussing equity and social exclusion in acessibility evaluations. *European Journal of Transport and Infrastructure Research*, *11*(4), 350–367.

# 9 Appendix 1

**Python script which handles the request and response to and from Google Directions API**

```
from gmaps import Directions
import json, time

api=Directions(api_key='AIzaSyDXuiOSoEJk6Vb7Rm8ksjJSNIMrh54ijPQ')

def steps_breakdown(steps):
    steplist=[]
    for step in range(len(steps)):
        step_travel_mode=steps[step]['travel_mode']
        step_travel_distance=steps[step]['distance']['value']
        step_travel_time=round(steps[step]['duration']['value']/60.0,3)

        if step_travel_mode=='TRANSIT':
            step_line_type=steps[step]['transit_details']['line']['vehicle']['type']
            step_num_stops=steps[step]['transit_details']['num_stops']
            steplist.append([step_travel_distance,
                             step_travel_time,
                                 step_line_type])
            try:
                step_line_shortname=steps[step]['transit_details']['line']['short_name']
                steplist[step].append(step_line_shortname)
            except:
                pass

            steplist[step].append(step_num_stops)

        else:
            step_tm=steps[step]['travel_mode']
            steplist.append([
                     step_travel_distance,
                     step_travel_time,
                       step_tm])
    return steplist

with open("RemainingOriginsforOD3.csv") as o:
    for line in o.readlines():
```

```
        org=line[:-1].split("|")
        with open("HospitalDestinations3.csv") as d:
            for dline in d.readlines():
                try:
                    dest=dline[:-1].split("|")
                    directions = api.directions(origin=org[1],
                                         destination=dest[1],
                                         mode="transit",
                                         departure_time='1513751600')

                    data=json.loads(json.dumps(directions))

                    print(org[0],dest[0],data[0]['legs'][0]['distance']['value'],
                  round(data[0]['legs'][0]['duration']['value']/60.0,3),
                    len(data[0]['legs'][0]['steps']),
                  steps_breakdown(data[0]['legs'][0]['steps']), sep="|")
                    time.sleep(0.3)
                except Exception as e:
                    print(org[0],dest[0],"no transit available", sep="|")
                    time.sleep(0.3)
                    if dest[0]=='1':
                        break
                    else:
                        continue
```

Series from Lund University

Department of Physical Geography and Ecosystem Science

**Master Thesis in Geographical Information Science**

1. *Anthony Lawther:* The application of GIS-based binary logistic regression for slope failure susceptibility mapping in the Western Grampian Mountains, Scotland (2008).
2. *Rickard Hansen:* Daily mobility in Grenoble Metropolitan Region, France. Applied GIS methods in time geographical research (2008).
3. *Emil Bayramov:* Environmental monitoring of bio-restoration activities using GIS and Remote Sensing (2009).
4. *Rafael Villarreal Pacheco:* Applications of Geographic Information Systems as an analytical and visualization tool for mass real estate valuation: a case study of Fontibon District, Bogota, Columbia (2009).
5. *Siri Oestreich Waage:* a case study of route solving for oversized transport: The use of GIS functionalities in transport of transformers, as part of maintaining a reliable power infrastructure (2010).
6. *Edgar Pimiento:* Shallow landslide susceptibility – Modelling and validation (2010).
7. *Martina Schäfer:* Near real-time mapping of floodwater mosquito breeding sites using aerial photographs (2010).
8. *August Pieter van Waarden-Nagel:* Land use evaluation to assess the outcome of the programme of rehabilitation measures for the river Rhine in the Netherlands (2010).
9. *Samira Muhammad:* Development and implementation of air quality data mart for Ontario, Canada: A case study of air quality in Ontario using OLAP tool. (2010).
10. *Fredros Oketch Okumu*: Using remotely sensed data to explore spatial and temporal relationships between photosynthetic productivity of vegetation and malaria transmission intensities in selected parts of Africa (2011).
11. *Svajunas Plunge:* Advanced decision support methods for solving diffuse water pollution problems (2011).
12. *Jonathan Higgins:* Monitoring urban growth in greater Lagos: A case study using GIS to monitor the urban growth of Lagos 1990 - 2008 and produce future growth prospects for the city (2011).
13. *Mårten Karlberg:* Mobile Map Client API: Design and Implementation for Android (2011).
14. *Jeanette McBride:* Mapping Chicago area urban tree canopy using color infrared imagery (2011).
15. *Andrew Farina:* Exploring the relationship between land surface temperature and vegetation abundance for urban heat island mitigation in Seville, Spain (2011).
16. *David Kanyari*: Nairobi City Journey Planner: An online and a Mobile Application (2011).

