Final Report CUTR-2020-06

Deficiencies in Public Transit Accessibility of Healthcare Facilities in Chicago

Project Number
79063-02

Prepared For
National Center for Transit Research
USF Center for Urban Transportation Research

Prepared by
Paul Metaxatos and P.S. Sriraj
with assistance from
Sneha Dattatraya Shet and Kewal Shah
Urban Transportation Center (M/C 357)
University of Illinois at Chicago

September 2020

ACKNOWLEDGEMENT

This study is based on research conducted under the sponsorship of the Metropolitan Transportation Support Initiative (METSI) from the Illinois Department of Transportation. Additional support has been provided by Grant: DTRT12-G-UTC18 from the U.S. Department of Transportation.

The contents of this report reflect the view of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Illinois Department of Transportation, or the U.S. Department of Transportation. This report does not constitute a standard, specification, or regulation.

SUMMARY

Public transportation has the potential to impact the accessibility of individuals to healthcare facilities and affect the quality of life and the livability of a community. It is critical, therefore, to understand the service efficiency of public transportation as a mode of access to healthcare facilities. In this study, we develop performance measures/indices that will reflect on both the supply (availability) and the demand side (accessibility). The indices are then applied in the Chicago area.

While there are different indicators for measuring availability, for the purposes of this research we focus on three indicators - the frequency of service, hours of service, and the service coverage. These measures are aggregated into an index for transit availability. The index is a measurement of the percent of person-minutes served. For a given geographic area, the index multiplies the percent of area served by transit by the percent of an hour that a station or stop is served (assuming a five-minute wait time) by the percent of a day that the area is served by transit. The index range is between zero and one for each census block group.

The public transit accessibility serves as a proxy for the travel demand at (or near) the locations of healthcare facilities using public transit. Among the several approaches to measuring transportation accessibility this research uses the generalized gravity model framework with public data in the Chicago region to develop a public transit accessibility index. The index measures the aggregate peak-period public transit accessibility potential to the locations of healthcare facilities for each residential zone in the Chicago area. The neighborhood with the highest such accessibility measure is the one with best public transit access, as measured by friction factors to all healthcare facilities in the region.

Both the indices ranging from zero to 100% are then split into four groups using the median values of the two indices: high-high group, high-low group, low-high group, and low-low group. Census Block groups in the low-low category are deemed deficient in both public transit accessibility and availability of healthcare facilities in the Chicago area, and several policy interventions are proposed to improve and address the situation. In this regard, this project has successfully demonstrated techniques that could add to the battery of tools available to study public transportation barriers to healthcare.

TABLE OF CONTENTS

1. INTRODUCTION

The welfare of the transportation disadvantaged may be at risk when access to healthcare for routine physicals, medical treatment and follow-ups is unnecessarily inhibited by poor public transit access (Rittner and Kirk, 1995; Glaeser, et al., 2008; Silver, et al. 2012; Syed, et al. 2013; Graham, et al. 2015). In this regard, transit planners and community stakeholders could be interested to identify such origin neighborhood clusters and destination hospitals/clinics, as well as specific types of spatial separation that tend to impede or enhance the likelihood of interactions between patients or care givers/care takers and treatment facilities. This exercise along with taking stock of available transit options could inform stakeholders of potential deficiencies.

In this study, deficiencies in public transit access to healthcare facilities will be determined by comparing two indices related to the demand for and supply of public transportation used as a mode connecting residential origin zones and healthcare facilities located in trip destination zones: (a) the first index relates to the demand for public transportation and measures the average accessibility potential of each residential zone in the Chicago metropolitan area to healthcare facilities in destination zones; and (b) the second index relates to the supply of public transportation and measures the average availability of transit at bus stops, rail stations, route corridors or system wide in neighborhoods with healthcare facilities. Such areas measuring at the low end of both accessibility availability indices will be deemed deficient and in need of policy intervention to improve public transit access to healthcare facilities.

The study will be using methodologies well established in transportation planning practice, and data that are routinely available in metro areas with transit presence. The techniques described in this study will likely add to the battery of tools available to planning authorities, researchers and practitioners. It would only be beneficial if similar studies were carried out in such areas. As the U.S. population ages, the number of people who rely on public transportation to access health services is expected to grow substantially.

2. PUBLIC TRANSPORTATION AND ACCESS TO HEALTHCARE

Only a few studies have examined the relationship between public transportation and access to healthcare and depending on the population studied and information available on public transport use, they have found differing associations. Rask et al. (1994) studied obstacles to care for 3,897 urban, low socioeconomic status (SES) adults in Atlanta and found that walking or using public transportation to receive medical care was an independent predictor of not having a regular source of care (Odds Ratio, OR 1.44). Patients who did not use private transportation were also more likely to delay care (OR 1.45).

Flores et al. (1998) studied 203 children's caretakers and found that 21 % of inner-city children faced transportation barriers to timely health care. Of these, 62 % cited lack of a car as the specific barrier, which exceeded other reasons including excessive distance, expense, or inconvenience of public transportation.

One study investigated transit accessibility to health care by either public transit or by foot in various low-income counties in the Bay Area (2002). Results revealed that transit accessibility to a hospital, defined as getting to a hospital or clinic in 30 min or less by public transit or ½ mile by foot, varied from 0 to 28 %. Additionally, 55 % of missed appointments or late arrivals were due to transportation problems. Similarly, in a study of 698 low-income adult patients Silver et al. (2012) found that 25 % of missed appointments/rescheduling needs were due to transportation problems and bus users were twice as likely to miss their appointments compared to car users.

In a survey of adults in rural and small urban areas in the U.S. Great Plains states Mattson (2011) found most respondents used private vehicles to get to physicians' offices, but 5% used public transport, 3% used volunteer driver services, and 2% used services provided by human service agencies. Use of these alternatives to private vehicles increased the probability of having difficulties traveling to care, but this relationship was not statistically significant. Additionally, 35% of respondents who did not have access to public transit indicated that they would use public transit to travel to care if it were available.

Graham, et al. (2015) examined public transportation travel time barriers to mammography facilities for women without access to a private vehicle and for women with especially long public transit times in six urban areas. Although only 2% of women had both characteristics (transit marginalized) they comprised a large number of women across the 6 large urban areas. While black women were less likely to have private vehicle access, and both Hispanic and black women were more likely to be transit marginalized, this outcome varied by urban area. White women constituted the largest number of transit marginalized. Although Hispanics are less segregated than other population groups overall, many live in isolated enclaves that are less favorably situated with respect to public transportation.

Ruggiano, et al. (2017) argue that for disadvantaged populations such as older adults, minorities, low-income individuals, and individuals with disabilities, not having reliable transportation can create a barrier to accessing and engaging in health services. The unavailability of reliable transportation makes it difficult for disadvantaged groups to manage chronic health care conditions and can negatively affect health outcomes. For example, not having reliable transportation can lead to missing health care appointments, postponing treatment, and difficulty visiting a pharmacy to pick up medications and subsequently results in unmet health care needs. The study examined perceptions of transportation and health self-management among older adults with chronic conditions (i.e., chronic illnesses and disabilities) in central and south Florida. Overall, the findings aligned with prior assertions that transportation is necessary for disease/disability management activities (e.g., transportation to doctor's offices, picking up medication) and that a lack of reliable transportation creates challenges to disease management, which can negatively affect health and result in unmet health care needs. Respondents reported difficulty using public transit due to difficulty getting to stops, increased travel time, and discomfort while traveling on the transportation system.

3. METHODOLOGY

The methodology to assess deficiencies in public transit accessibility of healthcare facilities in Chicago consists of separate measurements of public transit accessibility of and public transit availability at the locations of the healthcare facilities. The public transit accessibility serves as a proxy for the travel demand at (or near) the locations of healthcare facilities using public transit. The public transit availability serves as a proxy for the supply of public transit at (or near) the locations of healthcare facilities. We begin the presentation with the measurement of public transit accessibility.

3.1 MEASUREMENT OF PUBLIC TRANSIT ACCESSIBILITY

The measurement of public transit accessibility required data about the location of healthcare facilities as well as data about the separation between residential areas and healthcare facilities using public transit. Then we computed the accessibility of each residential area to each healthcare facility location by estimating generalized gravity models of spatial interaction (Sen and Smith, 1995). This framework of analysis is very flexible since it allows consideration of residential neighborhood characteristics associated with accessing particular healthcare facilities, as well as characteristics of healthcare facilities (Lowe and Sen, 1996).

There are four main approaches to measuring accessibility: opportunity-based, gravity-type, utility-based, and space-time approaches (Kwan, 1998; Liu and Zhu, 2004; Geurs and van Wee, 2004; Benenson et al., 2016; Lee and Miller, 2018). Among these approaches, traditional gravity models have been used to measure potential spatial access to healthcare facilities in previous studies (Joseph and Bantock, 1982; Luo and Wang, 2003; Minosha et al., 2008;

Schuurman et al., 2010; Crooks and Schuurman, 2012). In this paper, we demonstrate the generalized gravity model framework with public data in the Chicago region. The methodology is readily transferable to other regions in the country as discussed below.

Public Transit Demand Data

We obtained public transit demand data from an inventory available by the Chicago Metropolitan Agency for Planning (CMAP) as shown in the Appendix. The agency publishes data that are prepared for or output from the regional model used for the air quality conformity analysis in the Chicago region. We used the following data from the third quarter 2019 Air Quality Conformity Analysis:

- Study Area and Zonal Geography: In this paper, we have focused on the six-county service area of the Regional Transportation Authority (RTA) comprised by Cook County, DuPage County, Lake County, Kane County, McHenry County, and Will County. This service area is covered by 2,926 modeling zones or traffic analysis zones (TAZs). The size of TAZs range from one quarter-mile by one quarter-mile, and one half-mile by one half-mile in the Chicago Central Area to one square mile outside of the Central Area, and to four, nine, and thirty-six square-mile zones the farther away from the City of Chicago.
- Public transit trip table: We used an origin-destination trip table estimated for the morning peak hour of traffic (7:00 a.m. to 9:00 a.m.) with 682,152 home-based work transit person trips. Moreover, the trip table had 94 origin and destination zones with no trips going from/to there. These origin and destination were subsequently eliminated, as explained below, and we ended up with a 2,832 by 2,832 trip table. Note that, initially, the trip table was in production-attraction format in which all trips (i.e., the trips from home to work and the trips back from work to home) are represented as starting at the production (home) end. As a result, the trip table was converted to an origin-destination format by adding its transpose and taking one-half of the sum.
- Transit separation measures (skims): Several transit origin-destination skim/impedance matrices for the same peak period of travel were considered as separation measures as follows: (a) an in-vehicle travel time table (in minutes); (b) a table (in minutes) that included walking time between transfers on a path plus the egress walk time; and (c) a table representing twice the total wait time (in minutes), and divided by two to obtain the total wait time. Obviously, all separation measures considered have the same 2,832 by 2,832 dimensionality as the trip table above.

There are tradeoffs in using the data above. On one hand, (a) the data are publicly available from CMAP and are updated biannually; and (b) the separation measures reflect, overall, realistic traffic conditions accounting for the competition of highway and transit modes (see CMAP's travel model documentation https://www.cmap.illinois.gov/documents/10180/911391/FINAL+Travel+Demand+Model+Documentation+Appendix.pdf/f3b1322c-2e60-2513-720f-38ee68b799d1 – accessed 11/25/19). On the other hand, in the absence of more information about the travel behavior of prospective clients to the study area healthcare facilities, this particular group of transit users is thought of as a segment of the overall transit demand during the morning peak period.

3.1.1 Healthcare Facilities Data

We subset the location of the healthcare facilities for the study area from an automated inventory of the proper names and locations of hospitals located throughout the United States and its territories. The inventory was last updated in February 2020 and is made available as a layer package by ESRI (https://www.arcgis.com/home/item.html?id=f114757725a24d8d9ce203f61eaf8f75 – accessed 4/27/20). The inventory for the study area includes 206 healthcare facilities including medical centers, trauma centers, healthcare centers, hospitals, clinics, and immediate care centers. The location of each healthcare facility was then identified within its nearest TAZ centroid. As a result, the location of each healthcare would be assigned the same accessibility index score as its associated TAZ.

3.1.2 Estimation of Public Transit Accessibility

Estimation of public transit accessibility required to run a program implemented in Python (see Appendix for details). The procedure estimates a flow of trips, T_{ij}, between an origin zone $i = 1, \dots, I$ and a destination zone $j = 1, \dots, J$ given an observed trip table N_{ij} and various impedance factors $c_{ij}^{(k)}, k = 1, \dots K$ (e.g., travel time, travel distance, etc.) between i and j. In fact, the procedure estimates the gravity model $T_{ij} = A_i B_j F_{ij}$, with $F_{ij} = e^{\sum_1^K \theta_k c_{ij}^{(k)}}$. The parameters A_i and B_j are, respectively, origin specific factors for each zone i, and destination specific factors for each zone j. Maximum Likelihood estimation of the parameters A_i , $i = 1, \dots, I$, $B_j, j = 1, \dots, J$ and $\theta_k, k = 1, \dots, K$ is obtained by solving the following system of $I + J + K$ linear equations: $\sum_j T_{ij} = \sum_j N_{ij}\ \ \forall i$, $\sum_i T_{ij} = \sum_i N_{ij}\ \ \forall j$, and $\sum_i \sum_j c_{ij}^{(k)} T_{ij} = \sum_i \sum_j c_{ij}^{(k)} N_{ij}\ \ \forall k$ (Metaxatos, 2004).

