

## **Rochester Public Transit Accessibility Analysis - RMAPI**

Simon Business School, University of Rochester

Team 2 : Jiawei Gao, Yining Jia, Hanyang Jin, Longying Lai, Yuhe Zhang

Dec 12, 2023

# Table of Content:

<b>Table of Content:</b> .....	<b>2</b>
<b>1 Executive Summary</b> .....	<b>3</b>
1.1 Background and Study Objectives:.....	3
1.2 Methodology:.....	3
1.3 Key Findings:.....	3
<b>2 Methodology</b> .....	<b>4</b>
2.1 Tools Selection.....	4
2.2 Public Transit Analysis and Accessibility Analysis.....	6
2.3 Accessibility Analysis Based on Neighborhood Profile.....	7
2.3.1 Demographic Data Selection.....	7
2.3.2 Poverty Classification:.....	7
2.3.3 Accessibility Analysis by Income.....	8
2.4 Accessibility Differences by Weekday and Weekend.....	11
<b>3 Findings</b> .....	<b>12</b>
3.1 Public Transit and Accessibility Analysis.....	12
3.2 Accessibility Analysis by Poverty Level.....	17
3.3 Accessibility Analysis by Income and Underserved Area.....	18
3.4 Accessibility Analysis by Weekdays and Weekends.....	23
<b>4 Limitations</b> .....	<b>26</b>
4.1 Data Limitations:.....	26
4.2 Scope of Data:.....	26
4.3 Temporal Boundaries:.....	26
<b>Work cited</b> .....	<b>27</b>

# 1 Executive Summary

This study presents a comprehensive analysis of public transit accessibility in Rochester, focusing on its impact on accessing various essential resources. The executive summary outlines the background, objectives, methodology, and key findings of the research.

## 1.1 Background and Study Objectives:

### Background:

The study delves into Rochester's public transit system, aiming to evaluate its effectiveness in facilitating access to critical resources.

### Objectives:

The primary goal is to assess the accessibility provided by the public transit system to various resources and analyze the correlation between transit accessibility and socio-economic factors in different neighborhoods.

## 1.2 Methodology:

### Model and Accessibility Calculation:

Utilized General Transit Feed Specification (GTFS) data and ArcGIS Network Analyst for modeling the transit network. Conducted a network analysis to measure accessibility to eight resource types: Banks, Education, Emergency Food, Food, Healthcare, Jobs, Retail, and Parks. Key inputs included origins (like census tracts), destinations (resource locations), network data sources, and time constraints. Outputs included total destinations accessible, percentage of accessible destinations, and accessibility at various time thresholds.

### Accessibility Analysis:

Analyzed public transit and resource distribution using ArcGIS mapping.

Integrated demographic data from the American Community Survey to examine correlations with public transit metrics. Employed Python and Tableau for data processing and visualization, highlighting relationships between median household income and resource accessibility.

## 1.3 Key Findings:

1. Accessibility varies for different resources, with central areas generally having better access.
2. High-poverty areas showed better accessibility to all 8 types of resources via public transit than low-poverty areas.
3. Lower-income communities had greater access to emergency food and retail services.
4. Notable differences in accessibility to emergency food, retail, and healthcare facilities between weekdays and weekends.
5. Identified specific areas with limited access to certain resources, suggesting the need for targeted improvements in public transit.

## 2 Methodology

### 2.1 Tools Selection

The quantified ability to reach desired destinations, owing to density or ease of travel, is one consideration for an accessibility metric (Handley et al., 2019). To quantify the accessibility to the desired destinations for Rochester, the public transit system and street needed to be modeled in a way that variables such as time, date, and location can be controlled.

One popular solution for modeling transit networks is through General Transit Feed Specification (GTFS) files, which consist of a system of text files corresponding to transit trips, service calendars, route geometry, and other attributes (Higgins et al., 2022). The GTFS was introduced in 2005 as part of a collaboration between Google and the Portland, Oregon, public transit agency (TriMet). To facilitate data sharing and access to information for users, Google defined a publishing standard for transit agency operational data (e.g., stops, stop times, routes) (Fortin et al., 2016). The GTFS consists of a series of text files that describe public transit information. Each file models a specific aspect of the public transit information, including transit stops, routes, trips, and other schedule data. Besides static data, GTFS also supports real-time data that provides up-to-date information about the current arrival and departure times for transit lines. It'll be more complex to assess GTFS realtime data than GTFS static data. For this purpose, only GTFS static is used in this analysis.

To analyze the accessibility for public transit using the GTFS data, a model needs to be built to quantify the accessibility provided by the transit system. The Network Analyst component of ArcGIS Pro is a tool that enables users to generate a transportation network using GTFS data. A network is a system of interconnected elements, such as edges (lines) and connecting junctions (points), that represent possible routes from one location to another (ArcGIS, What is the arcgis network analyst extension? 2023). By modeling travel paths with a network, it is possible to perform analysis on the movement of agents such as persons, buses, or oil on a network. Finding the shortest paths between two points is the most common network analysis.

To create a network dataset, three types of network sources need to participate: Edge feature sources, junction feature sources, and turn feature classes. Edge is the path the agent is traveling along to, such as roads or rail lines. Junction join edges, such as street intersections or train stations. Turn stores information that can affect movement between two or more edges, such as restricted turns. The network dataset for this study is constructed using the GTFS static data of RGRTA (Rochester-Genesee Regional Transportation Authority) provided by open mobility data, and the NYS streets data provided by NYS GIS Clearinghouse. The RGRTA transit system is connected to the NYS street data to build a network in which agents can travel both by foot and by bus.

After the network dataset is constructed, network analysis can be conducted. ArcGIS Network Analyst extension supports six types of solvers that allow you to perform analysis on transportation networks, such as finding the best route across a city, finding the closest

emergency vehicle or facility, identifying a service area around a location, or servicing a set of orders with a fleet of vehicles (ArcGIS, Network analyst solvers 2023). The routing solvers in the ArcGIS Network Analyst extension utilized Dijkstra's algorithm for finding the shortest routes.

To calculate the accessibility between sets of origins and destinations, the Transit Network Analysis Tools are utilized. The Transit Network Analysis Tools is a toolset developed by Esri developers to further enhance the network analysis. The Transit Network Analysis Tools include the function calculating Calculate Accessibility Matrix. Calculate Accessibility Matrix solves an Origin-Destination Cost Matrix analysis incrementally over a time window and summarizes the results. It can be used to count the number of jobs accessible to a set of origins within a reasonable commute time (Morang, 2019). Given a set of origins and destinations, the Calculate Accessibility Matrix tool calculates the number and percent of destinations reachable from each origin by public transit and by foot within the time constraint. The tools calculate the accessibility by solving the Origin-Destination Cost Matrix multiple times for the time period to account for the changing schedule throughout the day.

The key inputs for the model include:

**Origins:** A point or polygon feature that the agent departs from. For example, the origin can be a point, a parcel, or a census tract. When using a polygon as an origin, the centroid for the polygon is used. Using a census tract, for instance, does not necessarily yield an accurate result as a census tract sometimes contains a large area while the accessibility is calculated from the centroid of the census tract.

**Destination:** The point or polygon feature that the agent will travel to. For example, a census tract with many jobs located in it or hospitals.

**Network Data Source:** The network dataset or service URL that includes the network dataset. This refers to the network dataset that the calculation will be based on.

**Cutoff Time:** The maximum allowed time for the agents to travel in the network.

**Cutoff Time Units:** The units for the cutoff time. It can be second, minutes, hours, or days

**Start Day (Weekday or YYYYMMDD date):** The start day for the calculation. This can be a day in a week or an exact day.