17. *Laura V. Drews:* Multi-criteria GIS analysis for siting of small wind power plants - A case study from Berlin (2012).
18. *Qaisar Nadeem:* Best living neighborhood in the city - A GIS based multi criteria evaluation of ArRiyadh City (2012).
19. *Ahmed Mohamed El Saeid Mustafa:* Development of a photo voltaic building rooftop integration analysis tool for GIS for Dokki District, Cairo, Egypt (2012).
20. *Daniel Patrick Taylor*: Eastern Oyster Aquaculture: Estuarine Remediation via Site Suitability and Spatially Explicit Carrying Capacity Modeling in Virginia's Chesapeake Bay (2013).
21. *Angeleta Oveta Wilson:* A Participatory GIS approach to *unearthing* Manchester's Cultural Heritage '*gold mine'* (2013).
22. *Ola Svensson:* Visibility and Tholos Tombs in the Messenian Landscape: A Comparative Case Study of the Pylian Hinterlands and the Soulima Valley (2013).
23. *Monika Ogden:* Land use impact on water quality in two river systems in South Africa (2013).
24. *Stefan Rova:* A GIS based approach assessing phosphorus load impact on Lake Flaten in Salem, Sweden (2013).
25. *Yann Buhot:* Analysis of the history of landscape changes over a period of 200 years. How can we predict past landscape pattern scenario and the impact on habitat diversity? (2013).
26. *Christina Fotiou:* Evaluating habitat suitability and spectral heterogeneity models to predict weed species presence (2014).
27. *Inese Linuza:* Accuracy Assessment in Glacier Change Analysis (2014).
28. *Agnieszka Griffin:* Domestic energy consumption and social living standards: a GIS analysis within the Greater London Authority area (2014).
29. *Brynja Guðmundsdóttir:* Detection of potential arable land with remote sensing and GIS - A Case Study for Kjósarhreppur (2014).
30. *Oleksandr Nekrasov:* Processing of MODIS Vegetation Indices for analysis of agricultural droughts in the southern Ukraine between the years 2000-2012 (2014).
31. *Sarah Tressel:* Recommendations for a polar Earth science portal in the context of Arctic Spatial Data Infrastructure (2014).
32. *Caroline Gevaert:* Combining Hyperspectral UAV and Multispectral Formosat-2 Imagery for Precision Agriculture Applications (2014).
33. *Salem Jamal-Uddeen:* Using GeoTools to implement the multi-criteria evaluation analysis - weighted linear combination model (2014).
34. *Samanah Seyedi-Shandiz:* Schematic representation of geographical railway network at the Swedish Transport Administration (2014).
35. *Kazi Masel Ullah:* Urban Land-use planning using Geographical Information System and analytical hierarchy process: case study Dhaka City (2014).
36. *Alexia Chang-Wailing Spitteler:* Development of a web application based on MCDA and GIS for the decision support of river and floodplain rehabilitation projects (2014).

