# Spatial accessibility to healthcare: an exploration

1. Abstract
2. Introduction
3. Research aims
4. Conceptual frameworks
   1. Healthcare and specific types of healthcare needs
   2. Healthcare accessibility
   3. Spatial accessibility
5. Data and methodology
   1. Data
      1. *Excel data*
      2. *Geographical data*
   2. Methodology
6. Results
   1. Descriptive results
   2. Model results

*6.2.1.Scenario 1*

*6.2.2.Scenario 2*

1. Conclusion
2. Limitations
3. References

# Abstract

Healthcare is one of the fundamental infrastructures that necessitates constant adjustment and improvements because it is strongly connected to the changing environmental socio-economic-demographic land use patterns. Access to healthcare is also one of the most important dimensions of spatial inequalities. However, spatial accessibility to healthcare is not a widely investigated research area. Therefore, this study aims to become an initiator and baseline for the conceptualization and operationalization of spatial healthcare accessibility by analyzing accessibility in Istanbul on the province level. The research uses the Rational Agent Access Model (RAAM) for operationalizing accessibility by taking the population as the measure of demand and the number of available doctors per healthcare facility as the measure of supply. The research analyzes two scenarios based on the travel time that people are willing to bear to reach a healthcare facility. The results are analyzed with Python via the Access package developed by James Saxon and colleagues. The data is prepared and visualized via the Quantum GIS (QGIS).

*Keywords: Spatial healthcare accessibility, RAAM, Istanbul*

# Introduction

In a city with more than 15 million residents, access to healthcare is becoming increasingly difficult. There are many reasons resulting in different types and degrees of inaccessibility that concern the whole nation such as deficiencies in the health infrastructure mainly denoted by the insufficient amount of specialist doctors and hospital beds compared to OECD averages (Gencelli et al., 2023); and concern Istanbul such as the uneven distribution of health facilities among districts and the provision of (quality) services predominantly by private facilities.

In the international literature, healthcare accessibility has been a widely investigated area. In the Turkish literature, on the other hand, this issue has been approached from the perspective of the existence of hospitals in certain areas. The Istanbul Municipality (*İstanbul Büyükşehir Belediyesi*) has prepared an index to measure the accessibility of health in walking distance on the neighborhood level. However, this research does not have any documentation to examine the data and method(s) used in index construction. So, this research will be a stepping stone and baseline in developing a healthcare accessibility measure on the province level in Istanbul.

Considering the fact that healthcare accessibility has been a critical issue in Istanbul which is a megacity with an ever-growing population, this research aims to conceptualize and operationalize spatial healthcare accessibility to pinpoint possible inequalities in the health infrastructure mainly due to the uneven placement of hospitals and differing travel costs.

Within the scope of this research, **accessibility** refers to the **fastest drive time** to hospitals assuming that people would be more inclined to use vehicles (ambulance, bus, taxi, personal automobile, etc.) and less crowded facilities to receive **primary** and **emergency care** from public or private healthcare facilities. The level of accessibility is measured on the province level using the neighborhoods as the **location(s) of demand**, neighborhood populations as the **measure of demand**, healthcare facilities as the **location(s) of supply**, and the number of specialist doctors in each facility as the **measure of supply**. **The Rational Agent Access Model (RAAM)** measures spatial healthcare accessibility using the PySAL developed by Saxon and colleagues (2020). Quantum GIS (QGIS) is used for visualization and calculating travel time cost matrix.

# Research aims

This research investigates how spatial healthcare accessibility in Istanbul can be computed and communicated. Thus, it is guided by the following research question:

- How to conceptualize and operationalize spatial healthcare accessibility in Istanbul on the province level?

Spatial accessibility is a difficult area to operationalize due to the limitations with data, errors with working approximate geographical values, and the potential values of demand as well as supply. Therefore, this research will be explorative in nature.

# Conceptual frameworks

## Healthcare and specific types of healthcare needs

In a nutshell, healthcare refers to all the services that aim to diagnose, prevent, and treat mental or physical illnesses or impairments. Healthcare takes many forms and durations such as primary and emergency based on patient needs and facility capabilities.

Primary care refers to all the circumstances that necessitate treatment for acute illnesses such as infections or chronic conditions such as diabetes. Emergency care, on the other hand, refers to all the circumstances that necessitate treatment for sudden and most possibly fatal injuries such as trauma or syndromes such as heart attacks.

As with many other studies in the literature, this research limits its healthcare scope to primary and emergency care to focus on relatively urgent treatment needs.