In the end, public transit accessibility, PTA_i, for each origin zone i, is defined by the formula $PTA_i = \sum_j B_j F_{ij}$. The index describes all opportunities available to a resident of origin zone i to access the location of any healthcare facility in the area. As more healthcare facilities are encountered in destination zones j, or as the separation between a residential zone i and destination zones j is decreased, the accessibility at each residential zone i will increase. This is the well-known accessibility definition first proposed by Hansen (1959) in a different planning context.

The aforementioned framework can be further extended to other times-of-day periods, and generalized to accommodate origin-specific factors associated with accessing particular healthcare facilities (e.g. demographic and socioeconomic mix, housing type, transit availability, car availability, etc.), destination-specific factors related to healthcare facilities (e.g., familiarity with the area, facility type, consultation hours, etc.), as well as factors solely dependent on the separation between residential neighborhoods and healthcare facilities (e.g., composite costs, social distance, etc.). Moreover, the methodology is readily transferable to other regions with substantial public transit presence and available data.

3.1.3 Computational Issues

The solution algorithm uses an (iterative) hill climbing method called Modified Scoring procedure as described by Yun and Sen (1994). At each iteration of the modified scoring procedure, the method obtains a new set of θ_k parameters once the parameters A_i and B_j are estimated by an iterative proportional fitting procedure (also known as Deming-Stephan-Furness or DSF procedure, the RAS Method, or two-dimensional balancing) as described elsewhere (Metaxatos, 2004).

If an origin zone i sends no trips to any other zone, i.e., $N_{i+} = \sum_j N_{ij} = 0$ then $N_{ij} = 0$ for each value of j, and such a row of the origin-destination matrix would play no role in the DSF procedure. Similarly, if a destination zone j receives no trips from any other zone, i.e., $N_{+j} = \sum_i N_{ij} = 0$ then $N_{ij} = 0$ for each value of i, and such a column of the origin-destination matrix would, also, play no role in the DSF procedure.
The fact that most of the Modified Scoring procedure running time is consumed by repeated calls to the DSF procedure can result in longer than necessary total running time, especially with the large size of matrices involved in this study. We found that stopping the DSF procedure once it had sufficiently converged, say, when (at iteration n) $\sum_i \sum_j \left| T_{ij}^{(n)} - N_{ij} \right| + \sum_j \sum_i \left| T_{ij}^{(n)} - N_{ij} \right| = 10E - 7$, and there was little change, say $10E - 20$, of the DSF stopping

criterion between additional successive DSF iterations, shortened the total running time to less than 30 minutes per model run on a business laptop (Intel Core i7-4800MQ CPU @2.70GHz, 16GB RAM, Windows 10 64-bit).

An overall assessment of the goodness of fit of the gravity model is the so-called Chi-square ratio (Chi-square divided by the degrees of freedom), $\mathrm{X}^2/df = \sum_i \sum_j \frac{(T_{ij}-N_{ij})^2}{T_{ij}} / [(I-1)(J-1)-K]$ (Sen and Smith, 1995). In practice, if the Chi-square ratio is less than 2 then the model fits well. This observation has been exploited in a number of empirical studies (Lowe and Sen, 1996; Metaxatos 2004, 2009). In the end, the model specification $T_{ij} = A_i B_j e^{-0.187\sqrt{c_{ij}^{(1)}}-0.001c_{ij}^{(3)}}$, where $c_{ij}^{(1)}$ is the in-vehicle travel time (in minutes) and $c_{ij}^{(3)}$ is the total wait time (in minutes), obtained the best overall fit and was used to estimate the spatial index of public transit accessibility, PTA_i .

The origin-destination trip and cost tables lacked intrazonal values. We examined two procedures recommended for intrazonal travel time estimation. The first method (Martin and McGuckin (1998) assumes that intrazonal travel times can be expressed as a function of the zonal area and the intrazonal speed, i.e.,

$$\text{intrazonal time} = 0.5 \times \sqrt{area} \times 60 / intrazonal\ speed\ (by\ area\ type),$$

where the intrazonal time is expressed in minutes, the zonal area is expressed in square miles, and the intrazonal speed in miles per hour varies by the area type of the zone. For example, the intrazonal (auto) speed for a CBD zone could be set at 15 miles per hour, and the intrazonal speed for a rural zone could be set at 30 miles per hour. For transit, we could take half or less the auto speed but this adjustment did not produce realistic intrazonal times.

The second method, called nearest neighbor method, assumes that the travel time within a zone is equal to one-half the average travel time to the nearest adjacent zones. In this paper, we computed one-half the average travel cost to the six nearest zones, which resulted in more realistic travel times (i.e., less than two percent of all trips estimated were intrazonal).

3.1.4 Discussion of Results

The implication of using the methodology described above to measure public transit accessibility is that, other factors aside, physical proximity to a healthcare facility does not necessarily indicate clear locational advantages for a typical resident searching for a healthcare facility. If there are numerous other individuals who also are in close proximity or who are conveniently linked by transportation access routes to the same healthcare facility, then there is little locational advantage.

Figure 1 illustrates the regional variations of peak-period public transit accessibility to healthcare facilities, PTA_i, in the Chicago area neighborhoods visualized as TAZs. The five groups of index values using the Jenks natural breaks optimization method are shown in graduated colors. Neighborhoods with the lowest accessibility potential score lower than 23% of the maximum 100% score that some neighborhoods receive. Similarly, the second lowest group scores between 22.7% and 49.8% of the maximum. The 'medium' group scores between 49.9% and 62.7%, and the second-to-highest group scores between 62.8% and 76.5% of the maximum. The top group includes neighborhoods receiving between 76.6% and 100% of the maximum score. Locations with the highest accessibility score (100%) do not have a healthcare facility.

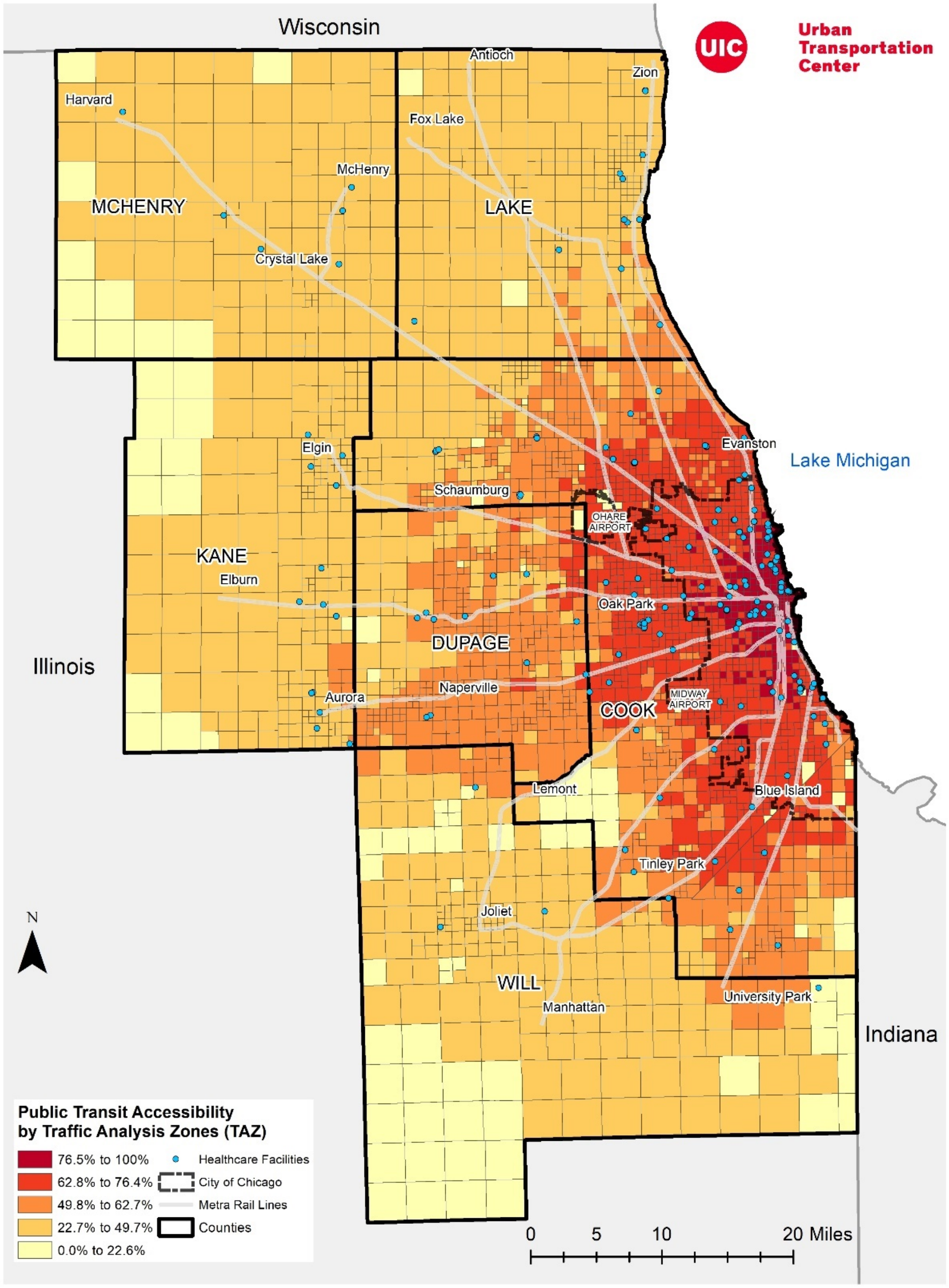


Figure 1. Public transit accessibility to healthcare facilities in the Chicago area.

Table 1 shows the five highest and five lowest scoring healthcare locations. Healthcare facility locations in neighborhoods in downtown Chicago and north/northwest of downtown score the highest accessibility values. On the other end of the spectrum, healthcare facility locations in areas in the far northwest suburbs (Woodstock, in two different TAZs, and Harvard in McHenry County), Hoffman Estates in northwest Cook County, and Crete in the far southeast part of Will County score the lowest accessibility values.

Table 1. Highest and Lowest Transit Accessibility Healthcare Facility Locations

Top Five Scoring Healthcare Locations		**Lowest Five Scoring Healthcare Locations**	
Area ZIP Code	**PTA_i Score**	**Area ZIP Code**	**PTA_i Score**
Chicago 60610	96.3%	Crete 60417	19.3%
Chicago 60612	95.8%	Hoffman Estates 60169	31.0%
Chicago 60657	95.6%	Harvard 60033	33.9%
Chicago 60616	95.2%	Woodstock 60098	37.7%
Chicago 60607	94.3%	Woodstock 60098	38.8%

The index PTA_i measures the aggregate peak-period public transit accessibility potential to the locations of healthcare facilities for each residential zone i. The neighborhood with the highest such accessibility measure is the one with best public transit access, as measured by F_{ij}, to all healthcare facilities in the region. Areas with the highest such values were found in the near northwest side neighborhoods of Chicago. While proximity to the CBD appears to be the most important factor, the high values extend further in a northwesterly direction than in a southwesterly direction. Also, the index values are significantly higher in the west side of the city community than in the south side community.

In the suburban areas, highest values are found in west suburban Cook County extending even into DuPage County and, to a lesser extent, Kane country. Here, again, the high values do not extend very far south. While southwest suburban communities have low values, the lowest values are located in the corner of the study area where the communities clearly suffer from remoteness. These are examples of the classical boundary effect problem. If the study area were larger, their values would be higher, but they would still likely be lower than most other values within the present study area.

Neighborhoods, especially in the northern half of the Central area of Chicago are highly accessible by public transit to healthcare facilities. Moreover, neighborhoods on the south side along CTA's Red Line also have relatively high accessibility values. On the other end of the spectrum, neighborhoods in the south east corner of the city of Chicago have by far the lowest accessibility by public transit to healthcare facilities. This is a result of a minimal number of nearby healthcare facilities and relative remoteness from downtown and other areas with many healthcare facilities in the city. Clearly, workers, visitors or patients residing in this part of the city have a locational disadvantage in accessing healthcare facilities by public transit compared to other users of healthcare facilities in the rest of the city.

3.2 MEASUREMENT OF PUBLIC TRANSIT AVAILABILITY

The public transit accessibility index, as described in the previous section, provides a measure to assess how accessible each neighborhood with a healthcare facility is from every other residential area using public transit. The measurement of a public transit availability index, in this section, measures the level of availability of public transit at each area with a healthcare facility.

The concept of transit availability is discussed within the framework provided by the Transit Capacity and Quality of Service Manual (Kittelson & Associates et al., 2013). The concept is further operationalized spatially and temporally by different measures that describe how often service is provided (frequency), how long service is provided (hours of service), and where service is provided (access), and several implementations have been realized.

The Florida Department of Transportation (FDOT), for example, calculated a transit availability index at various time points along the route of a transit facility. For a given geographic area, the index multiplies the percent of area served by the percent of an hour that a station or stop is served, assuming a 5-min wait time, by the percent of a day that the area is served by transit (Ryus et al., 2000). The index ranges between zero and one for each geographic unit and measures the percent of person minutes served by the bus or train facility at that point in time.

In a similar study (NYPIRG, 2001), the quality of service index for the New York subway system for 2001 was calculated as a composite measure of scheduled headways, service regularity, mean distance between failures, chances of finding a seat, and passenger responses on "cleanliness" and "adequacy of routine in-car announcements." The study determined improvement in the measures above by comparing them with the quality of service status in 1996-1997. On the basis of this, the study rated Subway Line Q as the best and C as the worst, in level of improvement based on these quality indicators. The latest report card is available at https://www.straphangers.org/reports/2016/StateoftheSubways2016.pdf (accessed 5/8/20).

Minosha et al. (2008) proposed a method to relate route-level information with U.S. Census tracts, as well as a method to combining availability scores from different public transit systems in the Chicago area developing thereby a composite index of transit availability and frequency and transit station asset information. The index used traditional transit availability measures, such as number of residents or number of jobs within walking distance of bus stops or rail stations, as well as additional information such as frequency, hours of service, and service coverage.