37. *Alessandro De Martino:* Geographic accessibility analysis and evaluation of potential changes to the public transportation system in the City of Milan (2014).
38. *Alireza Mollasalehi:* GIS Based Modelling for Fuel Reduction Using Controlled Burn in Australia. Case Study: Logan City, QLD (2015).
39. *Negin A. Sanati:* Chronic Kidney Disease Mortality in Costa Rica; Geographical Distribution, Spatial Analysis and Non-traditional Risk Factors (2015).
40. *Karen McIntyre:* Benthic mapping of the Bluefields Bay fish sanctuary, Jamaica (2015).
41. *Kees van Duijvendijk:* Feasibility of a low-cost weather sensor network for agricultural purposes: A preliminary assessment (2015).
42. *Sebastian Andersson Hylander:* Evaluation of cultural ecosystem services using GIS (2015).
43. *Deborah Bowyer:* Measuring Urban Growth, Urban Form and Accessibility as Indicators of Urban Sprawl in Hamilton, New Zealand (2015).
44. *Stefan Arvidsson:* Relationship between tree species composition and phenology extracted from satellite data in Swedish forests (2015).
45. *Damián Giménez Cruz*: GIS-based optimal localisation of beekeeping in rural Kenya (2016).
46. *Alejandra Narváez Vallejo:* Can the introduction of the topographic indices in LPJ-GUESS improve the spatial representation of environmental variables? (2016).
47. *Anna Lundgren:* Development of a method for mapping the highest coastline in Sweden using breaklines extracted from high resolution digital elevation models (2016).
48. *Oluwatomi Esther Adejoro:* Does location also matter? A spatial analysis of social achievements of young South Australians (2016).
49. *Hristo Dobrev Tomov:* Automated temporal NDVI analysis over the Middle East for the period 1982 - 2010 (2016).
50. *Vincent Muller:* Impact of Security Context on Mobile Clinic Activities A GIS Multi Criteria Evaluation based on an MSF Humanitarian Mission in Cameroon (2016).
51. *Gezahagn Negash Seboka:* Spatial Assessment of NDVI as an Indicator of Desertification in Ethiopia using Remote Sensing and GIS (2016).
52. *Holly Buhler:* Evaluation of Interfacility Medical Transport Journey Times in Southeastern British Columbia. (2016).
53. *Lars Ole Grottenberg*: Assessing the ability to share spatial data between emergency management organisations in the High North (2016).
54. *Sean Grant:* The Right Tree in the Right Place: Using GIS to Maximize the Net Benefits from Urban Forests (2016).
55. *Irshad Jamal:* Multi-Criteria GIS Analysis for School Site Selection in Gorno-Badakhshan Autonomous Oblast, Tajikistan (2016).
56. *Fulgencio Sanmartín:* Wisdom-volkano: A novel tool based on open GIS and time-series visualization to analyse and share volcanic data (2016).

57. *Nezha Acil:* Remote sensing-based monitoring of snow cover dynamics and its influence on vegetation growth in the Middle Atlas Mountains (2016).
58. *Julia Hjalmarsson:* A Weighty Issue: Estimation of Fire Size with Geographically Weighted Logistic Regression (2016).
59. *Mathewos Tamiru Amato:* Using multi-criteria evaluation and GIS for chronic food and nutrition insecurity indicators analysis in Ethiopia (2016).
60. *Karim Alaa El Din Mohamed Soliman El Attar:* Bicycling Suitability in Downtown, Cairo, Egypt (2016).
61. *Gilbert Akol Echelai:* Asset Management: Integrating GIS as a Decision Support Tool in Meter Management in National Water and Sewerage Corporation (2016).
62. *Terje Slinning:* Analytic comparison of multibeam echo soundings (2016).
63. *Gréta Hlín Sveinsdóttir:* GIS-based MCDA for decision support: A framework for wind farm siting in Iceland (2017).
64. *Jonas Sjögren:* Consequences of a flood in Kristianstad, Sweden: A GIS-based analysis of impacts on important societal functions (2017).
65. *Nadine Raska:* 3D geologic subsurface modelling within the Mackenzie Plain, Northwest Territories, Canada (2017).
66. *Panagiotis Symeonidis*: Study of spatial and temporal variation of atmospheric optical parameters and their relation with PM 2.5 concentration over Europe using GIS technologies (2017).
67. *Michaela Bobeck:* A GIS-based Multi-Criteria Decision Analysis of Wind Farm Site Suitability in New South Wales, Australia, from a Sustainable Development Perspective (2017).
68. *Raghdaa Eissa*: Developing a GIS Model for the Assessment of Outdoor Recreational Facilities in New Cities Case Study: Tenth of Ramadan City, Egypt (2017).
69. *Zahra Khais Shahid*: Biofuel plantations and isoprene emissions in Svea and Götaland (2017).
70. *Mirza Amir Liaquat Baig*: Using geographical information systems in epidemiology: Mapping and analyzing occurrence of diarrhea in urban - residential area of Islamabad, Pakistan (2017).
71. *Joakim Jörwall*: Quantitative model of Present and Future well-being in the EU-28: A spatial Multi-Criteria Evaluation of socioeconomic and climatic comfort factors (2017).
72. *Elin Haettner*: Energy Poverty in the Dublin Region: Modelling Geographies of Risk (2017).
73. *Harry Eriksson*: Geochemistry of stream plants and its statistical relations to soil- and bedrock geology, slope directions and till geochemistry. A GIS-analysis of small catchments in northern Sweden (2017).
74. *Daniel Gardevärn:* PPGIS and Public meetings – An evaluation of public participation methods for urban planning (2017).
75. *Kim Friberg:* Sensitivity Analysis and Calibration of Multi Energy Balance Land Surface Model Parameters (2017).
76. *Viktor Svanerud:* Taking the bus to the park? A study of accessibility to green areas in Gothenburg through different modes of transport (2017).