## Healthcare accessibility

Healthcare accessibility has been a widely examined field due to its impact on well-being, policy, and infrastructure design. The patterns of accessibility can be regarded as different forms of opportunities to reach and benefit from services with diverse quality.

In the most general perspective, healthcare accessibility takes two forms. “Aspatial and spatial” (Luo & Wang, 2005, 131). Aspatial accessibility refers to the socio-economic-demographic-cultural factors (e.g. income, age, gender, ethnicity, etc.) that affect individuals’ or households’ resources and capabilities to reach quality treatment (Luo & Wang, 2005, 131). Spatial accessibility refers to the geographic (e.g. distance, topology, etc.) and infrastructural (e.g. road networks, availability of public transportation lines and points, etc.) factors that constrain peoples’ choices limited by travel time and cost (Luo & Wang, 2005, 131).

## Spatial accessibility

Spatial accessibility refers to the ease through which people can reach and benefit from healthcare services based on their geographic location and cost of travel. In the most general perspective, spatial accessibility involves **demand** (e.g. patients in terms of population) and **supply** (e.g. number of hospitals) as the main variables. Depending on their complexity and objectives, some studies use “dynamic variables” (Park & Goldberg, 2021, p.2) that involve travel time according to different “travel modes” (Park & Goldberg, 2021, p.2) such as public transportation, walking, biking, etc.

The demand can be defined by the actual utilization of the services (“revealed accessibility”) (Luo & Wang, 2003, p.865) or its potential for services (“potential accessibility”) (Luo & Wang, 2003, p.865). The supply can be defined by the number of healthcare facilities, specialists, or treatments available in a given location. Defining supply is heavily determined by the desired level of detail and availability of data. The number of specialists is generally used within the literature. The reachability of demand locations to supply locations is usually confined to a “catchment area” (Saxon et al., 2020, p.6) which defines “the maximum range” demand can travel to supply locations (Saxon et al., 2020, p.6). Different means of travel determine the reasonable distance and time a person can take to reach facilities. Usually, studies focus on one means of travel (e.g. driving or walking). However, recent studies incorporate “multimodal transportation modes” (Zhou et al., 2023, p.1) that account for differing costs for different groups and the “attractiveness” (García-Albertos et al., 2019, p. 295) of supply locations. For example, Ahmed and colleagues (2019) as well as Mutono and colleagues (Mutono et al., 2022) account for different traffic scenarios and congestion.

In terms of workflow, defining and measuring spatial accessibility is straightforward. First, supply and demand values and locations must be defined. Second, either the distance or travel time must be measured. Third, spatial access models must be utilized to operationalize accessibility (Saxon et al., 2019). There are six fundamental access models within the literature, each differing according to factors they consider in operationalizing reachability and the complexity of their calculations. From the least to most complex models, they can be listed as follows (Saxon et al., 2020, pp. 7-10):

* “**Gravity model**”: Created by Harris (1954), this model is based on the summation of all supply locations and weighting them according to a “distance decay function” (Saxon et al., 2020, p. 7). The distance decay function can be considered a penalty function that reduces the attractiveness of a supply point with increasing distance or travel costs.
* “**Floating catchment area (FCA)**”: Created by Huff (1963), this model is based on the ratio of total supply points within a catchment area to its demand (population) (Saxon et al., 2020, p. 7). The catchment area, in this context, refers to an area within which people can easily reach services.

The processes of the FCA model constitute a benchmark that can be adjusted or built upon. The very first step in the FCA model is to define a catchment area which defines a specific area between demand and supply locations based on a travel time threshold such as 30 minutes of reachability. Then, within each catchment area, the ratio between supply locations and demand is calculated. Finally, the ratios are translated into accessibility scores defined according to each catchment area. Simple as it is, this model does not account for the distance decay function. However, more advanced versions of FCA apply penalties for increasing distances and travel time (Saxon et al., 2020, p. 8).