**Start Time (HH:MM) (24 hour time):** The start time for the calculation.

**End Day (Weekday or YYYYMMDD date):** The end day for the calculation. This can be a day in a week or an exact day.

**End Time (HH:MM) (24 hour time):** The end time for the calculation.

**Time Increment (minutes):** Increment the OD Cost Matrix's time of day by this amount between solves. For example, if this amount is 1 minute, the OD Cost Matrix will be solved for every 1 minute for the time window.

**Destinations Weight Field:** This will be the weight for the destination. For instance, if the destination includes 50 jobs, adding jobs as the weighted field will count the destination as 50 destinations reachable.

The key outputs for the model include:

**Total Dests:** The total number of destinations reachable by the origin in the time window. For example, if Origin 1 can reach Destination A within 30 minutes at any time of day but can only

reach Destination B within 30 minutes if the travel starts at exactly 10:03 AM, Destination A and Destination B still each contributes equally to TotalDests(Melinda Morang, Transit Network Analysis Tools User's Guide 2023).

**Perc Dests:** The percentage of the destinations reachable from the origin.

**DsAL10Perc, DsAL20Perc, ..., DsAL90Perc:** The total number of destinations reachable from this origin within the time limit at least x% of start times within the time window.

**PsAL10Perc, PsAL20Perc, ..., PsAL90Perc:** The total percentage of destinations reachable from this origin within the time limit at least x% of start times within the time window.

For the accessibility analysis, we collected data from various sources as the data input. The data source for this study includes:

**Origins:** 111 Rochester Census tracts

**Destination:** Bank and Credit Union, Education, Emergency Food, Food, Healthcare, Jobs, Retail, Park in Monroe County

**Cutoff Time:** 45

**Cutoff Time Units:** Minutes

**Start Day (Weekday or YYYYMMDD date):** Monday/ Sunday

**Start Time (HH:MM) (24 hour time):** 7 AM

**End Day (Weekday or YYYYMMDD date):** Monday/ Sunday

**End Time (HH:MM) (24 hour time):** 7 PM

**Time Increment (minutes):** 1 minute

The Accessibility Matrix is calculated for every point of interest for Monday and Sunday. The accessibility to jobs is only calculated for Monday due to Monday being a more common commute day.

## 2.2 Public Transit Analysis and Accessibility Analysis

GTFS data of Rochester are used to generate bus stop points and frequency data on the map. The frequency and distribution are mapped on the ArcGIS map. The frequency is measured for both weekdays and weekends and is averaged at the number of runs per hour.

To study the distribution of the key resources of Rochester, eight types of points of interest were incorporated in this study. Those eight types of points of interest were: Bank and Credit Union, Education, Emergency Food, Food, Healthcare, Jobs, Retail, and Park. We mapped those points of interest in ArcGIS using geolocation data to study the distribution.

## 2.3 Accessibility Analysis Based on Neighborhood Profile

### 2.3.1 Demographic Data Selection

To assess how well public transportation services meet the needs of residents, especially in high-poverty areas, the demographic features of census tracts are included in the analysis. The neighborhood profile is constructed with data from the American Community Survey (ACS), focusing on the years 2017-2021 to capture recent socioeconomic trends. The ACS datasets

employed include Median Household Income (2022), Vehicle Availability (B08201, 2017-2021), and Poverty Status (S1701, 2017-2021). The data are analyzed at the census tract level, which enables a detailed exploration of neighborhood characteristics and identifies variations within Rochester.

The selected ACS datasets provide a spectrum of economic indicators: population, income levels, vehicle accessibility, poverty rates, and employment status. These variables characterize the socioeconomic conditions of the neighborhoods, influencing residents' quality of life and accessibility of resources.

The study further integrates this neighborhood profile with the GTFS data and public transportation accessibility results to all eight types of interests. By comparing the ACS data with public transit metrics, we aim to identify correlations between the socioeconomic status of neighborhoods and the efficiency of public transportation systems.

### 2.3.2 Poverty Classification:

An area is considered in persistent poverty if it had a poverty rate of 20.0% or higher during the three decades period from 1989 to 2015-2019".(Benson, Bishaw, & Glassman, 2023) This definition applies to both county and sub-county geographies, including census tracts. Rochester does not have an area classified as "persistent poverty", which indicates the lack of policy-based classification for poverty level. To investigate the accessibility of public transportation among populations with varying degrees of poverty, the Jenks natural breaks are applied to classify the census tracts. The Jenks natural breaks optimization method was employed for its efficacy in revealing inherent patterns within skewed geographic data distributions. This method minimizes the variance within classes and maximizes the variance between classes, ideal for identifying natural groupings in data (Chen et al., 2013). Census tracts are manually segmented into two categories based on poverty rates: high-poverty and low-poverty areas.

In the next step, by averaging the accessibility to public transportation across populations with different levels of poverty, a comparison was derived between areas of high poverty and low poverty.

### 2.3.3 Accessibility Analysis by Income

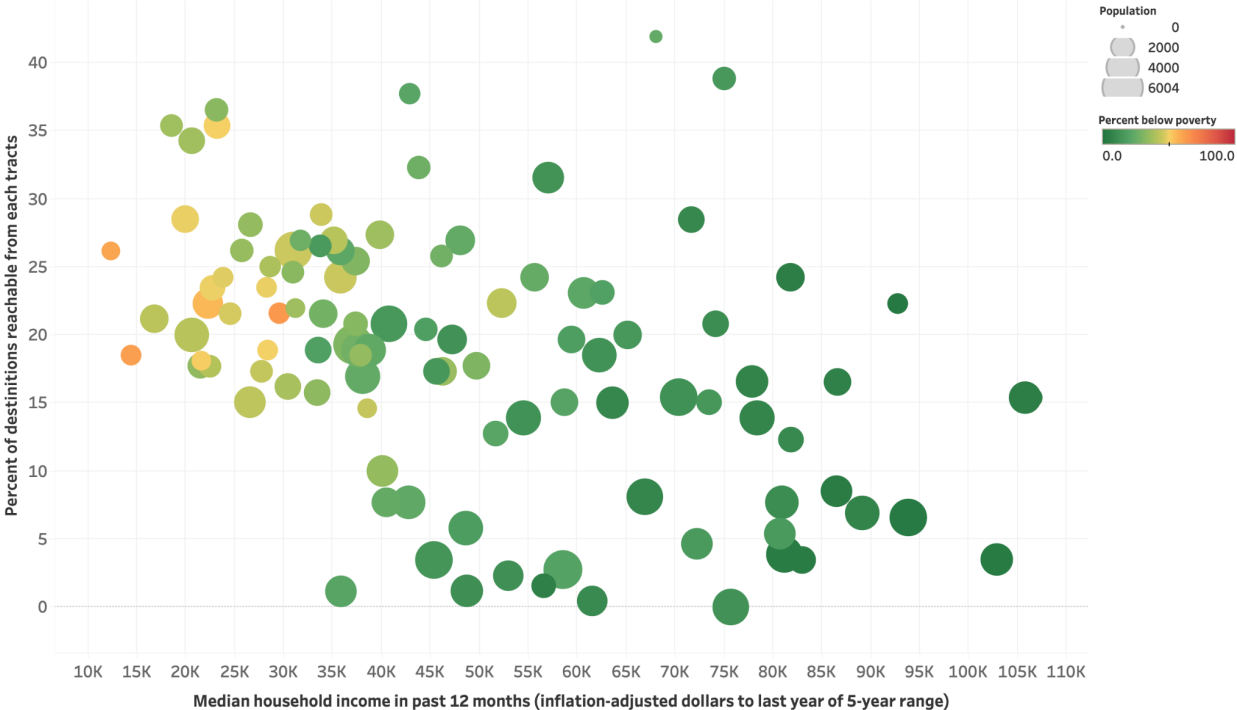
For data processing, we created a dataset for each type of resource analysis by merging demographic metrics (population, percentage below poverty in each community, unemployment rate, poverty classification using the Jenks method) with accessibility results generated by our ArcGIS model (TotalDests, PercDests, DsAL...Perc, PsAL...Perc) using Python. In total, we compiled eight distinct datasets.