Two relevant indices developed by the Center of Neighborhood Technology are: (a) the Transit Access Shed (TAS) determines the distance and the number of transit stops that can be reached by public transportation from a given location in 30 minutes; and (b) the Transit Connectivity Index (TCI) measures the number of transit stops available within or near a block group. The measures are available at the block group level for all cities in the United States with a population of at least 50,000 people (Haas et al., 2013). Both indices were used to examine the role of public transportation in access to care of older adults in the United States (Zuckerman, 2016).

The discussion below continues with data acquisition details, and details about the methodology used. Later the methodology is demonstrated in the six-county Chicago study area.

3.2.1 Data Acquisition

GTFS Data

Three public transit agencies operate in the six-county Chicago study area. The Chicago Transit Authority (CTA) operates two modes, bus and rapid rail transit, in the city of Chicago, and provides connections with suburban systems. The Metropolitan Rail (Metra) commuter rail service operates 12 routes connecting the Chicago suburbs with the Chicago Central Business District. The third agency, Pace Bus, operates mostly in the six-county area except for the city of Chicago.

We downloaded General Transit Feed Specification (GTFS) data for each public transit provider in the Chicago area CTA, Metra and Pace from https://transitfeeds.com/l/146-chicago-il-usa. The data is collected for October 2019. From this data we are using records for only Wednesday that fall in peak time.

Each provider's data were included in a compressed ZIP file containing comma-separated values (CSV) files. Each CSV file models a particular aspect of transit information, that is, stops, routes, trips, and other schedule data. The following information is provided in the CSV files: (a) The agency which provides the data in this; (b) a schedule of when the service is available; (c) the transit routes available to riders within a single service; (d) the individual locations where vehicles pick up and drop off passengers; (e) the specific times that a vehicle arrives and departs from a stop location; and (f) information about a trip defined as a sequence of two or more stops that occurs at a specific time.

Census Data

We downloaded the following data from the U.S. Census Bureau https://www2.census.gov/:

- Block groups geography: We used the geography for the 5,841 block groups included in the six-county study area because it is the smallest geography for which the bureau publishes sample data.
- Household units: We used the 2013-2017 American Community Survey (ACS) 5-year estimates. The data contain information about the number of housing units in each block group.
- Employment data: We downloaded information about the number of jobs in each of the above block groups from the Longitudinal Employer-Household Dynamics (LEHD) https://onthemap.ces.census.gov/. Search data based on the following search criteria:
    - States – Illinois
    - Home/Work Area – Work
    - Area Comparison
        - Areas to compare – Census Block Groups
        - Labor Market Segment – All Workers
    - Year – 2017
    - Job Type – All Jobs

Other Geographic Data

We used the geographic files for routes included in GTFS system for CTA, Metra and Pace (over email). The routes cover the six-county Chicago study area.

3.2.2 Methodology Used

Of the measures applicable to transit availability at the station/stop level in the Transit Capacity and Quality of Service Manual, hours of service, frequency of service, as well as the reach of the service (percent of the population within a certain buffer of the transit line) can be most relevant from the users' perspective in evaluating the level of service. In fact, FDOT proposed the development of an index of transit availability that takes into account frequency and hours of service and service coverage (Ryus et al., 2000).

In this study the transit availability index is a measurement of the percent of person-minutes served. For a given geographic area, the index multiplies the percent of area served by transit by the percent of an hour that a station or stop is served (assuming a five-minute wait time) by the percent of a day that the area is served by transit. The index range is between zero and one for each geographic unit, since it is a product of the percentages of the three factors described above. The analysis is done at the census block group level.

In this study we adapted the algorithmic approach first described in Minosha et al. (2008) to estimate stop-level, route-level, and system-level availability for each block group. Three measures are used to develop an index of transit availability: frequency of service, hours of service, and service coverage. In this application, transit service was restricted to the morning peak (7:00 a.m. to 9:00 a.m.) of a typical Wednesday in September 2019 to be consistent with the range of the accessibility index in the previous section.

Frequency of Service Calculation

The first step is to estimate the percent of an hour that a transit stop is served. Since the frequency of transit vehicles changes throughout the day, it is necessary to determine a measure that accurately portrays the average frequency for an hour of the day. The methodology used by the New York Public Interest Research Group (NYPIRG) Straphangers Campaign attributes 40% of the measure to the frequencies during the AM and PM peak intervals each and the remaining 20% of the measure is attributed to the midday frequency (NYPIRG, 2001). In a similar manner, the breakdown of the frequency measure for the Chicago region was weighted so that AM and PM peak times counted for 35% each, midday times were assigned a 20% weight, and the overnight period was weighted at 10%. In this study we will be using the AM peak factor of 0.35.

All frequencies were taken from the weekday (Wednesday) schedules from the GTFS data. The frequencies of transit vehicles were then multiplied by five (the minutes of wait time) to determine the minutes per average hour that transit serves the individual transit station/stop. This number was then divided by sixty to determine the percent of an average hour transit serves the station/stop.

In cases where multiple transit lines serve the same stops/stations we used the minimum frequency for the worst case scenario. Therefore, each station/stop was assigned the lowest frequency measure of the routes that serve it. Since there are multiple stops in a given census block group, the average frequency of these various stops was computed using buffering techniques in ArcGIS 10. The end result is the percent of an average hour that transit is available at the station/stop at each block group.

Hours of Service Calculation

While Ryus et al., (2000) measured the hours of service as a percentage of the day that the service was operating, in this study this measurement was assigned to each transit route based on the schedules posted on the GTFS data for each transit agency. To mitigate the issue of multiple routes serving the same stop/station we assigned at each stop/station the hours of service measure from the route with the most hours of service. As a result, the particular station/stop is served by transit for the corresponding percentage of the day.

Quarter-mile buffers were used to show the service coverage of each station/stop. The buffers were then assigned the hours of service measure corresponding to the station/stop being buffered. Using the spatial join function of ArcGIS 10 the maximum hours of service measure of all the buffers intercepting the individual census tracts was assigned to each census block group.

Service Coverage Calculation

The third component of the transit availability index is service coverage. We used ARC GIS to calculate the area of each census block group that is covered by transit station/stop buffers. We then converted the area covered into a percentage of the total area of the census block group. This measure satisfies the service coverage area component as described by Ryus et al. (2001).

The three components (frequency, hours of service, and service coverage), which are represented as percentages, were multiplied to each other to calculate a composite index score. Since each component is a percent value (between 0 and 1), the combined transit availability index is also between 0 and 1 for each census block group. The presentation below lists all the steps taken to calculate the stop/station-level availability, the route-level availability, the system-level availability, and the combined stop/station-, route-, and system-level availability

Stop/Station-level Availability

Stop/Station-level availability is defined to be the percent of an hour a stop/station is served. Note that GTFS data had an arrival time at certain stops that went beyond 24:00:00. Such times needed to be converted to the 24-hour format and the day of service was changed to the next day. We then executed the following algorithmic steps:

1. The arrival time of buses/trains for each stop/station is sorted in ascending order.
2. A headway is calculated: Headway = current arrival time – previous arrival time.
3. If the headway is less than 5 minutes, then minutes served is the value of headway; if the headway is greater than or equal to 5 minutes then minutes served is 5 minutes.
4. As the minutes served is calculated for peak time, multiply the minutes served by 0.35.
5. Sum up the weighted minutes served. Divide this value by the length of the period and multiply by 100 to get the percent of hour that stop/station is served.

The Stop/station level availability is calculated using the above algorithm through a python script. This python script can be easily configured to take as input different GTFS data and peak time values. The output of the python script is an Excel file with stop/station level availability values for each stop/station in the six-county area.

Route-level Availability

Route-level availability is defined to be the percent of day the entire route is served. Note that GTFS data had an arrival time at certain stops that went beyond 24:00:00. Such times needed to be converted to the 24-hour format and the day of service was changed to the next day. We only kept records that fall in the peak time. We then executed the following algorithmic steps:

1. For each route, the start times for all its runs are taken.

2. Difference between consecutive start times are measured.
3. For each run, we find intervals where the service is provided for at least an hour.
4. An hour is added to each time interval where the service is for at least an hour.
5. These values are added to get the total hours of service for that route.
6. The percent hours of service is calculated with respect to the considered period – peak time (7 a.m. to 9 a.m.)

The route-level availability is calculated using the above algorithm through a python script. This python script can be easily configured to take as input different GTFS data. The output of the python script is an Excel file with stop/station level availability values for each route in the six-county area.

System-level Availability

System-level availability is defined to be the percent of eligible block groups served by the transit systems. A block-group is eligible to receive public transportation services if the number of jobs per acre is greater than or equal to 4, or the number of housing units per acre is greater than or equal to 3. We then used ArcMap to join the block-groups with the housing units and employment data based on the GEOID attribute of each block group. Then we calculated the area for each block group in acres. Finally, we created a layer for each of CTA Bus, CTA Rail, Pace Bus and Metra Rail stops/stations, as well as a layer for each of CTA Bus, CTA Rail, Pace Bus and Metra Rail routes. We then executed the following algorithmic steps:

1. Create a buffer of 0.5 miles for CTA Bus, CTA Rail and Pace Bus routes.
2. Create a buffer of 2.5 miles for Metra Rail routes.
3. Combine the buffers for CTA Bus, CTA Rail, PACE Bus and Metra rail routes into a single buffer region.
4. Determine transit-supportive areas. All block-groups with household density of 3.0 or more households per acre or a job density of 4.0 or more jobs per acre or both are identified.
5. Calculate the intersection analysis of this service coverage with the transit supportive areas. Estimate how much area of each transit supportive block-group is covered by the service coverage.
6. Sum up the areas of all the transit supportive block-groups.
7. Sum up the areas covered by the service coverage of all the transit supportive block-groups.

8. Calculate the percentage of area that is covered by the combined buffer region with respect to the overall area of the eligible regions. This percentage is the system-level availability.

The system-level availability is calculated using the above algorithm through a python script. This python script needs to be executed inside ArcMap. The end result is a score that signifies the system level availability for each block group.

Combined Stop/Station-, Route-, and System-level Availability

A combined stop/station-, route-, and system-level availability value is the product of stop/station-level, route-level and system-level availabilities for each block-group. The following algorithmic steps were executed.

1. Perform a spatial join between block-groups and stops/stations in ArcMap. This will provide us with a stop-level availability value for each block-group. If a block-group has multiple stops/stations, we select the stop/station with the minimum value of stop-level availability.
2. Perform a spatial join between block groups and routes in ArcMap. This will provide us with a route-level availability value for each block group. If a block group has multiple routes, we select the route with the minimum value of route-level availability.
3. We have skipped certain stops which were not getting served at all. These are the stops whose stop-level availability value is 0.
4. Multiply the stop/station-level, route-level and system-level availabilities for each block-group to obtain one value for each block group.
5. Create a thematic map of these values using different color shades for different range of values (Figure 2).

Note that combining different availability values was implemented within the ArcMap environment. This is because the particular ArcMap version used did not have the functionality in python (in ArcPy) that can perform a spatial join by selecting the minimum in situations where there are multiple matching entities.

Discussion of Results

Figure 2 illustrates the regional variations of peak-period public transit availability to healthcare facilities in the Chicago area neighborhoods visualized as census block groups. The five groups of index values using the Jenks natural breaks optimization method are shown in graduated colors. Block groups with the lowest transit availability score up to 3.1% of the maximum 100% score that some block groups receive. Similarly, the second lowest group scores between 3.2% and 9.4% of the maximum. The 'medium' group scores between 9.5% and 16.8%, and the second-to-highest group scores between 16.9% and 28.7% of the maximum. The top group includes block groups receiving between 28.8% and 100% of the maximum score. Block groups with the highest availability score (100%) do not have a healthcare facility.

Overall, it is quite evident from the map that census block groups in Cook County and parts of the suburban counties that are close to Cook as well as along the major Metra routes are well served by transit in the region. The block groups toward the north and west of the study region have a more uniform distribution in contrast to the block groups in the south because most of these block groups are served by frequent and multiple systems at the same time, especially within the city of Chicago.

Interestingly, more than half (109 out 206) healthcare facilities in the study region are located in census block groups that score very poorly on the transit availability index. On the contrary, only 27 out of 206 healthcare facilities in the six-county region are located in block groups with the highest transit availability scores. Most of these facilities are located in the city of Chicago and in the path of Metra's north and southwest lines. Such observations seem to be in line with past studies as discussed earlier. However, before discussing policy implications we would need to examine the availability question in conjunction with the accessibility issue discussed in the previous section. The joint analysis of the transit availability and transit accessibility indices will give rise to potential deficiencies that warrant intervention. This is the focus in the next section.