* “**Two-step** **floating catchment area (2SFCA)**”: This is an upgraded variant of FCA that accounts for the distance decay function (Saxon et al., 2020, p. 8).
* “**Three-step** **floating catchment area (3SFCA)**”: This is the most advanced variant of FCA. It adds weights to the distance decay function (Saxon et al., 2020, p. 8), allowing it to account for varying capabilities for reaching supply locations.
* “**Rational access agent model (RAAM)**”: This is different than FCA-based models. It is based on the optimization of travel based on the time it takes to reach a supply location and “congestion” (Saxon et al., 2020, p. 9). The model defines the distance function by calculating the catchment areas. Then, it divides catchment areas according to travel times (Saxon et al., 2020). It assumes that agents or individuals choose the shortest and least occupied supply or service locations. However, they are constrained by the degree of occupation of the supply locations (congestion) and travel time. This allows RAAM to assume and incorporate “competition” (Saxon & Snow, 2020, p.5) among service locations. Thus, this model better incorporates the possible choices regarding which facility to use. The creators of this model, Saxon and Snow, argue that this model is differentiated FCA methods by accounting for the “demand elasticity to supply” (Saxon & Snow, 2020, p.5). This means that the RAAM does not necessarily assume that demand increases with each increase in service supply (e.g. doctors) (Saxon & Snow, 2020, p.6). Rather, it assumes that “ if one location costs twice as much for a patient as another, it is completely avoided—not utilized half as much” (Saxon & Snow, 2020, p.7).

This research takes neighborhood centroids in Istanbul as the **demand** and hospital locations as the **supply** **locations**. The **demand values** are the neighborhood populations fixated on the centroids. The **supply values** are the number of available specialist doctors in each healthcare facility in each province. The **travel cost** is based on the driving road network in Istanbul considering average and maximum driving speed. These independent variables are applied to the RAAM established by Saxon and Snow (2020) and practicalized by Saxon and colleagues as a Python package called “Access” (Saxon et al., 2020, p. 10). The variables and the usage of the Python package will be detailed in the following chapters.

# Data and methodology

## Data

### Excel data

The Istanbul Municipality open data portal and the Turkish Statistical Institute (TÜİK) database are the main data sources.

*İstanbul Sağlık Kurum ve Kuruluşları Verileri* (Istanbul Health Institutions and Organizations Data) gathered from the IBB open data portal is the main dataset that provides healthcare facility name (*sağlık tesisi adı*), sub-category (*alt kategori*), province name (*ilçe adı*), neighborhood name (*mahalle adı*), address (*adres*), longitude, and latitude as variables. **Healthcare sub-category, province name, longitude, and latitude** are the variables of interest within the scope of this research. Also, **the number of specialist doctors** in each establishment is an important variable not included in the IBB open data portal. The number of available specialist doctors data in each facility is gathered by crawling the websites of the establishments, which is a very long process.

The sub-category variable 45 items including **hospitals** (private, university, public, training and research, city), **municipal health centers** (*belediye sağlık merkezi*), **private polyclinics** (*poliklinik özel*), and **private medical centers** (*tıp merkezi özel*). Within the scope of this research, these items will be subsetted. The main reason behind this decision is to include the easily identifiable centers on the map that include information about the number of specialist doctors working, and facilities where people can get primary and emergency treatment. This excludes beauty clinics, individual establishments, pharmacies, labs, vets, etc. Establishments like family health centers (*aile sağlığı merkezi*), province polyclinics (*semt polikliniği*), dialysis centers, and community health centers (*toplum sağlığı merkezi*) are excluded because their websites do not include information about the number of specialist doctors included. Moreover, since this is a small-scope research, establishments such as dental care centers are excluded because this research focuses on primary and emergency treatment.

The whole dataset includes 20469 items. The subsetted dataset includes over 800 items. After the subsetting, incorporating the centers that include the number of available specialist doctor information, and removing duplicate values, 426 items left for the study. Thus, this constitutes the supply dataset[[1]](#footnote-1).

From the TÜİK database, the Address-Based Population Registration System (*Adrese Dayalı Nüfus Kayıt Sistemi*) dataset was accessed for the year 2022. From this dataset, total population data in 2022 at the neighborhood and province levels were gathered. The neighborhood population data constitutes the demand dataset[[2]](#footnote-2). The province population dataset is used for creating composite accessibility scores on the province level which will be explained in the following chapters.

### Geographical data

OSM Standard map, gathered from the Quick Map Services plugin in QGIS, is selected as the base map. The administrative boundaries for neighborhoods and provinces in Istanbul are gathered through the *Overpassturbo API* which gathers data from the Open Street Map database as shapefiles in terms of polygon layers.

With the join operation in QGIS, the population data are appended to the polygon layers. Before the join operation, polygon boundaries regarding Adalar are removed because there is no hospital location data in the subsetted dataset used in this research. After the adjustments[[3]](#footnote-3), each polygon layer involves information regarding ID, name (province name and neighborhood name), and population. Thus, 950 neighborhood (originally 963 for Istanbul) and 38 province items are used for the research.