We primarily utilized Tableau to create visualizations that more effectively illustrate the relationship between median household income and the accessibility of eight types of resources.

Each census tract is plotted on the graph, with the x-axis representing median household income and the y-axis indicating the key metric from our ArcGIS Model, as stated in the report before. This metric indicates the percentage of destinations reachable at least 50% of start times during the time window.

Green dots signify wealthier areas, while yellow and orange dots denote poorer communities. The larger the dots, the greater the population within the census tract. We generated this graph eight times, once for each type of resource.

### Bank and Credit Union Accessibility



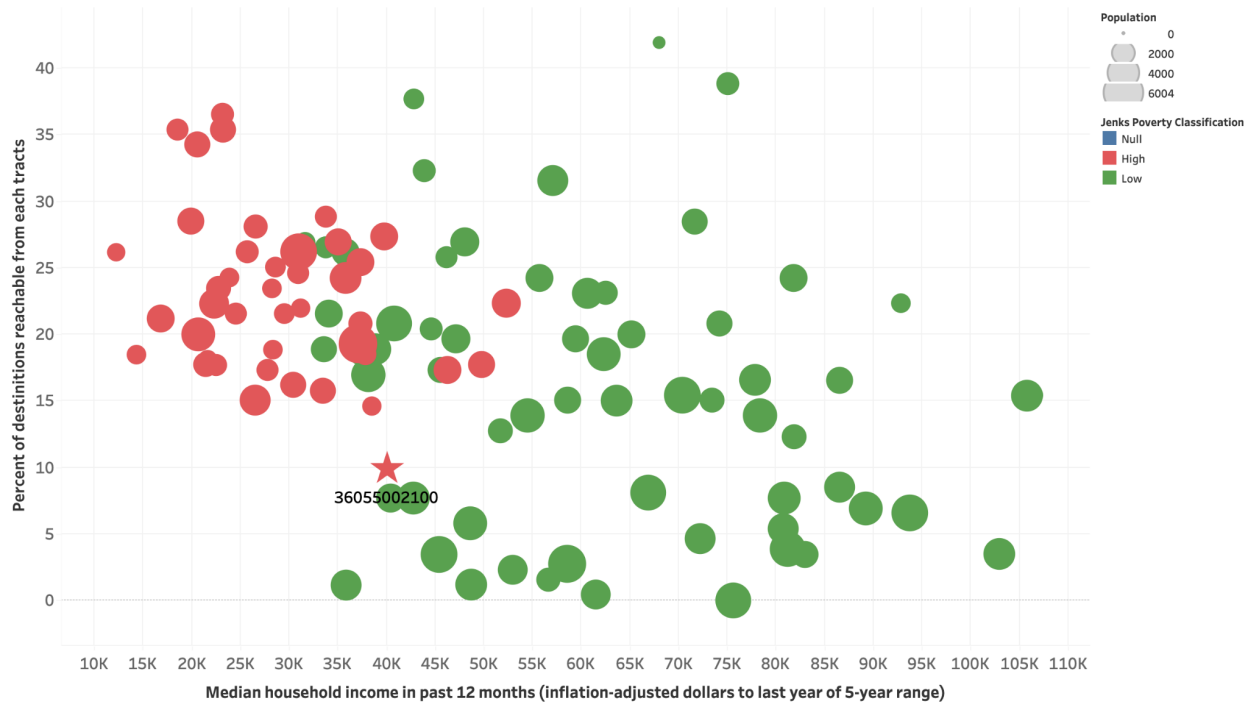
Since public transportation is likely more crucial in low-income areas because of the limited vehicle ownership, we are considering a more in-depth analysis targeting those communities. We have employed the Jenks method to segment all census tracts. This helps us gain a better understanding of communities affected by poverty. As indicated in the graph, the red dots represent census tracts identified as high-poverty areas. We will delve deeper into these specific tracts.



## Bank and Credit Union



## Bank and Credit Union



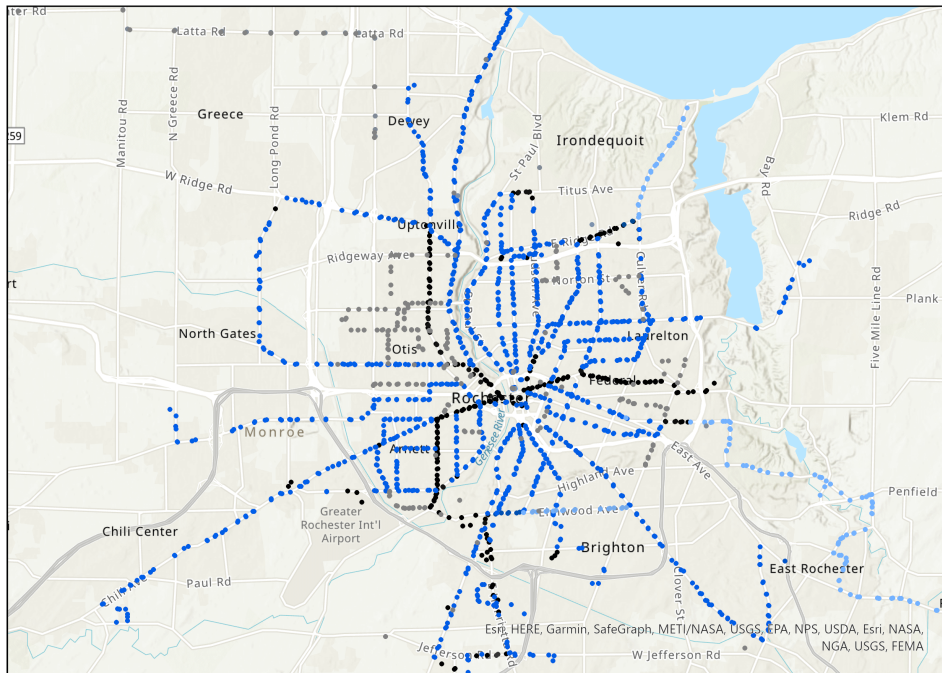
and adds depth to the analytical findings. This spatial analysis is instrumental in providing a clearer picture of the regional dynamics at play.

### 3 Findings

#### 3.1 Public Transit and Accessibility Analysis

The public transit system of Rochester is in a radioactive shape where multiple lines intersect at the RTS transit center, in downtown Rochester. The alignment of darker dots with the orange line validates the higher frequency of the frequent lines, matching the officially displayed information by RTS. The weekday and weekend maps illustrate the disparity in stop frequency, revealing increased frequency on weekdays for many routes. This insight is valuable for understanding the dynamics of public transportation usage throughout the week.