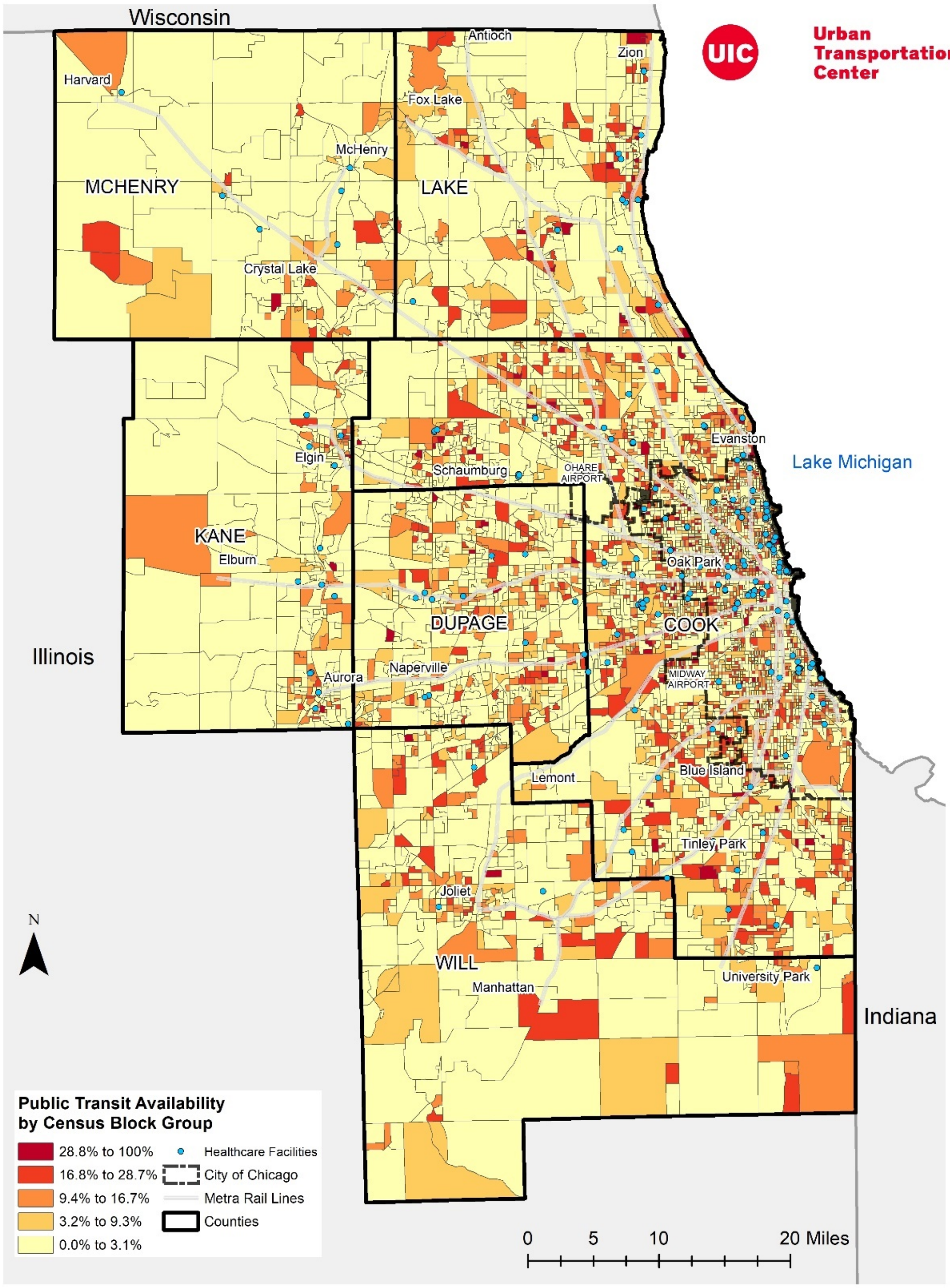


Figure 2. Public transit availability to healthcare facilities in the Chicago area.

4. DEFICIENCY ANALYSIS

The previous section discussed the development of two indices: a public transit accessibility index and a public transit availability index. The joint use of both indices relates the following information to each neighborhood with a healthcare facility: the level of public transit accessibility from every residential zone, and the level of public transit availability. A comparative analysis of the two indices will inform the deficiency analysis in this section.

We found that 28% (58 out of 206) of healthcare facilities in the six-county region are located in areas with below average transit accessibility and transit availability (Table 2). Interestingly, areas with such deficiencies can be found in all six counties of the study area.

Table 2. Healthcare Facilities in Areas with Poor Transit Availability and Accessibility

Healthcare Facility Name	County Name	Transit Availability	Transit Accessibility
Adventist Bolingbrook Hospital	Will County	0.0%	53.4%
Advocate Good Samaritan Hospital	DuPage County	0.0%	61.7%
Advocate Good Shepherd Hospital	Lake County	0.0%	44.2%
Advocate South Suburban Hospital	Cook County	0.0%	62.4%
Advocate Trinity Hospital	Cook County	0.0%	58.7%
Alden Poplar Creek Convalescent Center	Cook County	0.0%	48.4%
Alexian Brothers Behavioral Health Hospital	Cook County	0.0%	48.4%
Alexian Brothers Medical Center	Cook County	0.0%	55.2%
Ambutal Hospital Trauma Center	McHenry County	0.0%	45.0%
American International Hospital	Lake County	0.0%	39.5%
Amita Health Adventist Medical Center GlenOaks	DuPage County	0.0%	56.0%
Brock Medical Plaza	Cook County	0.0%	55.2%
Centegra Memorial Medical Center	McHenry County	0.0%	39.9%
Centegra Memorial Medical Center South Street Campus	McHenry County	0.0%	37.7%
Central Dupage Hospital	DuPage County	0.0%	54.6%
Community Hospital (historical)	Kane County	0.0%	45.1%
Concentra Immediate Care Center	DuPage County	0.0%	46.0%
Condell Medical Center	Lake County	0.0%	46.4%
Copley Hospital (historical)	Kane County	0.0%	45.3%
Delnor Community Hospital	Kane County	0.0%	44.8%
Delnor Hospital (historical)	Kane County	0.0%	41.2%
DuPage County Home	DuPage County	0.0%	52.9%
Edward Hospital	DuPage County	0.0%	57.9%
Elgin Mental Health Center	Kane County	0.0%	44.5%
Elmhurst Memorial Hospital	DuPage County	0.0%	42.8%

Healthcare Facility Name	County Name	Transit Availability	Transit Accessibility
Franciscan Health Olympia Fields	Cook County	0.0%	58.4%
Glenbrook Hospital	Cook County	0.0%	52.2%
Glendale Heights Community Hospital	DuPage County	0.0%	56.0%
Holy Family Medical Center	Cook County	0.0%	61.9%
Kane County Home	Kane County	0.0%	43.8%
Lake County Sanatorium (historical)	Lake County	0.0%	42.4%
Lake Forest Hospital	Lake County	0.0%	46.5%
Linden Oaks Hospital	DuPage County	0.0%	57.9%
Manor Care Nursing Center	Cook County	0.0%	55.8%
McHenry Hospital	McHenry County	0.0%	42.1%
Mercyhealth Hospital Harvard	McHenry County	0.0%	33.9%
Mercyville Sanitarium	Kane County	0.0%	48.0%
Midwest Physicians Center	Cook County	0.0%	61.9%
Midwestern Regional Medical Center	Lake County	0.0%	39.5%
Niehoff Pavilion	Cook County	0.0%	51.6%
Northern Illinois Medical Center	McHenry County	0.0%	40.2%
Northwest Community Hospital	Cook County	0.0%	55.1%
Oak Forest Hospital of Cook County	Cook County	0.0%	63.4%
Palos Primary Care Medical Center	Cook County	0.0%	59.7%
Provena Mercy Medical Center	Kane County	0.0%	48.0%
Provena Saint Joseph Medical Center	Will County	0.0%	42.9%
Rush Copley Medical Center	Kane County	0.0%	47.7%
Saint Alexius Medical Center	Cook County	0.0%	48.4%
Saint James Hospital	Will County	0.0%	19.3%
Saint James Hospital and Health Center	Cook County	0.0%	58.1%
Sherman Hospital	Kane County	0.0%	39.1%
Silver Cross Hospital	Will County	0.0%	44.7%
Tinley Park Mental Health Center	Will County	0.0%	47.0%
Vista Health System Victory Memorial Hospital	Lake County	0.0%	42.0%
Vista Medical Center West Campus	Lake County	0.0%	42.8%
Wimmer Medical Plaza	Cook County	0.0%	55.2%
Wyngarden Health Center	DuPage County	0.0%	53.7%
Zace Sanitarium	DuPage County	0.0%	49.2%

The implication for the healthcare facilities in Table 2 is that their clients would need to be more reliant on private automobiles to access their services. It would certainly require locally targeted interventions to mitigate such deficiencies in public transit accessibility of healthcare facilities in the Chicago area. Such interventions can potentially reduce social inequities regarding the healthcare accessibility, enhance transportation safety and environment in the neighborhoods with healthcare facilities, reduce mental stress on work and healthcare trips, reduce greenhouse emissions from auto-dependent trips, and foster collaboration among stakeholders and communities, which together contribute to local and regional sustainability and better health outcomes for all in the society.

In addition, to the previous problematic from a transit availability and accessibility (during the morning peak) viewpoint, our analysis found other areas that do not exhibit deficiencies to the same extent. For example, the 50 (out of 206, or 24%) healthcare facilities in Table 3 are located in areas (all in Cook County) with below average transit availability but above average transit accessibility.

Table 3. Healthcare Facilities in Areas with Poor Transit Availability and Better Transit Accessibility

Healthcare Facility Name	County Name	Transit Availability	Transit Accessibility
Advocate Illinois Masonic Medical Center	Cook County	0.0%	86.3%
Bethany Brethren Hospital	Cook County	0.0%	78.3%
Booth Memorial Hospital (historical)	Cook County	0.0%	67.9%
Cardinal Bernardin Cancer Center	Cook County	0.0%	68.5%
Columbus Hospital (historical)	Cook County	0.0%	78.2%
Cuneo Hospital	Cook County	0.0%	70.6%
Edward Hines Junior Veterans Affairs Hospital	Cook County	0.0%	68.5%
Fahey Center	Cook County	0.0%	68.5%
Frank Cuneo Hospital (historical)	Cook County	0.0%	70.6%
Garfield Park Hospital	Cook County	0.0%	87.1%
Hal Sanitarium	Cook County	0.0%	96.4%
Hartgrove Behavioral Health System	Cook County	0.0%	71.6%
Ingalls Memorial Hospital	Cook County	0.0%	64.6%
Jesse Brown Veterans Administration Medical Center	Cook County	0.0%	82.6%
John H Stroger Junior Hospital of Cook County	Cook County	0.0%	86.3%
John J Madden Mental Health Center	Cook County	0.0%	68.5%
La Grange Memorial Hospital	Cook County	0.0%	68.8%
La Rabida Children's Hospital	Cook County	0.0%	75.8%
Lakeside Veterans Administration Hospital	Cook County	0.0%	86.2%
Little Company of Mary Hospital	Cook County	0.0%	72.4%
Loretto Hospital	Cook County	0.0%	71.6%
Loyola Center for Health and Fitness	Cook County	0.0%	68.5%
Loyola Outpatient Center	Cook County	0.0%	68.5%
Loyola University Medical Center	Cook County	0.0%	68.5%
Maguire Center	Cook County	0.0%	68.5%
Mercy Hospital and Medical Center	Cook County	0.0%	81.3%
Metro South Medical Center	Cook County	0.0%	64.0%
Mulcahy Center	Cook County	0.0%	68.5%

Healthcare Facility Name	County Name	Transit Availability	Transit Accessibility
Northwestern Memorial Hospital - Prentice	Cook County	0.0%	86.2%
Old Orchard Hospital (historical)	Cook County	0.0%	64.0%
Passavant Hospital	Cook County	0.0%	86.2%
Rehabilitation Institute of Chicago	Cook County	0.0%	86.2%
Riveredge Hospital	Cook County	0.0%	69.1%
RML Specialty Hospital	Cook County	0.0%	78.3%
RML Specialty Hospital	Cook County	0.0%	65.1%
Ronald McDonald Childrens Hospital of Loyola University Medical Center	Cook County	0.0%	68.5%
Rush University Medical Center	Cook County	0.0%	88.7%
Rush-Presbyterian-Saint Lukes Hospital	Cook County	0.0%	88.7%
Saint Bernard Hospital and Health Care Center	Cook County	0.0%	80.9%
Saint Joseph Hospital	Cook County	0.0%	76.3%
Saint Lukes Hospital (historical)	Cook County	0.0%	87.5%
Skokie Hospital	Cook County	0.0%	64.0%
Skokie Valley Hospital	Cook County	0.0%	64.0%
The Robert H Lurie Medical Research Center of Northwestern University	Cook County	0.0%	86.2%
United States Health Public Hospital	Cook County	0.0%	76.9%
University of Illinois at Chicago Hospital	Cook County	0.0%	86.3%
Veterans Administration Lakeside Medical Center	Cook County	0.0%	86.2%
Weiss Memorial Hospital	Cook County	0.0%	70.6%
Woodlawn Hospital*	Cook County	0.0%	74.7%

*The hospital is included with identical name twice in the database, but with different facility ID each time.

The healthcare facilities in Table 3 are located in areas that benefit from relatively high transit accessibility, but still exhibit poor transit availability. These areas will certainly benefit from additional transit investments. As an example of a local intervention that aims improving, among other objectives, the transit availability, we could cite the case of the Illinois Medical District (IMD) which is located less than two miles west of Chicago's CBD and is the largest urban medical district in the country. The IMD has four core medical institutions that are also included in Table 3: (a) the University of Illinois at Chicago hospital; (b) the John H. Stroger Jr. Hospital of Cook County; (c) the Rush University Medical Center; and (d) the Jesse Brown Veterans Affairs Medical Center. Two of the transportation-related goals of the IMD strategic plan is (a) to improve multiple modes of transportation and access for all District stakeholders; and (b) to establish efficient transportation alternatives and connections throughout the District and coordinate transportation initiatives with institutional and community needs (Illinois Medical District Master Plan http://medicaldistrict.org/wp-content/uploads/pdf/IMD_Master_Plan_SCB_01_2016.pdf - accessed 6/26/19).

Another group of healthcare facilities are located in areas that exhibit good transit availability but poor transit accessibility. These 13 (out of 206, or 6%) healthcare facilities are shown in Table 4.