*Mean Coordinates* algorithm in QGIS was used to identify the mean centers (centroids) within each neighborhood boundary. Then, the neighborhood population data is joined to the centroids to create a demand dataset.

The main roads (*ana arterler*) data is gathered from the Transportation and Coordination Office (*Ulaşım ve Koordinasyon Müdürlüğü*). The main roads data is cleaned to create a whole network layer without disconnected islands. This process is conducted by using the *Disconnected Islands plugin* in QGIS to identify and group the disconnected road lines in the dataset. Then, deleting them from the dataset. After cleaning the data, name, road type, and road length (km) remained as the main variables. To create a travel cost dataset, a travel time variable was created by dividing the road length by average speed. The average speed is determined by taking the average of the hourly traffic density dataset (*saatlik trafik yoğunluk verisi*) for the year 2022[[4]](#footnote-4), which is 58 km/h.

## Methodology

As mentioned in the previous section. This study uses the RAAM which takes demand, supply, and travel time cost as its main functions. Within the scope of this research, the demand is determined as the population of each neighborhood alongside their locations, and the supply is determined as the number of available specialist doctors in each healthcare facility. Moreover, the travel time cost is determined with the *QNEAT3 plugin* via its *OD Matrix form Layers as Table (m:n)* algorithm. The OD Matrix algorithm takes the main roads dataset as the *network layer*, the demand dataset as the *from-point layer*, the supply dataset as the *to-point layer, fastest past* as the *optimization criterion, entry cost calculation* as ellipsoidal, *default direction* as *both directions, default speed* as58 km/h, and 50[[5]](#footnote-5) as the topology tolerance. The model assumes that people in a catchment area can reach healthcare facilities at any point (Saxon et al., 2020). Moreover, they prefer closer facilities with less congestion (Saxon et al., 2019).

PySAL Access package developed by Saxon and colleagues (2021) is used for preparing the datasets for analysis while utilizing RAAM, and scoring the values created via the RAAM function (Saxon et al., 2019). The results are visualized via QGIS.

The PySAL package[[6]](#footnote-6) is initialized by importing it as *Access*. Pandas, geopandas, and matplotlib are also imported to prepare the datasets for analysis. After importing the demand and supply datasets, the Access function is initialized to make the datasets understandable for the RAAM function. The RAAM function (*Access.raam*) (Saxon et al., 2019) is initialized by defining the tau value (“maximum willingness to travel”) (Saxon et al., 2021, p. 18), which is 60 minutes for scenario 1 and 90 for scenario 2 (will be explained in the following chapters). The scores function (*Access.score*) (Saxon et al., 2019) is initialized by converting RAAM results to a dictionary.

# Results

## Descriptive results

The maps below illustrate the number of available doctors and the population according to province[[7]](#footnote-7).

A map of turkey with blue and white colors

Description automatically generated

*Figure 1: A map depicting the number of available specialist doctors in each province*

The map illustrates there are more doctors available around Beşiktaş, Bakırköy, Kadıköy, and Ataşehir. The table results show that Şişli, Üsküdar, Kadıköy, Sarıyer, Küçükçekmece, Fatih, Bakırköy, Maltepe, Pendik, and Ataşehir are the top ten provinces with the highest number of available doctors.

A map of turkey with different colored areas

Description automatically generated

*Figure 2: A map depicting the population counts in each province*

The map illustrates that around Bakırköy on the European side and around Kadıköy are the most populous areas. The table results show that Şile, Çatalca, Beşiktaş, Silivri, Beyoğlu, Bakırköy, Beykoz, Bayrampaşa, Şişli, and Büyükçekmece are the top ten provinces with the lowest population count according to the 2022 Address-Based Population Registration System census results. This could indicate that congestion is lesser in these areas.

## Model results

### Scenario 1

The scores function in the PySAL Access package (2019) takes the RAAM results as inputs and creates score values based on aggregating the distance weights within catchment areas (Saxon et al., 2020). Scenario one assumes that people would not like to travel for more than 60 minutes to receive treatment. Therefore, the “tau” value (“ maximum willingness to travel”) (Saxon et al., 2020, p. 18) in the RAAM function is defined as 60. This value acts as a distance decay function to penalize travel times more than this threshold (Saxon & Snow, 2020, p. 13).

In the RAAM analysis, the smaller scores indicate greater accessibility contrary to the other models in the PySAL package. To create scores on the province level, several steps were followed. First, the scores were multiplied by the neighborhood populations to create weights. Then, the weights were divided by the total population (15,891,261) to create composite scores for each province[[8]](#footnote-8). The inverse relationship between the scores and the degree of accessibility applies here also since RAAM scores are utilized to establish these composite scores.