The frequency of stops on weekdays:

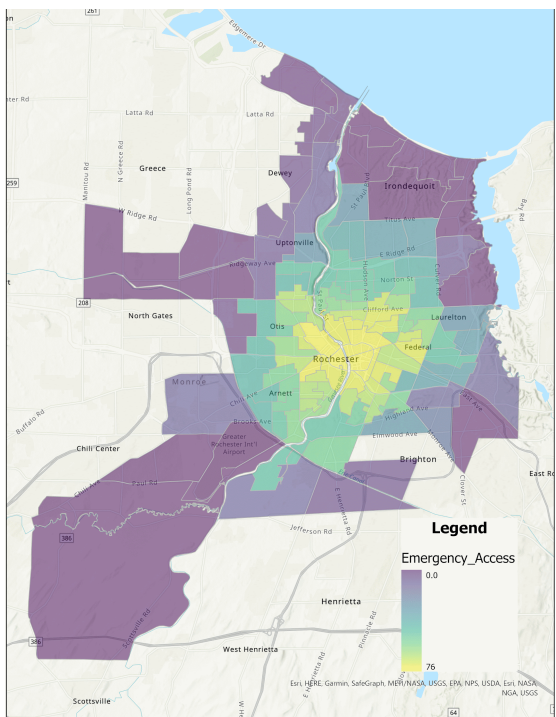
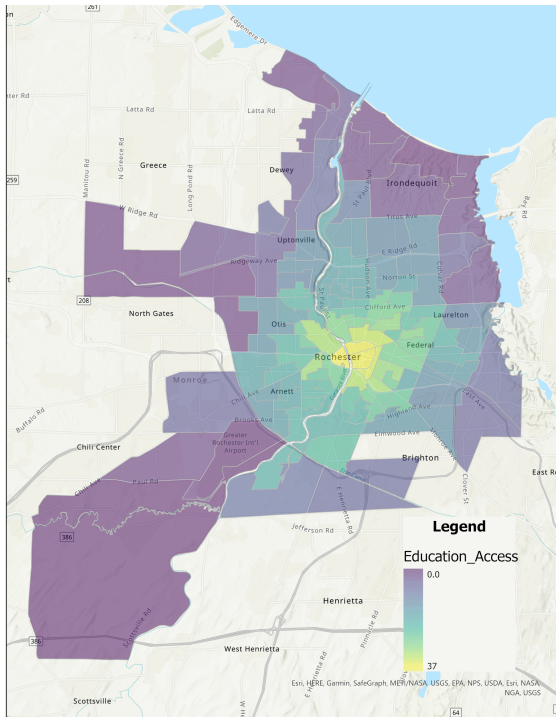
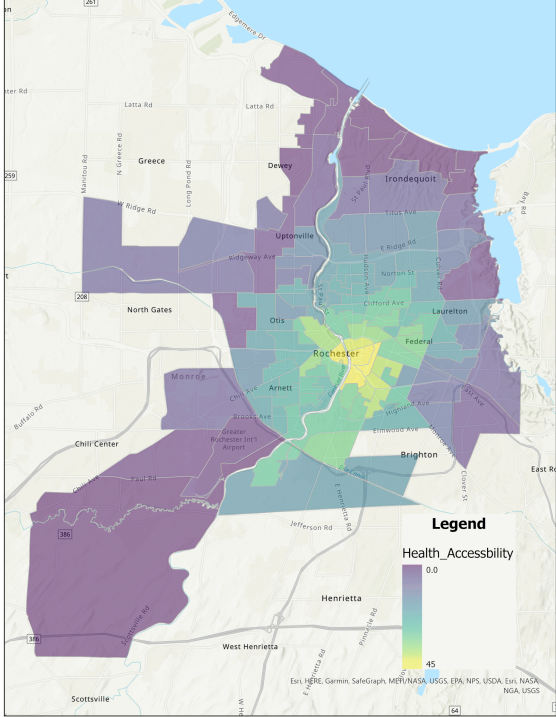
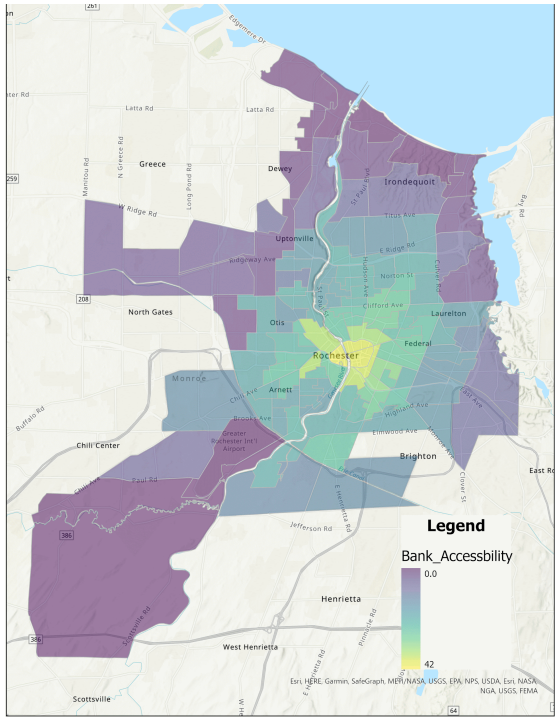


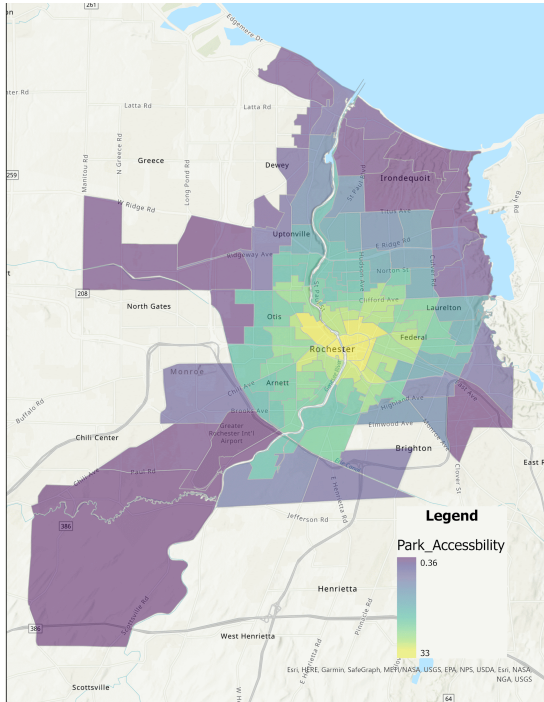
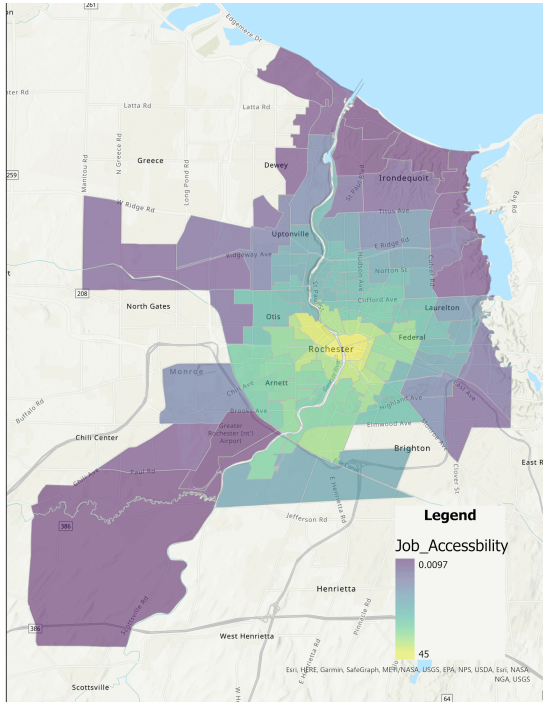
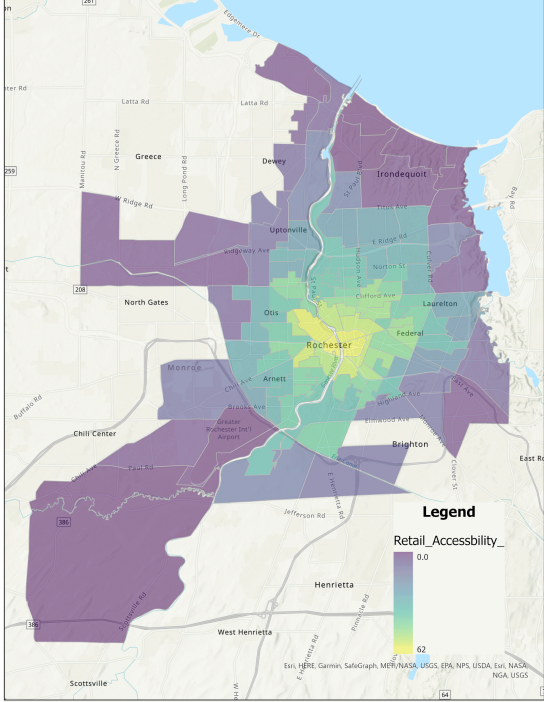
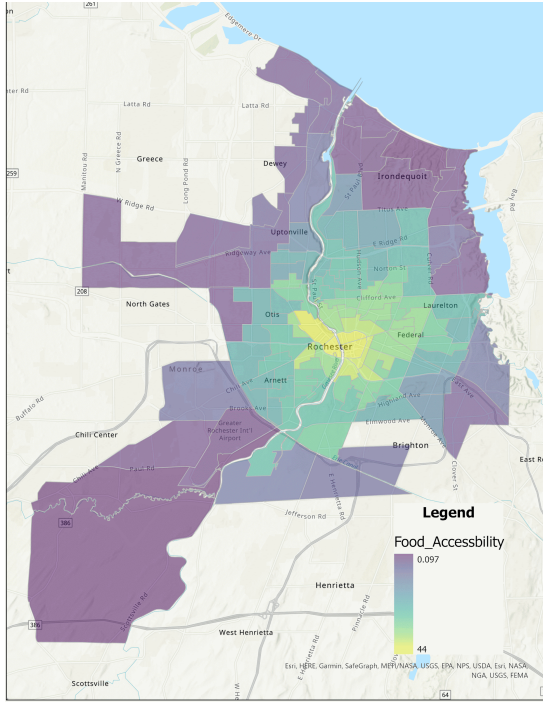
The frequency of stops on weekends:



This section provides various analyses focused on resource accessibility, with an emphasis on the graphical representation of spatial distribution. It features a color-coded map where lighter hues indicate higher levels of accessibility. The map reveals a pattern of diminishing accessibility from the city center outward, and this accessibility varies during weekdays, introducing a time-related aspect to the analysis.

The spatial distribution of each resource type, including banks, health services, jobs, and parks, is examined. This analysis uncovers significant trends, such as a higher density of banks in the downtown area of Rochester, increased accessibility to health services in the southern regions, and a central distribution of parks. These preliminary observations lay the groundwork for a more comprehensive study, which is expected to yield insights into how public transportation influences resource accessibility in Rochester.



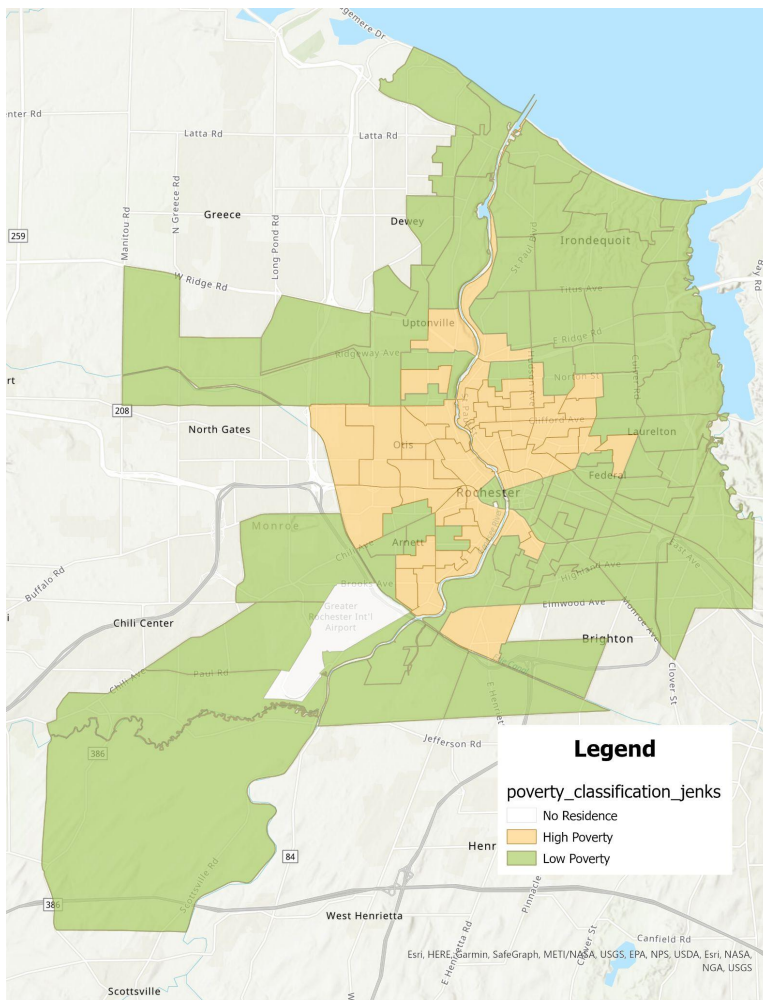


### 3.2 Accessibility Analysis by Poverty Level.

The delineation of poverty zones was conducted employing the Jenks natural breaks classification method:

Jenks natural break result: [0.0, 0.27417342482844664, 1.0]

Thresholds were established at zero to approximately twenty-seven percent (27%) for regions categorized as low poverty. In contrast, any value exceeding this benchmark was indicative of high-poverty areas. Among all 111 census tracts in Rochester, 45 were classified as high-poverty areas, 64 as low-poverty areas, and one census tract was deemed inapplicable to the analysis owing to an absence of residents, as shown in the chart below:



By calculating the average accessibility to 8 points of interest for high-poverty areas and low-poverty areas, the outcome demonstrates that: high-poverty areas exhibited superior accessibility to the specified points of interest via public transportation, surpassing that of the low-poverty regions:



### Average Accessibility to all 8 Points of Interest by Poverty Level

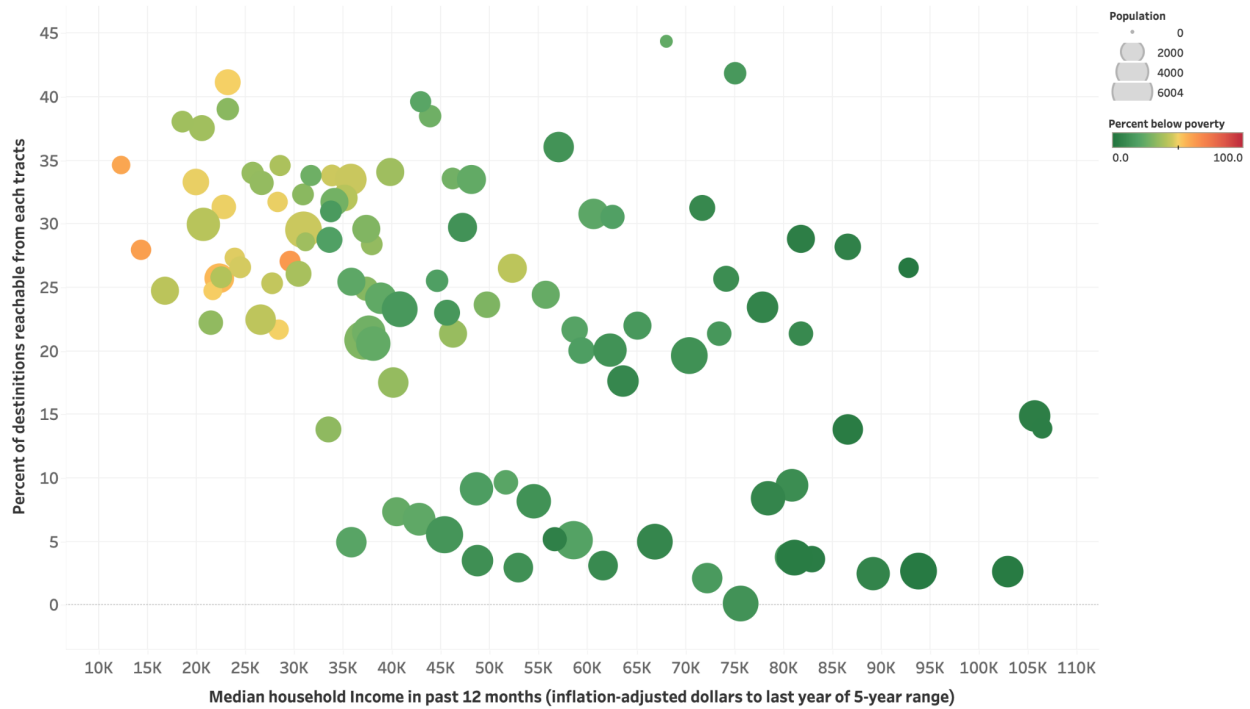


### 3.3 Accessibility Analysis by Income

Moving forward, we will conduct a thorough accessibility analysis for eight different types of resources.