Table 4. Healthcare Facilities in Areas with Good Transit Availability and Poor Transit Accessibility

Healthcare Facility Name	County Name	Transit Availability	Transit Accessibility
Adventist Hinsdale Hospital	DuPage County	17.0%	63.3%
Captain James A Lovell Federal Health Care Center	Lake County	1.4%	49.8%
Chicago Kindred Hospital Northlake	Cook County	6.4%	63.0%
Downey Veteran Administration Hospital	Lake County	1.4%	49.8%
Mercy Medical Center	Cook County	0.1%	55.6%
Naval Health Clinic	Lake County	1.4%	49.0%
NorthShore University HealthSystem - Highland Park Hospital	Lake County	7.6%	54.9%
Palos Community Hospital	Cook County	2.1%	62.8%
Provena Saint Joseph Hospital	Kane County	0.9%	43.8%
Saint Anns Infirmary	Cook County	0.4%	58.1%
Saint Joseph Hospital	Kane County	7.9%	48.3%
Saint Josephs Hospital	Kane County	0.4%	45.0%
Shriners Hospital for Children Chicago	Cook County	1.7%	61.2%

Transit accessibility to these healthcare facilities could improve through local economic development policies. An example would be to bring affordable housing closer to those healthcare facilities.

Finally, the rest of the healthcare facilities are located in areas with good transit availability and accessibility during the morning peak period. These 85 facilities (out of 206, or 41%) are shown in Table 5. All of these facilities are located in Cook County.

Table 5. Healthcare Facilities in Areas with Good Transit Availability and Transit Accessibility

Healthcare Facility Name	County Name	Transit Availability	Transit Accessibility
Advocate Christ Medical Center	Cook County	3.1%	71.0%
Advocate Lutheran General Children's Hospital	Cook County	3.0%	67.7%
Advocate Lutheran General Hospital	Cook County	3.0%	67.7%
Alexian Hospital	Cook County	20.5%	88.6%
Anne and Robert Lurie Children's Hospital of Chicago	Cook County	3.9%	86.2%

Healthcare Facility Name	County Name	Transit Availability	Transit Accessibility
Augustana Hospital	Cook County	10.2%	84.1%
Bernard Mitchell Hospital	Cook County	10.2%	77.7%
Bethany Hospital	Cook County	25.6%	75.5%
Bethesda Hospital	Cook County	5.1%	69.2%
Charter Barclay Hospital	Cook County	27.3%	77.3%
Chicago Lakeshore Hospital	Cook County	10.2%	75.8%
Chicago Lakeshore Hospital (historical)	Cook County	10.2%	75.8%
Chicago Lying-In Hospital	Cook County	10.2%	77.7%
Chicago Nursery and Orphan Asylum	Cook County	25.6%	71.9%
Chicago Read Mental Health Center	Cook County	13.6%	65.1%
Chicago Sanitarium	Cook County	5.1%	69.6%
Childrens Hospital	Cook County	26.8%	78.2%
Children's Memorial Hospital (historical)	Cook County	10.2%	84.1%
Community First Medical Center	Cook County	17.0%	66.1%
Doctors Hospital (historical)	Cook County	10.2%	72.8%
Edgewater Medical Center (historical)	Cook County	18.7%	70.9%
Englewood Hospital (historical)	Cook County	11.9%	73.3%
Evangelical Hospital	Cook County	10.2%	75.9%
Evanston Hospital	Cook County	8.5%	67.4%
Fairview Health Care Center	Cook County	7.1%	67.8%
Forest Hospital (historical)	Cook County	0.5%	64.5%
Forkosh Hospital	Cook County	18.7%	75.4%
Gottlieb Memorial Hospital	Cook County	6.9%	64.6%
Grant Hospital	Cook County	18.7%	84.1%
Henrotin Hospital (historical)	Cook County	17.0%	91.5%
Holy Cross Hospital	Cook County	13.6%	66.7%
Home for the Aged	Cook County	4.3%	69.0%
Hyde Park Hospital	Cook County	23.9%	70.8%
Illinois Eye and Ear Infirmary	Cook County	13.6%	94.3%
Jackson Park Hospital and Medical Center	Cook County	13.6%	67.7%
Jackson Park Hospital and Medical Center	Cook County	13.6%	67.7%
Kindred Chicago Central Hospital	Cook County	16.3%	74.2%
Kindred Hospital Chicago North	Cook County	18.7%	75.4%
Lakeside Hospital (historical)	Cook County	2.1%	95.3%
Lincoln West Medical Center	Cook County	18.7%	75.4%
Lutheran General Hospital (historical)	Cook County	3.0%	67.7%
MacNeal Hospital	Cook County	6.8%	72.5%
Martha Washington Hospital (historical)	Cook County	18.7%	77.8%
Mary Thompson Hospital (historical)	Cook County	6.0%	95.9%
Methodist Hospital of Chicago	Cook County	25.6%	75.5%
Michael Reese Developmental Institute and Childhood Development Center	Cook County	29.0%	73.7%

Healthcare Facility Name	County Name	Transit Availability	Transit Accessibility
Mount Sinai Hospital	Cook County	17.0%	76.6%
Mount Sinai Hospital (historical)	Cook County	17.0%	76.6%
Municipal Tuberculosis Sanitarium (historical)	Cook County	15.3%	66.1%
Northwest Hospital	Cook County	17.0%	66.1%
Northwestern Memorial Hospital	Cook County	3.9%	86.2%
Norwegian American Hospital	Cook County	20.5%	74.1%
Olson Hospital	Cook County	3.9%	86.2%
Presence Resurrection Medical Center	Cook County	2.8%	71.1%
Presence Saint Francis Hospital	Cook County	18.8%	70.0%
Presence Saints Mary and Elizabeth Medical Center - Saint Elizabeth Campus	Cook County	18.7%	82.5%
Presence Saints Mary and Elizabeth Medical Center - Saint Mary Campus	Cook County	17.0%	78.3%
Provident Hospital of Cook County	Cook County	17.0%	76.6%
Ravenswood Hospital and Medical Center (historical)	Cook County	23.9%	78.5%
Reese Hospital (historical)	Cook County	36.6%	79.2%
Rest Haven Hospital	Cook County	17.0%	76.6%
Roosevelt Memorial Hospital	Cook County	18.7%	92.3%
Roseland Community Hospital	Cook County	13.6%	72.5%
Rubloff Intensive Care Tower	Cook County	10.2%	77.7%
Rush Oak Park Hospital	Cook County	4.7%	71.0%
Sacred Heart Hospital	Cook County	6.8%	78.8%
Saint Annes Hospital (historical)	Cook County	27.3%	76.7%
Saint Anthony Hospital	Cook County	13.6%	78.3%
Saint Cabrini Hospital (historical)	Cook County	18.6%	84.3%
Saint Josephs Hospital (historical)	Cook County	46.4%	76.3%
Saint Vincents Hospital	Cook County	20.5%	95.6%
South Shore Hospital	Cook County	41.2%	70.6%
Surgery-Brain Research Pavilion	Cook County	10.2%	77.7%
Swedish Covenant Hospital	Cook County	25.6%	71.9%
Thorek Hospital and Medical Center	Cook County	10.2%	80.9%
Thorek Memorial Hospital	Cook County	10.2%	76.9%
University Hospital (historical)	Cook County	6.0%	75.8%
University of Chicago Medical Center	Cook County	10.2%	77.7%
Von Solbrig Hospital	Cook County	18.7%	73.6%
Walther Hospital (historical)	Cook County	6.0%	75.8%
Wesley Memorial Hospital	Cook County	3.9%	86.2%
West Suburban Medical Center	Cook County	6.8%	73.0%
Westlake Hospital	Cook County	27.3%	69.8%
Windemere Senior Health Center	Cook County	17.0%	70.8%
Wyler Childrens Hospital	Cook County	10.2%	77.7%

Certainly, healthcare facilities located in areas with high-level transit availability and accessibility would enjoy an advantage from a high level of access to their services by public transportation. In other words, public transportation is no longer a barrier to care for these facilities.

5. CONCLUSIONS

The welfare of the transportation disadvantaged may be at risk when access to healthcare for routine physicals, medical treatment and follow-ups is unnecessarily inhibited by poor public transit access. In urban areas with strong transit presence it is only prudent that further transit improvements target neighborhoods that lack such critical access. In this regard, this study proposed and demonstrated two indices of transit accessibility (on the demand side) and transit availability (on the supply side) that can be used as benchmarks toward measuring deficiencies in public transit accessibility of neighborhoods with healthcare facilities in the Chicago area.

Ideally, we would like to have more information about the market segments in Chicago who access healthcare facilities either as patients or workers. In the absence of more information about the travel behavior of prospective clients and/or workers to the healthcare facilities, these transit users are thought as segments of the overall transit demand during the morning peak period. If more relevant information becomes available, the study's methodological framework can readily accommodate it. This is because both indices are grounded on sound methodologies and can be replicated as needed in other areas with routinely available data from transit properties and planning authorities.

Future developments for both indices include: (a) adding an off-peak component to investigate differences in deficiencies from the peak period; (b) refining the spatial resolution of the indices; (c) expanding the inventory of healthcare facilities; and (d) improving visualization of the results to augment the user experience. Another issue that will be considered in a follow-up study is a comparative deficiency analysis between city of Chicago and Chicago suburbs. It is a fact that the population density (demand) and the transit service (supply) are significantly different between the city and the suburbs and will present these issues in a different light if the geographies were to be parsed and studied separately for the demand vs supply issues.

REFERENCES

Benenson, I., E., Ben-Elia, Rofé, Y., and D. Geyzersky, D. (2016). The Benefits of a High-resolution Analysis of Transit Accessibility. International Journal of Geographical Information Science, 8816(May), 1–24. http://doi.org/10.1080/13658816.2016.

Crooks, V.A. and N. Schuurman (2012). Interpreting the Results of a Modified Gravity Model: Examining Access to Primary Health Care Physicians in Five Canadian Provinces and Territories. BMC Health Services Research, 12 (1), p. 230.

Flores, G., M. Abreu, M.A. Olivar, and B. Kastner (1998). Access Barriers to Health Care for Latino Children. Archives of Pediatrics and Adolescent Medicine, 152(11):1119–1125.

Geurs, K.T. and B. van Wee (2004). Accessibility Evaluation of Land-use and Transport Strategies: Review and Research Directions. Journal of Transport Geography, 12(2), 127–140. http://doi.org/10.1016/j.jtrangeo.2003.10.005.

Hobson J., J. Quiroz-Martinez, and C. Yee (2002). Roadblocks to Health: Transportation Barriers to Healthy Communities. Center for Third World Organizing (CTWO), People United for a Better Oakland (PUEBLO), Transportation and Land Use Coalition (TALC).

Glaeser, E.L., M.E. Kahn, and J. Rappaport J. (2008). Why Do the Poor Live in Cities? The Role of Public Transportation. Journal of Urban Economics, 63:1-24.

Graham, S., B. Lewis, B. Flanagan, M. Watson, and L. Peipins (2015). Travel by Public Transit to Mammography Facilities in 6 US Urban Areas. Journal of Transport & Health, 2(4), 602-609, DOI: 10.1016/j.jth.2015.09.001.

Haas P., S. Morse, S. Becker, L. Young, and P. Esling (2013). The Influence of Spatial and Household Characteristics on Household Transportation Costs. Research in Transportation Business & Management, 7, 14-26, doi.org/10.1016/j.rtbm.2013.03.004.

Hansen W.G. (1959). How Accessibility Shapes Land Use. Journal of the American Institute of Planners, 25(2), 73-76, DOI: 10.1080/01944365908978307.

Joseph, A.E. and P.R. Bantock (1982). Measuring Potential Physical Accessibility to General Practitioners in Rural Areas: A Method and Case Study. Social Science & Medicine, 16(1), 85-90, DOI: 10.1016/0277-9536(82)90428-2.

Kittelson & Associates, Inc., Parsons Brinckerhoff, KFH Group, Inc., Texas A&M Transportation Institute, and Arup (2013). Transit Capacity and Quality of Service Manual, Third Edition. TCRP Report 165. Transit Cooperative Research Program, Transportation Research Board, National Academies of Sciences, Engineering, and Medicine, Washington, D.C.

Kwan, M.-P. (1998). Space-time and Integral Measures of Individual Accessibility: A Comparative Analysis Using a Point-based Framework. Geographical Analysis, 30(3), 191–216. Applied Geography, Volume 93, April 2018, Pages 47-63. doi.org/10.1111/j.1538-4632.1998.tb00396.x.

Lee, J. and H.J. Miller (2018). Measuring the Impacts of New Public Transit Services on Space-time Accessibility: An Analysis of Transit System Redesign and New Bus Rapid Transit in Columbus, Ohio, USA. doi.org/10.1016/j.apgeog.2018.02.012

Liu, S., and X. Zhu (2004). Accessibility Analyst: An integrated GIS tool for Accessibility Analysis in Urban Transportation Planning. Environment and planning B: Planning and Design, 31(1), 105–124. http://doi.org/10.1068/b305.

Lowe, J.M. and A. Sen (1996). Gravity Model Applications in Health Planning: Analysis of an Urban Hospital Market. Journal of Regional Science, 36(3), 437-461.

Luo W, Wang FH (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, 865-884, DOI: 865-884. DOI: 10.1068/b29120.

Martin, W. A. and N. A. McGuckin (1998). Travel Estimation Techniques for Urban Planning. NCHRP Report 365. American Association of State Highway and Transportation Officials, Washington, D. C.

Mattson, J. (2011). Transportation, Distance, and Health Care Utilization for Older Adults in Rural and Small Urban Areas. Transportation Research Record 2011, Transportation Research Board, National Research Council, Washington, D.C., 192-199.

Metaxatos, P. (2004). Estimation and Accuracy of Origin-Destination Highway Freight Weight and Value Flows. Journal of Transportation and Statistics, 6(2/3): 67-80.