77. *Lisa-Gaye Greene*: Deadly Designs: The Impact of Road Design on Road Crash Patterns along Jamaica's North Coast Highway (2017).
78. *Katarina Jemec Parker*: Spatial and temporal analysis of fecal indicator bacteria concentrations in beach water in San Diego, California (2017).
79. *Angela Kabiru*: An Exploratory Study of Middle Stone Age and Later Stone Age Site Locations in Kenya's Central Rift Valley Using Landscape Analysis: A GIS Approach (2017).
80. *Kristean Björkmann*: Subjective Well-Being and Environment: A GIS-Based Analysis (2018).
81. *Williams Erhunmonmen Ojo*: Measuring spatial accessibility to healthcare for people living with HIV-AIDS in southern Nigeria (2018).
82. *Daniel Assefa*: Developing Data Extraction and Dynamic Data Visualization (Styling) Modules for Web GIS Risk Assessment System (WGRAS). (2018).
83. *Adela Nistora*: Inundation scenarios in a changing climate: assessing potential impacts of sea-level rise on the coast of South-East England (2018).
84. *Marc Seliger*: Thirsty landscapes - Investigating growing irrigation water consumption and potential conservation measures within Utah's largest master-planned community: Daybreak (2018).
85. *Luka Jovičić*: Spatial Data Harmonisation in Regional Context in Accordance with INSPIRE Implementing Rules (2018).
86. *Christina Kourdounouli*: Analysis of Urban Ecosystem Condition Indicators for the Large Urban Zones and City Cores in EU (2018).
87. *Jeremy Azzopardi*: Effect of distance measures and feature representations on distance-based accessibility measures (2018).
88. *Patrick Kabatha*: An open source web GIS tool for analysis and visualization of elephant GPS telemetry data, alongside environmental and anthropogenic variables (2018).
89. *Richard Alphonce Giliba*: Effects of Climate Change on Potential Geographical Distribution of Prunus africana (African cherry) in the Eastern Arc Mountain Forests of Tanzania (2018).
90. *Eiður Kristinn Eiðsson*: Transformation and linking of authoritative multi-scale geodata for the Semantic Web: A case study of Swedish national building data sets (2018).
91. *Niamh Harty*: HOP!: a PGIS and citizen science approach to monitoring the condition of upland paths (2018).
92. *José Estuardo Jara Alvear*: Solar photovoltaic potential to complement hydropower in Ecuador: A GIS-based framework of analysis (2018).
93. *Brendan O'Neill*: Multicriteria Site Suitability for Algal Biofuel Production Facilities (2018).
94. *Roman Spataru*: Spatial-temporal GIS analysis in public health – a case study of polio disease (2018).
95. *Alicja Miodońska*: Assessing evolution of ice caps in Suðurland, Iceland, in years 1986 - 2014, using multispectral satellite imagery (2019).
96. *Dennis Lindell Schettini*: A Spatial Analysis of Homicide Crime's Distribution and Association with Deprivation in Stockholm Between 2010-2017 (2019).

97. *Damiano Vesentini*: The Po Delta Biosphere Reserve: Management challenges and priorities deriving from anthropogenic pressure and sea level rise (2019).
98. *Emilie Arnesten*: Impacts of future sea level rise and high water on roads, railways and environmental objects: a GIS analysis of the potential effects of increasing sea levels and highest projected high water in Scania, Sweden (2019).
99. *Syed Muhammad Amir Raza*: Comparison of geospatial support in RDF stores: Evaluation for ICOS Carbon Portal metadata (2019).
100. *Hemin Tofiq*: Investigating the accuracy of Digital Elevation Models from UAV images in areas with low contrast: A sandy beach as a case study (2019).
101. *Evangelos Vafeiadis*: Exploring the distribution of accessibility by public transport using spatial analysis. A case study for retail concentrations and public hospitals in Athens (2019).