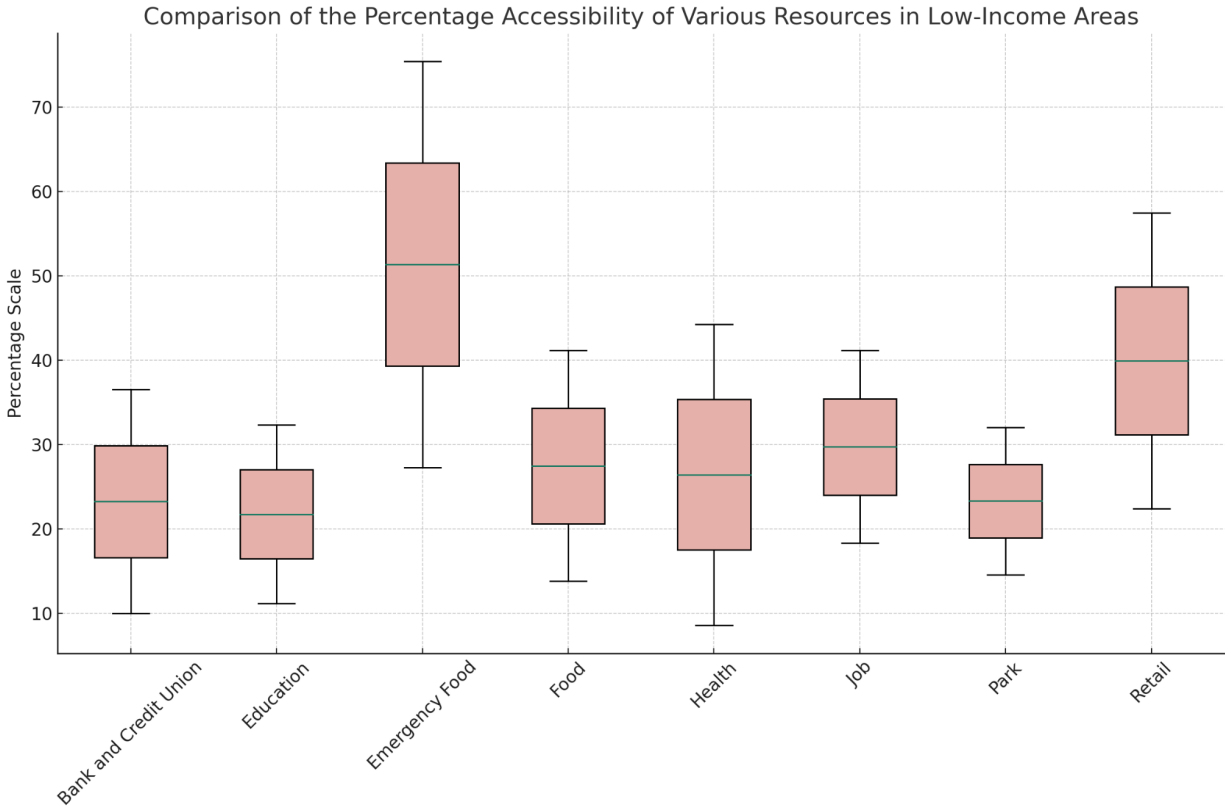
The initial insight from the overall distribution is readily apparent: Rochester offers a reasonable public transportation layout. Areas with enhanced accessibility to essential resources via public transport are often associated with higher poverty rates and lower median household incomes. In contrast, wealthier neighborhoods tend to have less access to public transportation. Such an arrangement suggests that public transportation can better assist those in the Rochester area who are most in need of accessing resources.

## Food Accessibility



Another insightful finding is that communities with lower incomes have greater access to emergency food and retail services compared to other resources. To compare the accessibility scale of 8 types of resources, we plot both the maximum and minimum accessibility percentages for all communities affected by poverty using a box plot. From this graph, we can determine which resources are most accessible to impoverished communities.

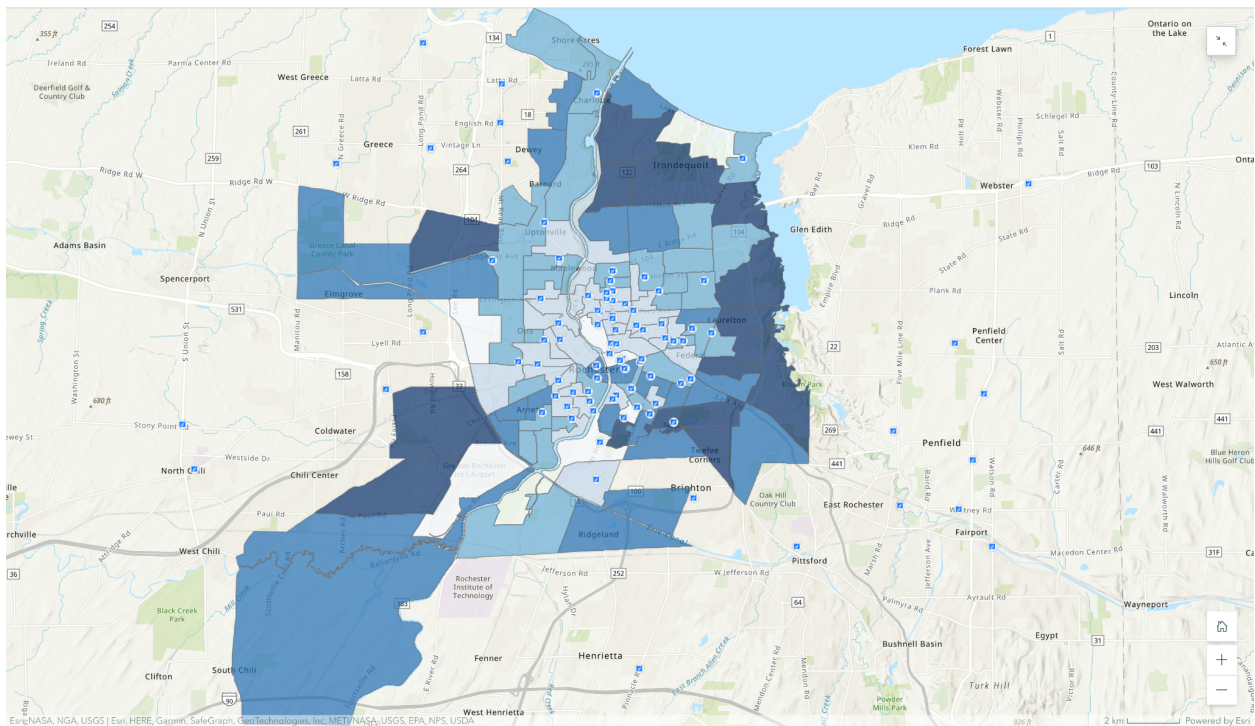
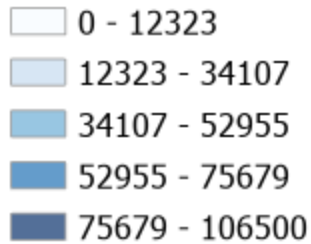
For example, areas struck by poverty have the highest accessibility to emergency food, with all poverty-stricken areas in Rochester having at least 35% and the most 75% of all resources can be reached within the time window. And for retail stores 25-60%, higher than other resource types. For retail stores, accessibility ranges from 25-60%, which is higher compared to other resource types. All the other resource types usually range from 10%-40%.