Metaxatos, P. (2009). Synthetic Data Generation for Small-Area Demand Forecasting of Freight Flows. Operations and Supply Chain Management, 2(1), 42-51.

Minosha, I., P.S. Sriraj, P. Metaxatos, and P. Thakuriah (2008). An Analysis of Transit Quality of Service and Employment Accessibility for the Greater Chicagoland Region. Transportation Research Record 2042, Transportation Research Board, National Research Council, Washington, D.C., 20-29.

New York Public Interest Research Group Straphangers Campaign (2001). State of Subways Report Card, Appendix 1, NY.

Rask, K.J., M.V. Williams, R.M. Parker, and S.E. McNagny (1994). Obstacles Predicting Lack of a Regular Provider and Delays in Seeking Care for Patients at an Urban Public Hospital. JAMA, the Journal of the American Medical Association, 271(24):1931–1933.

Rittner, B. and A.B. Kirk (1995). Health-Care and Public Transportation Use by Poor and Frail Elderly People. Social Work, 40:365-73.

Ruggiano, N., N. Shtompel, K. Whiteman, and K.thy Sias (2017). Influences of Transportation on Health Decision-Making and Self-Management Behaviors among Older Adults with Chronic Conditions. Behavioral Medicine, 43:1, 61-70, DOI:10.1080/08964289.2015.1065788.

Ryus, P., J. Ausman., D. Teaf, M. Cooper, and M. Knoblauch (2000). Development of Florida's Transit Level-of-Service Indicator. Transportation Research Record 1731, Transportation Research Board, National Research Council, Washington, D.C., 123–129.

Sen, A., P. Metaxatos, S. Sööt and V. Thakuriah (1999). Welfare Reform and Spatial Matching between Clients and Jobs. Papers of Regional Science, Vol. 78, No. 2, 195-211.

Sen, A. and T.E. Smith (1995). Gravity Models of Spatial Interaction Behavior. Springer-Verlag.

Schuurman, N., M. Berube, and Crooks V.A. (2010). Measuring Potential Spatial Access to Primary Health Care Using a Modified Gravity Model. The Canadian Geographer, 54: 29-45, DOI: 10.1111/j.1541-0064.2009.00301.x.

Silver, D., J. Blustein, and B.C. Weitzman (2012). Transportation to Clinic: Findings from a Pilot Clinic-Based Survey of Low-Income Suburbanites. Journal of Immigrant Minority Health, 14, 350–355, DOI: 10.1007/s10903-010-9410-0.

Sööt, S. and A. Sen (1991). A Spatial Employment and Economic Development Model. Papers in Regional Science, 70(2), 149-166.

Syed, S.T., B.S. Gerber, and L.K. Sharp (2013). Traveling Towards Disease: Transportation Barriers to Health Care Access. Journal of Community Health, 38(5), 976–993, DOI: 10.1007/s10900-013-9681-1.

Yun S. and A. Sen (1994). Computation of Maximum Likelihood Estimates of Gravity Model Parameters. Journal of Regional Science, 34, 199-216.

Zuckerman, R. (2016). The Role of Public Transportation in Access to Care for Older Adults in the United States. Ph.D. dissertation, Johns Hopkins University, Baltimore, Maryland. May, 2016 (https://jscholarship.library.jhu.edu/bitstream/handle/1774.2/40284/ZUCKERMAN-DISSERTATION-2016.pdf?sequence=1&isAllowed=y, accessed 6/26/19).

APPENDIX

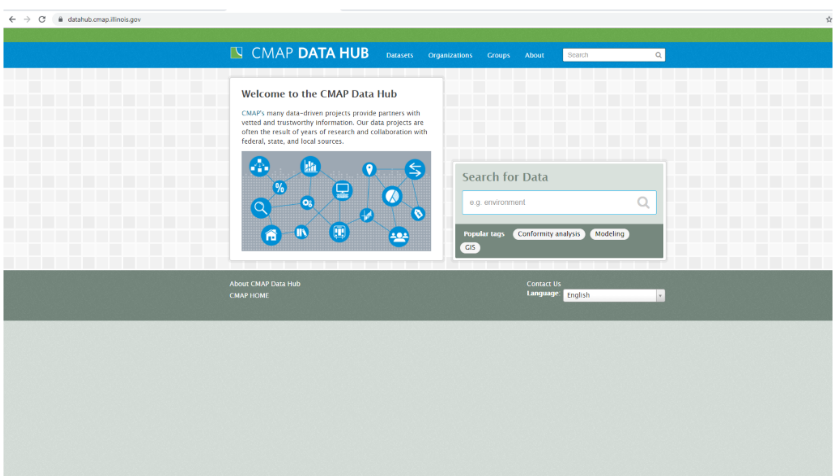
We are documenting details of the activities conducted to estimate public transit accessibility from data acquisition to code execution.

Getting the data

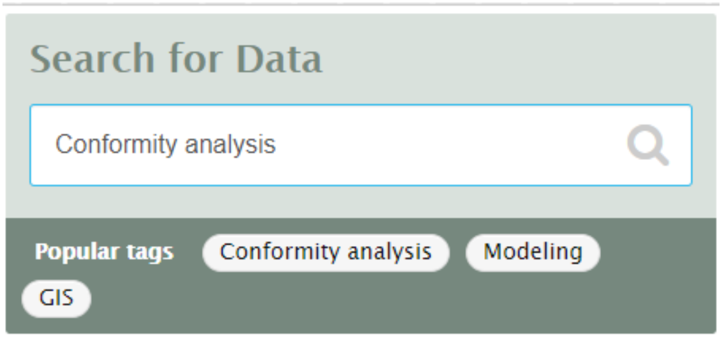
The program uses transit trip matrices and transit skim matrices (for impedances between zones) from Air Quality Conformity Analysis made available from CMAP. The air quality analysis is completed twice annually, in the first quarter and the third quarter. The data associated with the analysis is named based on the year the analysis was completed (C19 for 2019) and the quarter it was completed (Q1 or Q3). Therefore, the files in the dataset are referred to, say, C19Q3 data, for 2019 analysis year and Q3 quarter.

The dataset for any required analysis year can be found at CMAP Data Sharing Hub. Below are the steps to retrieve the files by manual downloading.

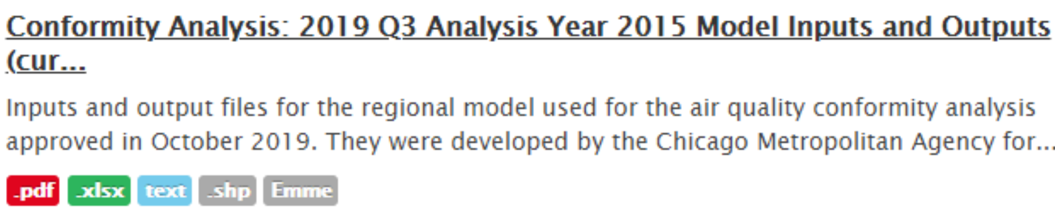
1. Go to CMAP data hub



2. In the block for **Search For Data** type **Conformity analysis** and press Enter

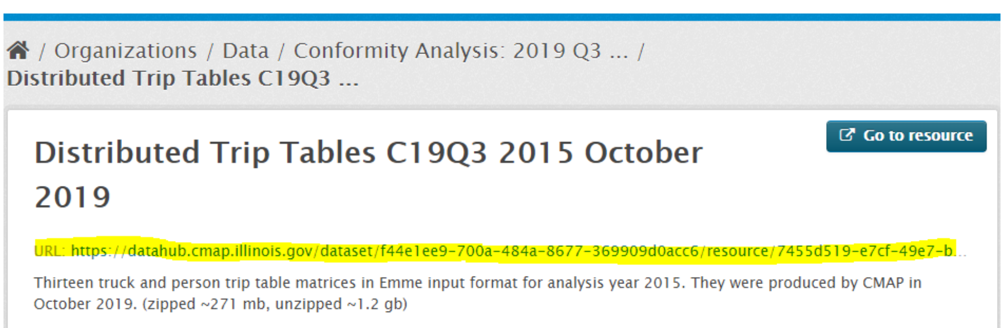


3. Each conformity cycle includes a bunch of analysis years. Click on the required dataset. For example, if you need the dataset that was completed in the year 2019 for the analysis year 2015 Q1 quarters then click on. Also, one can filter the search by typing in the year required in the search tab.

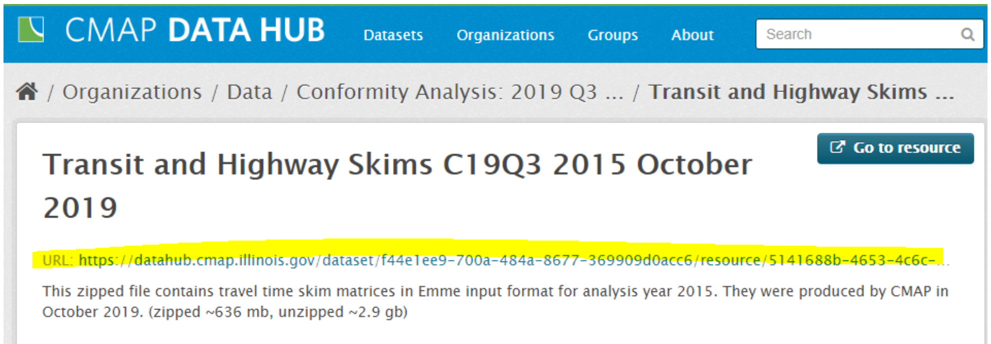


4. Once you click on the required link, you will find a number of subfolders. Following are the folders that the program uses. For example, for the above dataset,

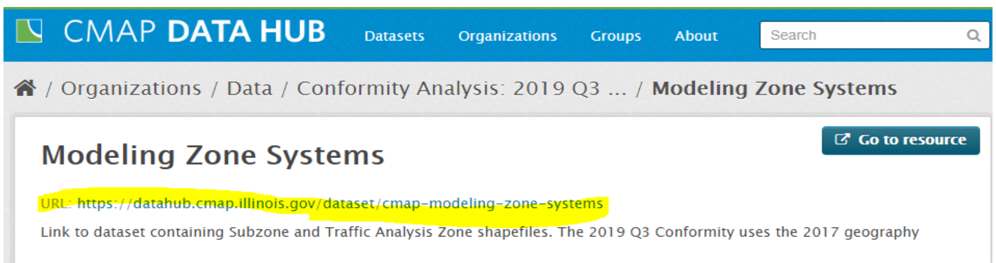
   a. Distributed Trip Tables C19Q3 2015 October 2019 (Transit Trip Matrices). Click on URL to download the entire folder



   b. Transit and Highway Skims C19Q3 2015 October 2019 (Transit Skim Matrices for Impedances between Zones). Click on URL to download the entire folder.



c. Modeling Zone Systems (To get the coordinates for Traffic analysis zones) – Optional if one already has coordinates for zone 2017. Then click on one of the four required resources.



d. Data Description C19Q3 2015 October 2019 (For information about the data files).

5. Once all the folders are downloaded including the Data Description file, unzip the folder and select the following from each of the folders: There are two categories of data: Peak and Off-peak.

Peak Data:

Distributed Trip Tables folder (named as tripsc19q3100 for e.g.) – pick:

a. **mf14.txt** (Home-based work transit person trips) – refer to the Data description file.

Distributed Trip Tables C19Q3 2015 October 2019 (trips_C19Q3_100.zip)
This compressed file includes thirteen ASCII text file matrices in Emme batchin format. They include four rows of headers, with the remaining trip records formatted:

Origin D_1:$Trips_1$ D_2:$Trips_2$ D_3:$Trips_3$ D_4:$Trips_4$.

The CMAP network assignment process requires the truck vehicle trips (actual matrix contents) to be converted to vehicle equivalents. The factors used are b-plate and light trucks=1 vehicle equivalent, medium trucks=2 vehicle equivalents, and heavy trucks=3 vehicle equivalents.

Because of the unique way the transit network is coded, the transit trip tables cannot be assigned directly to the transit network. Please contact me if you need more information on this topic.

File Contents and Control Totals

Filename Format	Contents	2015
mf1.txt P/A	Home-based work auto person trips	6,100,106
mf2.txt P/A	Home-based other auto person trips	9,941,131
mf3.txt O/D	Non-home based auto person trips	5,505,308
mf4.txt O/D	B-plate Truck vehicle trips	2,382,695
mf5.txt O/D	Light Truck vehicle trips	281,695
mf6.txt O/D	Medium Truck vehicle trips	262,266
mf7.txt O/D	Heavy Truck vehicle trips	454,200
mf8.txt O/D	Auto Point of Entry vehicle trips	255,445
mf9.txt O/D	Truck Point of Entry vehicle trips	109,476
mf10.txt O/D	Airport Trip vehicle trips	69,538
mf14.txt (P/A)	Home-based work transit person trips	692,193
mf42.txt (P/A)	Home-based other transit person trips	602,211
mf43.txt O/D	Non-home based transit person trips	186,724

These totals existed within the model. Rounding of decimals while exporting may cause your totals to vary slightly.

Transit Skims (e.g., transitskimsc19q3100) – pick:

- mf822.txt (indexed transit in-vehicle minutes (peak))
- mf823.txt (indexed transit walk transfer minutes (peak))
- mf838.txt (indexed transit total wait time x2 (peak))

Again, the file name may vary. Please refer to Data description file to pick the correct file.