A map of turkey with green and white text

Description automatically generated

*Figure 3: A map depicting the composite scores for scenario 1 per province*

Beşiktaş, Bakırköy, Bayrampaşa, Beyoğlu, Büyükçekmece, Esenler, Maltepe, Güngören, Silivri, and Kadıköy are the top ten provinces with the highest accessibility to healthcare facilities.

### Scenario 2

Considering Istanbul’s busy traffic environment, the second scenario assumes that people would not like to travel for more than 90 minutes to receive treatment. Therefore, the “tau” value (“maximum willingness to travel”) (Saxon et al., 2020, p. 18) in the RAAM function is defined as 90. The same process was followed in creating the composite scores.

A map of turkey with green and white text

Description automatically generated

*Figure 4: A map depicting the composite scores for scenario 2*

Beşiktaş, Bayrampaşa, Silivri, Beyoğlu, Şile, Büyükçekmece, Bakırköy, Çatalca, Zeytinburnu, and Maltepe are the top ten provinces with the highest accessibility to healthcare facilities. The inverse relationship between the scores and the degree of accessibility applies here also since RAAM scores are utilized to establish these composite scores.

Contrary to the presumption, the two results do not differ much when the maximum travel willingness variable (tau) is adjusted for different scenarios. The only notable fact is that seven provinces (Beşiktaş, Bakırköy, Maltepe, Bayrampaşa, Silivri, Beyoğlu, and Büyükçekmece) remain in the top ten. This indicates that the model is not very sensitive to the willingness to travel. Values may need to be adjusted more dramatically which may not be feasible.

When the composite scores are correlated with the number of available doctors in a province and the province population, for each scenario the following results occur.

|  |  |  |
| --- | --- | --- |
|  | **Scores by number of doctors** | **Scores by population** |
| **Scenario 1** | 0,15076234 | 0,544354592 |
| **Scenario 2** | 0,185387753 | 0,706893651 |

*Figure 5: A table depicting the correlations between the number of doctors, population, and composite scores according to each scenario*

# It seems that population and the composite scores are positively correlated. This result is reasonable in the sense that population, as a measure of demand, is also one of the functions of congestion. When the willingness to travel time is increased, congestion in the nearest healthcare facilities may decrease because people can travel longer distances. Thus, the composite scores or accessibility seem to increase dramatically for the provinces with the lowest populations (notably Şile, Çatalca, and Silivri) in this scenario.

# Conclusion

From the data processing to the analysis steps both on Python and QGIS, this research has been a long and educational process in understanding the concept of spatial accessibility and its measures.

The results of RAAM analyses in both scenarios produced results that are both interesting and somewhat contrary to the belief held before beginning this study. For example, one would expect healthcare accessibility higher in places with a higher number of specialist doctors and a lower population. Although this study does not prove this belief wrong, there seems no consistently strong relationship between the composite scores and the number of doctors as well as the population. Of course, the results of this study heavily depend on the limitations and errors in the process and analysis of data which might have confounded or misguided the results. Therefore, there seems to be misconnections between the descriptive results and model results.

Despite these unpredicted mishaps, this research can still be considered a stepping stone to reveal the processes, possible challenges, and basic problems that can be experienced in dealing with the measurement of spatial accessibility. However, to become a credible source and baseline, this research needs to address its limitations, compare its results, and find ways to validate its relations with the descriptive results.

# Limitations

This research involves many limitations with the datasets and analysis.

Limitations with the datasets,

* The hospital locations dataset is quite limited after subsetting
* The number of available doctors data was mostly gathered manually
* Roads network data was heavily reduced to make it ready for OD Matrix
* The time travel cost matrix does not incorporate real-life traffic information

Limitations with the analysis,

* Using potential demand and supply may create problems in real-life applications or interpretations,
* Centroids do not indicate a true demand location, defining them on the neighborhood level only allows more detailed results for the province-level
* Using road networks makes the analysis focus on certain groups. For example, those who can only use public transportation may be excluded
* The process of creating composite scores may be wrong
* Driving time does not necessarily include public transportation use. Therefore, using driving time in measuring accessibility may not be the most feasible method especially for incorporating socio-economic inequalities as a dimension within research.