Regarding emergency food, we discovered that most emergency food sites are located in low-income communities, which are represented by the gray areas on the map. This finding explains the high accessibility to emergency food in these poverty-stricken areas.

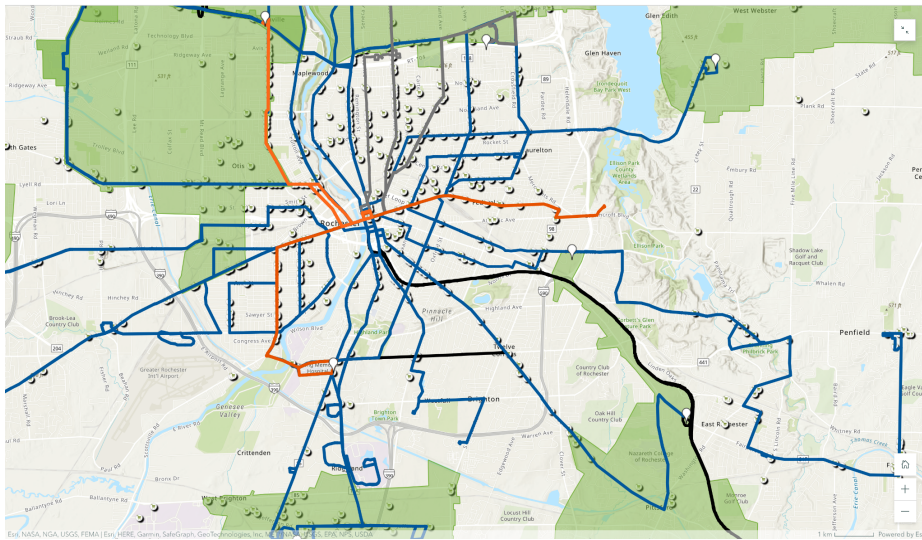
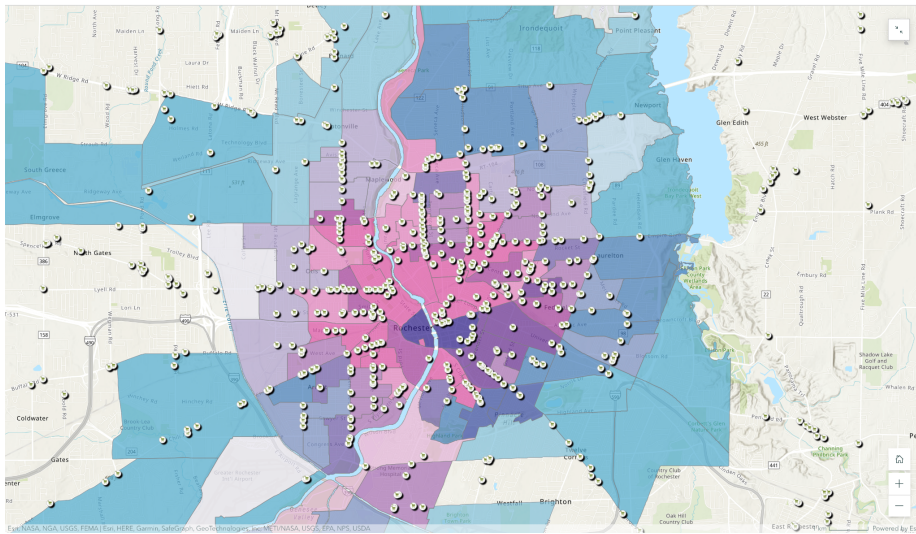
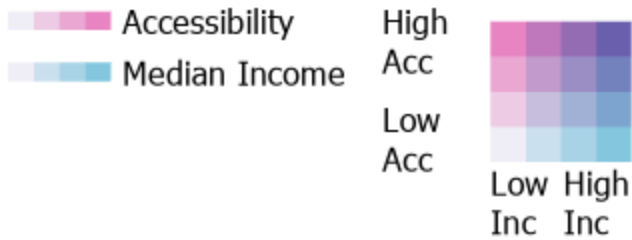
## Legend

### Median Household Income



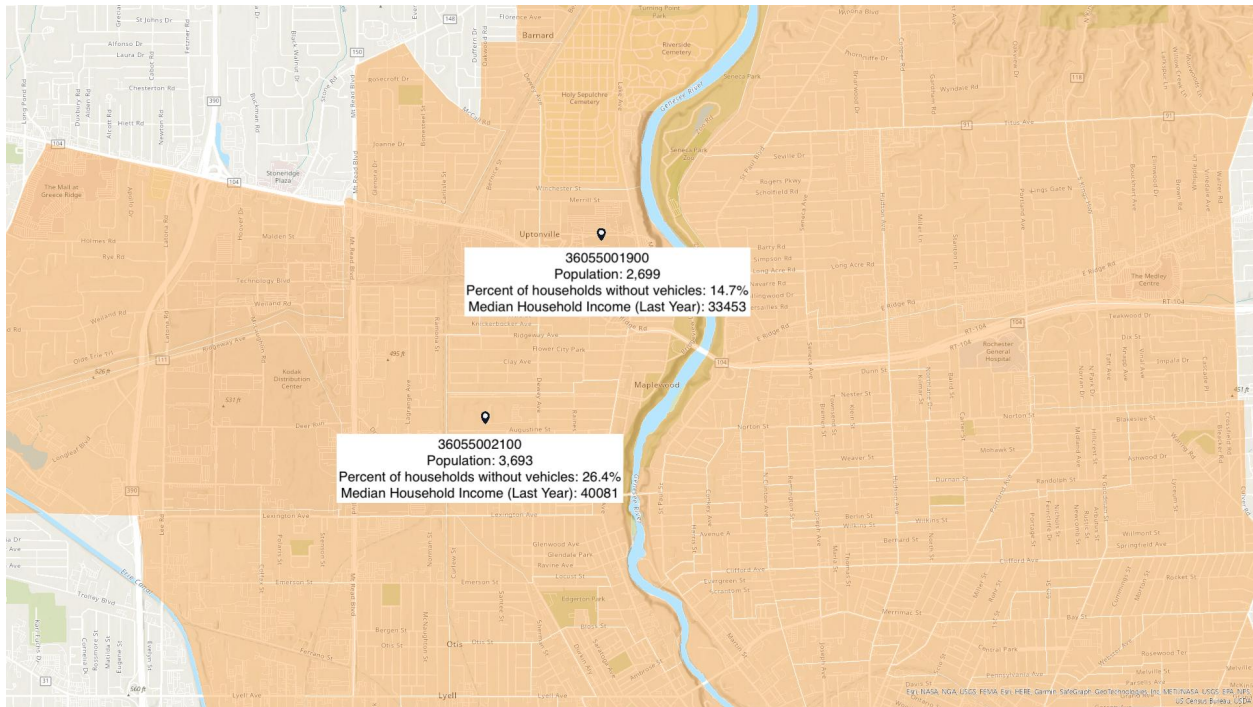
In terms of retail resources, there are two noteworthy observations on the map. Firstly, retail resources are predominantly concentrated in the downtown area, aligning with communities marked by low income yet high accessibility, which are those magenta areas as indicated in the legend. Secondly, the distribution of retail resources is closely aligned with public transit routes, with most retail stores situated near stations. Those two findings support the observation of high accessibility to retail resources.