Off-Peak Data:

Distributed Trip Tables folder (named as tripsc19q3100 for e.g.) – pick:

- mf42.txt (Home-based other transit person trips)
- mf43.txt (Non-home based transit person trips)

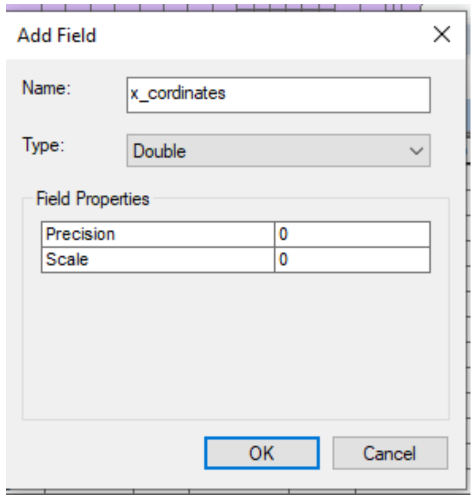
Transit Skims (namely for e.g. transitskimsc19q3100) – pick:

- mf922.txt (indexed transit in-vehicle minutes (off-peak))
- mf923.txt (indexed transit walk transfer minutes (off-peak))
- mf938.txt (indexed transit total wait time x2 (off -peak))

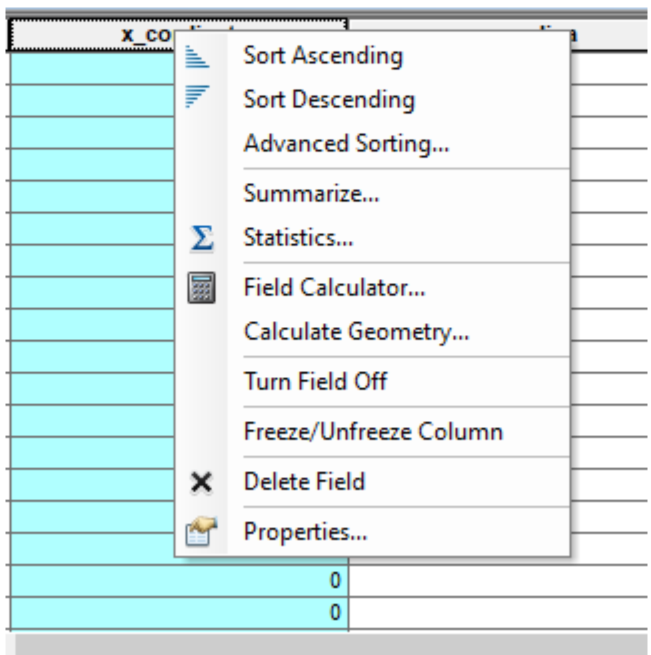
Getting the Coordinate File

After downloading Modeling Zone Systems file select zone17 folder that contains the shape file of the modeling zones geography. Open the shape file in ArcMap and follow the steps below:

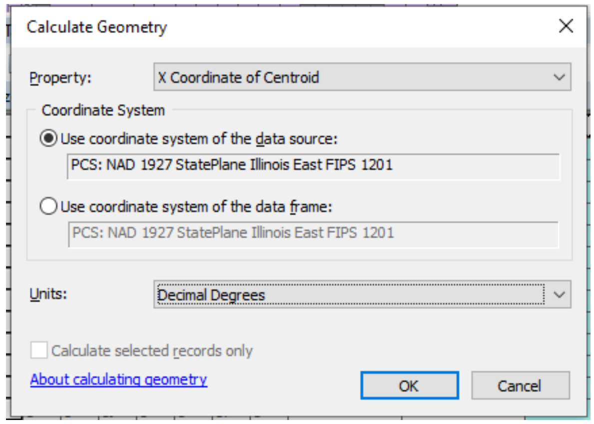
- Open the attribute table of the shape file.
- Create 2 fields, one for longitude (x coordinates) and one for latitude (y coordinates) – ee picture below for the x coordinates; repeat for the y coordinates.



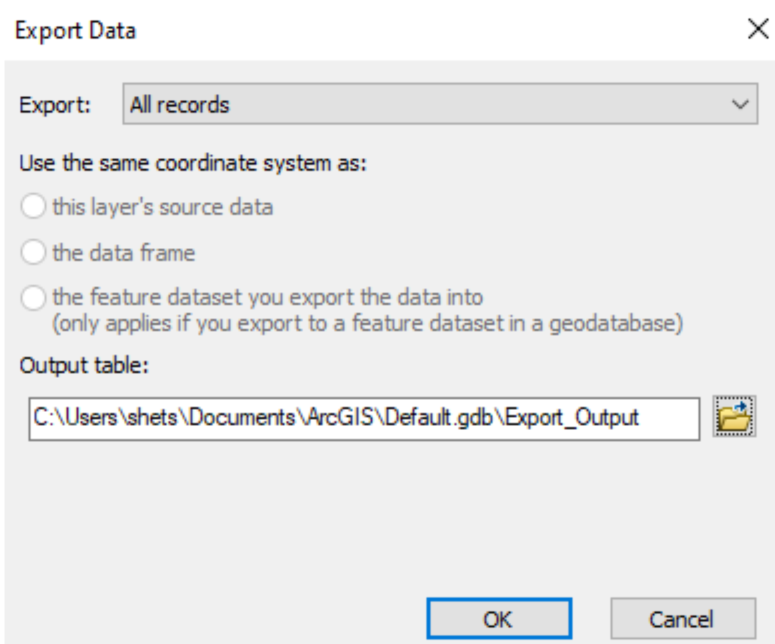
- Once both the fields are created, right click on each of the fields. You will be shown options like the one below:



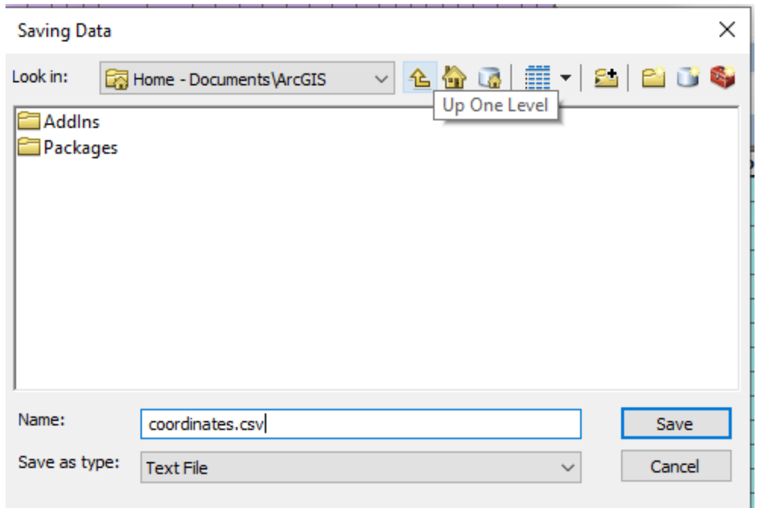
- Select the option "Calculate Geometry". This will open a dialogue box as follows. Fill in details as shown below (for y coordinates, select the property Y Coordinate of Centroid).



- The previous step will populate the two additional fields with the x, y coordinate values of each modeling zone centroid. Then export the table to a csv file. The export option below is found in the main option menu of the attribute table and will export all the fields.



- Save the file as text and add extension as .csv as shown below. Next use any tool like excel to delete extra fields, i.e. keep only the zone, xcordinates, ycordinates columns.



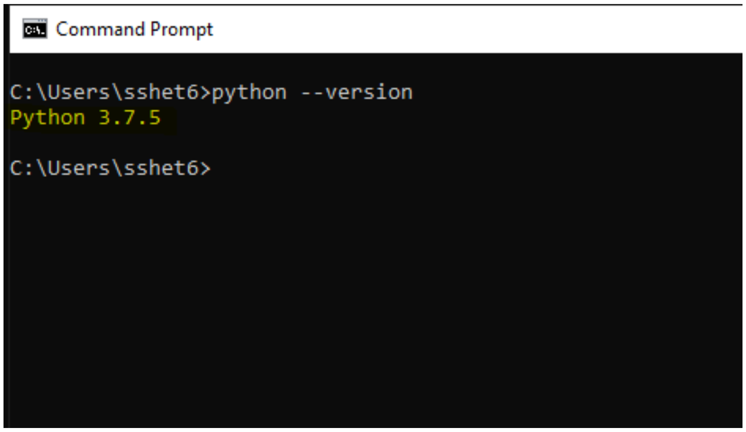
Program Prerequisites

The python program cannot run unless the following libraries are installed:

- **python – 3.7.5 version or above**

Test if already installed:

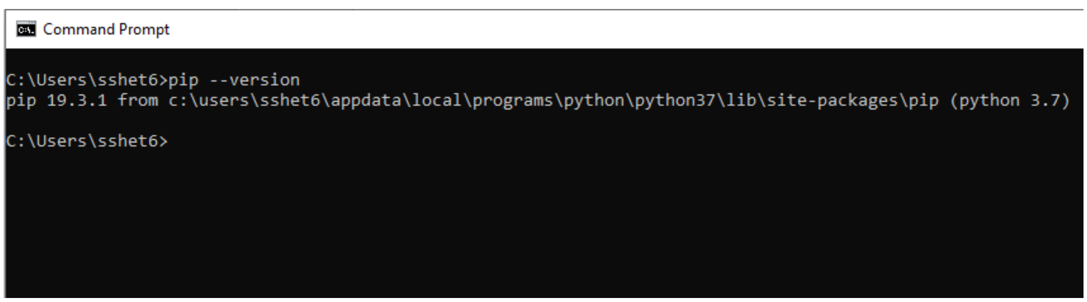
a. Open command prompt. In Windows this can be done by pressing **Windows**+R to **open the** “Run” box. Type **“cmd”** and then click “OK” to **open** a regular **Command Prompt**. In Ubuntu one can open by shortcut Ctrl - Alt + T.
b. Type into the command prompt the following: **“python –version”** and press Enter. If python3 is installed this should give the version as below. Else you will get command not found error.



- **pip – (19.3.1 or above) package installer for Python. Used to install any python libraries**

Test if already installed:

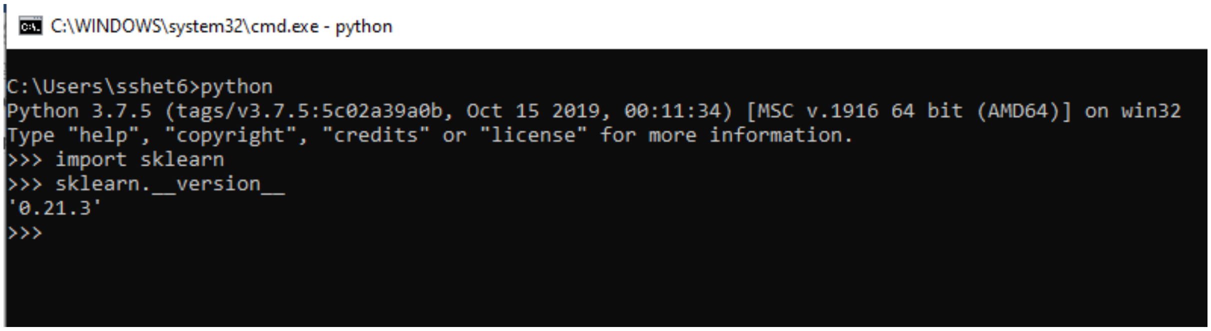
Open command prompt and type in **“pip --version”**. Same as above scenario, if the library is already installed, the version will be printed.



- **Sklearn(Scikit-learn) – (0.21.3 or above) python libraries for scientific computation**

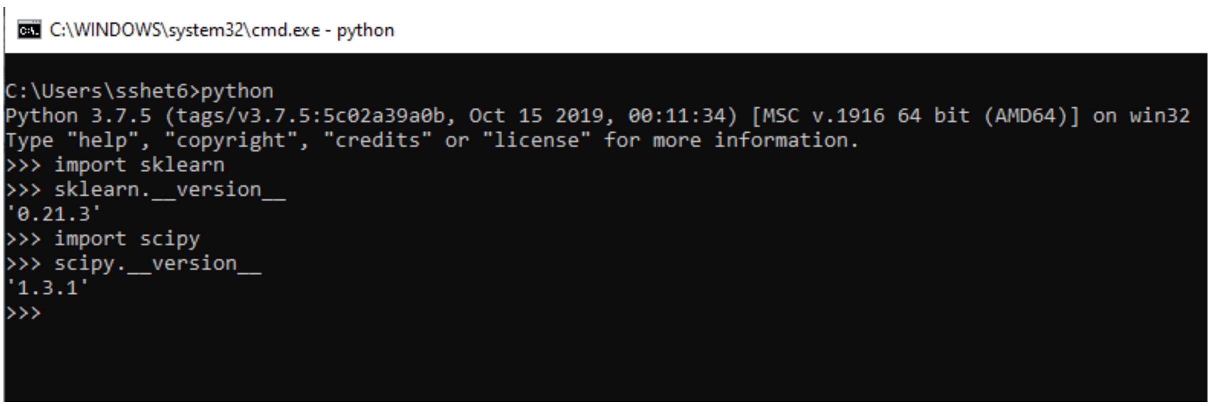
Test if already installed:

a. Open command prompt and type in **“python”**. The python command line opens as >>>.
b. Type **“import sklearn”**. If the library is already installed , at this point nothing will be printed else you get the error **“ModuleNotFoundError: No module named: sklearn”**
c. Type **“sklearn.__version__ “** to check the version of sklearn.



- **scipy – (1.3.1 or above) libraries for scientific computation**

In the same python command prompt (if python is already installed) type **"import scipy"**. As above, if the library is installed you get no output. Next type **"scipy.__version__"** to check the version.



- **pandas – (0.25.1 or above) library for data analysis, statistics, visualization.**

In the same python command prompt (if python is already installed) type **"import pandas"**. As above, if the library is installed you get no output. Next type **"pandas.__version__"** to check the version.