# References

Ahmed, S., Adams, A. M., Islam, R., Hasan, S. M., & Panciera, R. (2019). Impact of traffic variability on geographic accessibility to 24/7 emergency healthcare for the urban poor: A GIS study in Dhaka, Bangladesh. *PLOS ONE*, *14*(9), e0222488. https://doi.org/10.1371/journal.pone.0222488

García-Albertos, P., Picornell, M., Salas-Olmedo, M. H., & Gutiérrez, J. (2019). Exploring the potential of mobile phone records and online route planners for dynamic accessibility analysis. *Transportation Research Part A: Policy and Practice*, *125*, 294–307. https://doi.org/10.1016/j.tra.2018.02.008

Gencelli, A., Yaşar, G., & Demirkol, Ö. (2023). (report). *Bir insan hakkı olarak sağlığa erişim: İstanbul’da sağlık altyapısının sektörel bazda incelenmesi*. Istanbul Planlama Ajansı

Harris, C. D. (1954). The market as a factor in the localization of industry in the United States. *Annals of the Association of American Geographers*, *44*(4), 315–348. https://doi.org/10.1080/00045605409352140

Huff, D. L. (1963). A probabilistic analysis of shopping center trade areas. *Land Economics*, *39*(1), 81. https://doi.org/10.2307/3144521

Luo, W., & Wang, F. (2005). Assessing spatial and nonspatial factors for healthcare access: Towards an integrated approach to defining health professional shortage areas. *Health &amp; Place*, *11*(2), 131–146. https://doi.org/10.1016/j.healthplace.2004.02.003

Luo, W., & Wang, F. (2003). Measures of spatial accessibility to health care in a GIS environment: Synthesis and a Case Study in the Chicago Region. *Environment and Planning B: Planning and Design*, *30*(6), 865–884. https://doi.org/10.1068/b29120

Mutono, N., Wright, J. A., Mutunga, M., Mutembei, H., & Thumbi, S. M. (2022). Impact of traffic congestion on spatial access to healthcare services in Nairobi. *Frontiers in Health Services*, *2*, 788173. https://doi.org/10.3389/frhs.2022.788173

Park, J., & Goldberg, D. W. (2021). A review of recent spatial accessibility studies that benefitted from advanced geospatial information: Multimodal transportation and spatiotemporal disaggregation. *ISPRS International Journal of Geo-Information*, *10*(8), 532. https://doi.org/10.3390/ijgi10080532

Saxon, J., Koschinsky, J., Acosta, K., Anguiano, Snow D., Atlas Y., Wang B. (2019). (presentation). Making spatial access metrics more accessible: Documentation for the PySAL spatial access package

Saxon, J., Koschinsky, J., Acosta, K., Anguiano, V., Anselin, L., & Rey, S. (2020). An open software environment to make spatial access metrics more accessible. *Journal of Computational Social Science*, *5*(1), 265–284. https://doi.org/10.1007/s42001-021-00126-8

Saxon, J., & Snow, D. (2020). A rational agent model for the spatial accessibility of primary health care. *Annals of the American Association of Geographers*, *110*(1), 205–222. https://doi.org/10.1080/24694452.2019.1629870

Zhou, X., Yuan, L., Wu, C., Yu, Z., & Lei, W. (2023). Measuring spatiotemporal accessibility to healthcare with multimodal transport modes in the dynamic traffic environment. *Open Geosciences*, *15*(1), 20220461. https://doi.org/10.1515/geo-2022-0461

1. Please refer to the *Healthcare\_facilities* dataset on GitHub. [↑](#footnote-ref-1)
2. Please refer to the *Population* dataset on GitHub to examine both neighborhood and province populations. [↑](#footnote-ref-2)
3. There were inconsistencies regarding the number of neighborhoods in the polygon shapefile and population data. Therefore, the population data is adjusted by deleting some of the rows, changing the neighborhood names, and matching the IDs according to the shapefile. The reason for adjustment according to the shapefile is that it is more difficult to create and modify new polygons. This introduces an error as well as a limitation to the study. [↑](#footnote-ref-3)
4. The hourly traffic density data exists in monthly datasets. [↑](#footnote-ref-4)
5. This value is used as a proxy for the maximum driving speed in the city. [↑](#footnote-ref-5)
6. Please refer to the *Analysis* file on GitHub for a detailed examination. [↑](#footnote-ref-6)
7. Please refer to the *Healthcare\_facilities (Sheet 2)* dataset on GitHub to examine the number of available doctors per province. Please refer to the *Population (Population)* dataset to examine total population counts per province. [↑](#footnote-ref-7)
8. Please refer to the S1 and S2 datasets to examine the composite scores. [↑](#footnote-ref-8)