- **numpy – (1.17.3 or above) python library for efficient array computations, modeled after Matlab.**

In the same python command prompt (if python is already installed) type **"import numpy"**. As above, if the library is installed you get no output. Next type **"numpy.__version__"** to check the version.

- **xlsxwriter – (1.2.2 or above)Python module used to data to multiple worksheets in an Excel 2007+ XLSX file**

In the same python command prompt (if python is already installed) type **"import xlsxwriter"**. As above, if the library is installed you get no output. Next type **"xlsxwriter.__version__"** to check the version.

```
C:\WINDOWS\system32\cmd.exe - python

C:\Users\sshet6>python
Python 3.7.5 (tags/v3.7.5:5c02a39a0b, Oct 15 2019, 00:11:34) [MSC v.1916 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>> import sklearn
>>> sklearn.__version__
'0.21.3'
>>> import scipy
>>> scipy.__version__
'1.3.1'
>>> import pandas
>>> pandas.__version__
'0.25.1'
>>> import numpy
>>> numpy.__version__
'1.17.3'
>>> import xlsxwriter
>>> xlsxwriter.__version__
'1.2.2'
>>>
```

INSTALLATION OF PYTHON LIBRARY (WINDOWS)

This step can be skipped if all the required packages as mentioned above are already installed.

python (3.7.5 version or above)

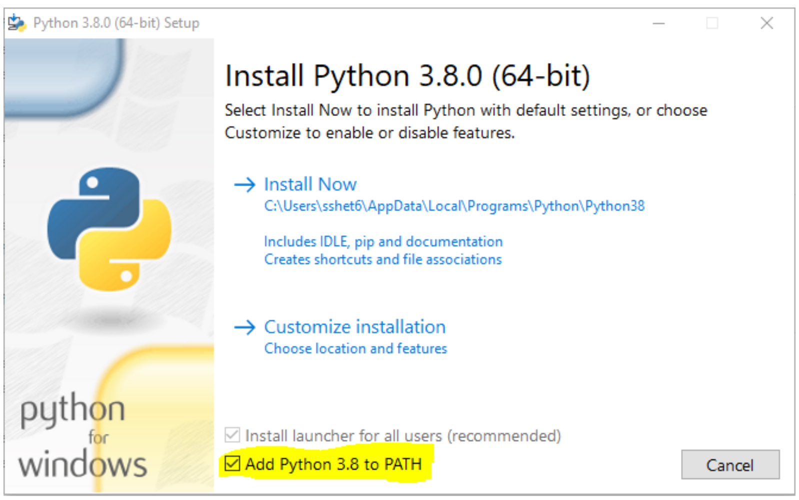
- **Step 1: Download the Python 3 Installer**
    a. Click on the link https://www.python.org/downloads/windows/
    b. Underneath the heading at the top that says **Python Releases for Windows**, click on the link for the **Latest Python 3 Release - Python 3.x.x**. (As of this writing, the latest is Python 3.7.5)
    c. Scroll to the bottom and under the section that says Files, select either **Windows x86-64 executable installer** for 64-bit or **Windows x86 executable installer** for 32-bit. (If you're unsure which version to pick, go with the 64-bit version.)

Files

Version	Operating System	Description	MD5 Sum	File Size	GPG
Gzipped source tarball	Source release		e18a9d1a0a6d858b9787e03fc6fdaa20	23949883	SIG
XZ compressed source tarball	Source release		dbac8df9d8b9edc678d0f4cacdb7dbb0	17829824	SIG
macOS 64-bit installer	Mac OS X	for OS X 10.9 and later	f5f9ae9f416170c6355cab7256bb75b5	29005746	SIG
Windows help file	Windows		1c33359821033ddb3353c8e5b6e7e003	8457529	SIG
Windows x86-64 embeddable zip file	Windows	for AMD64/EM64T/x64	99cca948512b53fb165084787143ef19	8084795	SIG
Windows x86-64 executable installer	Windows	for AMD64/EM64T/x64	29ea87f24c32f5e924b7d63f8a08ee8d	27505064	SIG
Windows x86-64 web-based installer	Windows	for AMD64/EM64T/x64	f93f7ba8cd48066c59827752e531924b	1363336	SIG
Windows x86 embeddable zip file	Windows		2ec3abf05f3f1046e0dbd1ca5c74ce88	7213298	SIG
Windows x86 executable installer	Windows		412a649d36626d33b8ca5593cf18318c	26406312	SIG
Windows x86 web-based installer	Windows		50d484ff0b08722b3cf51f9305f49fdc	1325368	SIG

- **Step 2: Run the Installer**

Once you have chosen and downloaded an installer, simply run it by double-clicking on the downloaded file. A dialog should appear that looks something like this:



Important Note. You want to be sure to check the box that says Add Python 3.x to PATH as shown to ensure that the interpreter will be placed in your execution path.

Then just click **Install Now**. That should be all there is to it. A few minutes later you should have a working Python 3 installation on your system.

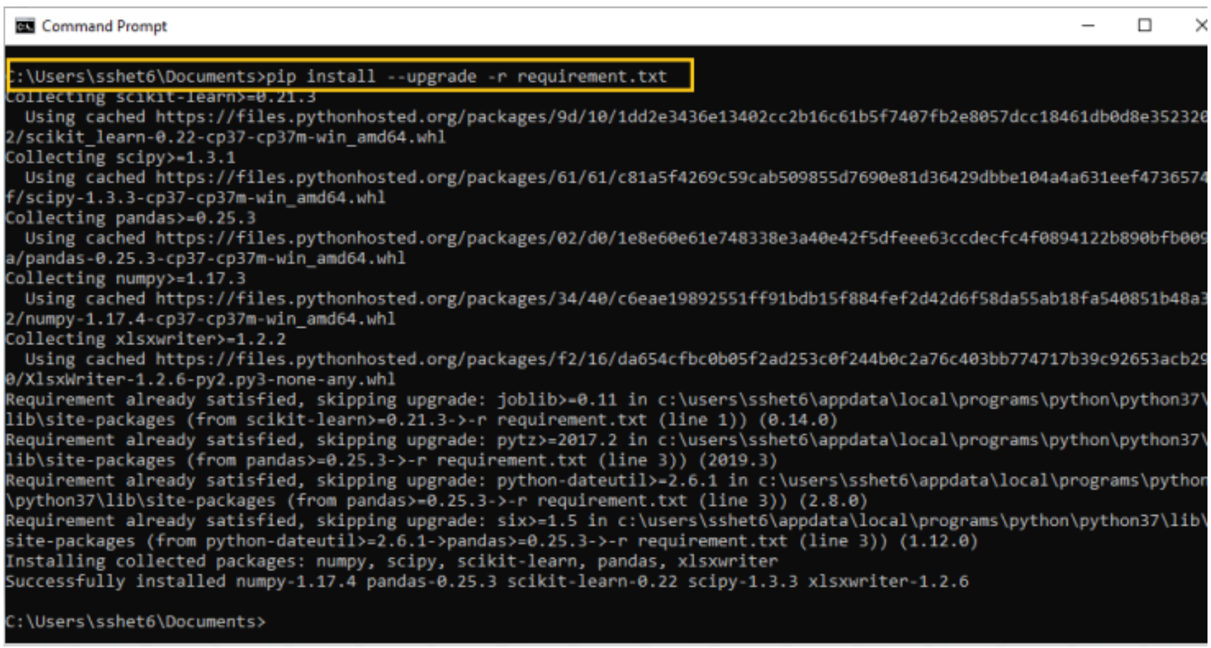
pip – (19.3.1 or above)

The latest python3 installers for Windows install pip automatically but make sure that you have clicked the checkbox “Add Python 3.x to PATH” during python installation.

Rest libraries can be easily installed using pip command as follows:

- Place the **requirement.txt** file in the folder of interest e.g. Documents.
- Open the command prompt and type the following command “**cd Documents**” to navigate to the folder containing the **requirement.txt**. Then type the following:

pip install --upgrade -r requirement.txt. The installation starts:



Note that the requirement.tx file is simply a list of the following libraries need to run the program: scikit-learn, scipy, pandas, numpy, xlsxwriter, openpyxl, geopy. If these libraries are already installed then typing “import” and the library name, while inside the python environment, you get no output; if not, you get the error “ModuleNotFoundError: No module named: sklearn”. If the latter happens, get out of the python environment by simply typing “exit()” and then use the pip installer to install each of these libraries, e.g., typing at the command prompt: “pip install” and the library name.

EXECUTING THE CODE

1. All the data files downloaded in the step 2: GETTING THE DATAFILES and the code should be placed in a single folder. So the folder should contain following files:

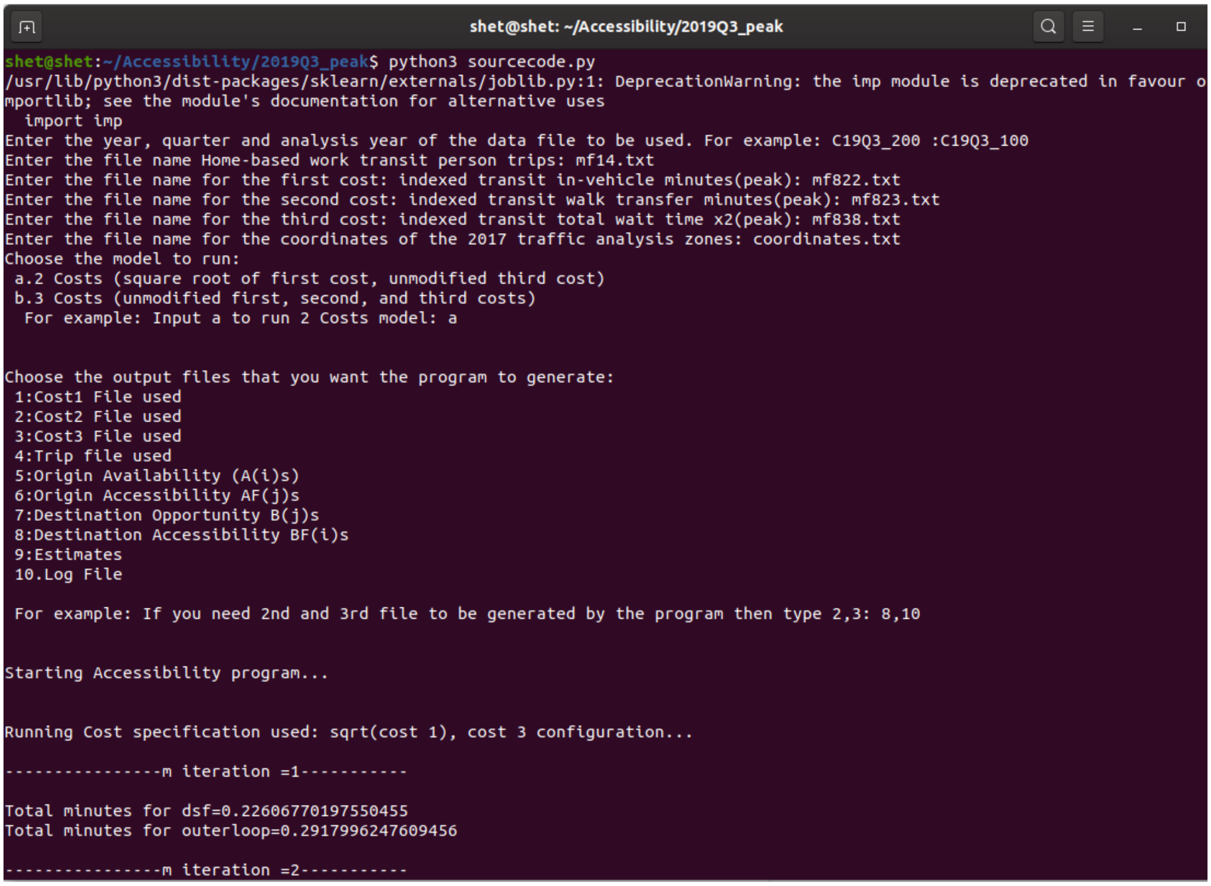
PEAK DATA:

a. mf14.txt (Home-based work transit person trips)
b. mf822.txt (indexed transit in-vehicle minutes (peak))
c. mf823.txt (indexed transit walk transfer minutes (peak))
d. mf838.txt (indexed transit total wait time x2 (peak))
e. coordinates.txt (that contains the coordinates of Traffic analysis zones) get coordinate file
f. source_code.py (python program)
g. helper_functions.py (python program that contains helper functions)

OFF-PEAK DATA:

a. mf42.txt (Home-based other transit person trips)
b. mf43.txt (Non-home based transit person trips)
c. mf922.txt (indexed transit in-vehicle minutes (off-peak))
d. mf923.txt (indexed transit walk transfer minutes (off-peak))
e. mf938.txt (indexed transit total wait time x2 (off -peak))
f. coordinates.txt (that contains the coordinates of Traffic analysis zones) get coordinate file
g. source_code.py (python program)
h. helper_functions.py (python program that contains helper functions)

2. Go to the folder created in step 1 that has all data files and code. Click on source_code.py to run the file. Enter the required parameters as prompted by the program and press enter. You can ignore the warnings. The program starts running as follows:



OUTPUT OF CODE

The program will create **result.xlsx** – This file holds all the iteration calculations, accessibility indices, cost matrices used, estimates as requested.

REFERENCES

- *CMAP DATA HUB* https://datahub.cmap.illinois.gov/
- *PYTHON* - https://www.python.org/
- *PIP* - https://pypi.org/project/pip/
- *Scikit-learn* - https://scikit-learn.org/stable/
- *Pandas* - https://pandas.pydata.org/
- *Numpy* - https://numpy.org/
- *Scipy* - https://www.scipy.org/
- *Guide to installation of python* - https://realpython.com/installing-python/