## Legend



We identified two specific areas that are potentially underserved. Area 36055002100 has limited bank and credit union, health, and park accessibility, while area 36055001900 shows weaknesses in emergency food and food accessibility. Both of the regions are located to the west of the river. We conducted a basic profile analysis of these two neighborhoods and found

that 26.4% and 14.7% of households do not own vehicles, respectively. We suggest that further analysis be conducted on potential underserved areas to enhance public transportation services in Rochester.

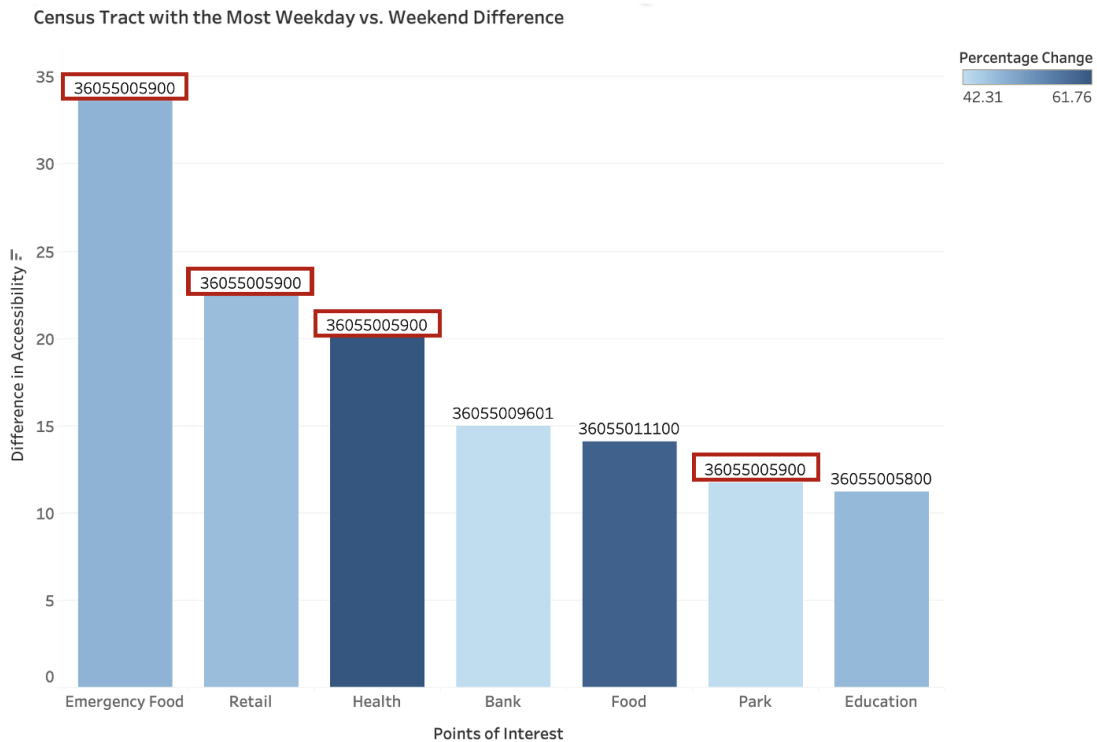


### 3.4 Accessibility Analysis by Weekdays and Weekends

Points of Interest	Mean of Weekday	Mean of Weekend	Difference of Mean
Emergency Food	42.4734	34.7748	7.69861
Retail	30.1384	23.0709	7.06747
Health	19.8026	14.5607	5.24187
Bank	18.5863	14.2134	4.37283
Food	22.4997	18.4247	4.07509
Education	16.3698	12.7937	3.5761
Park	17.2902	14.1547	3.13552
Job	22.0939	N/A	N/A

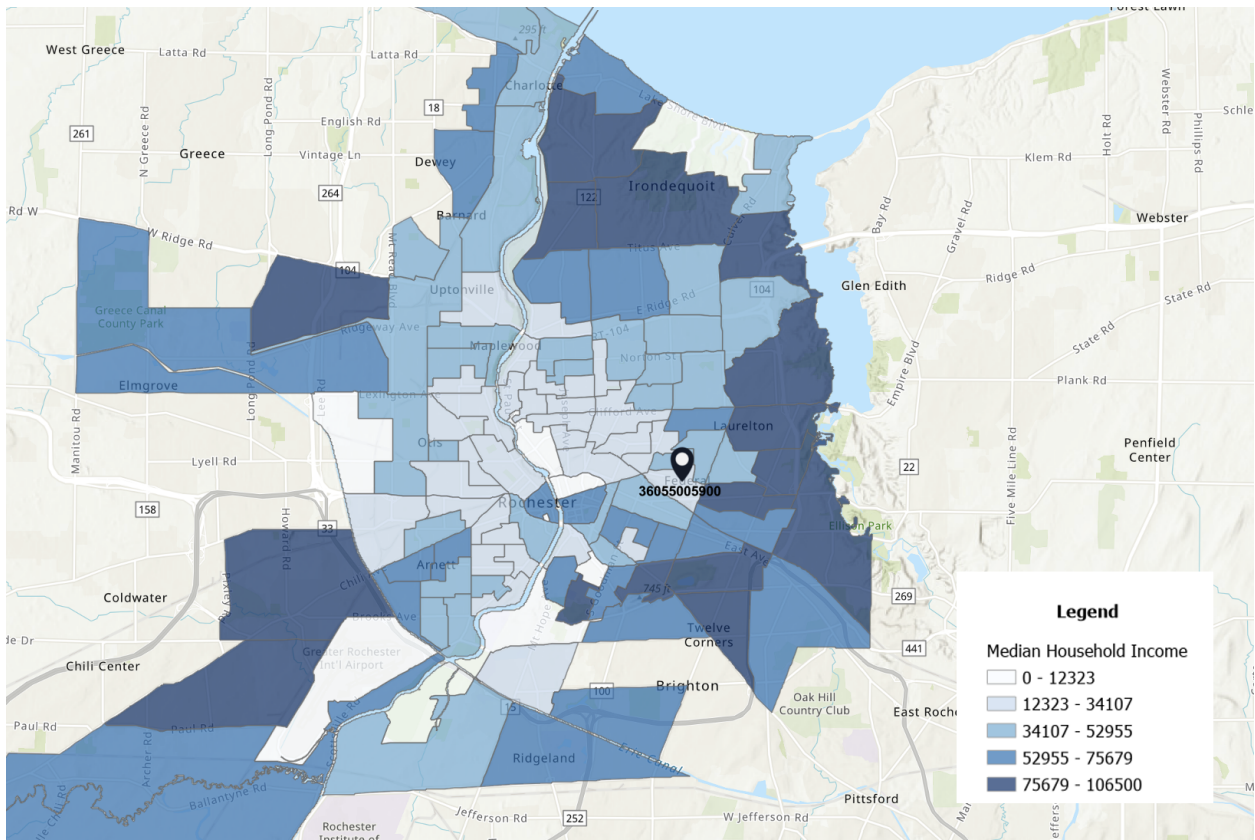
Points of Interest	Median of Weekday	Median of Weekend	Difference of Median
Emergency Food	48.1818	36.3636	11.8182
Retail	33.3725	22.7967	10.5758
Health	19.7452	14.3312	5.41401
Food	24.9027	19.7471	5.15564
Bank	18.8462	13.8462	5
Education	16.5394	12.4682	4.07125
Park	18.8612	15.3025	3.55872
Job	23.2726	N/A	N/A

The tables presenting mean and median values reveal clear patterns in accessibility, highlighting emergency food services, retails, and healthcare facilities as key points of interest with significant differences in accessibility between weekdays and weekends. This observation underscores a shift in accessibility trends during the weekends.



The included bar chart provides a visual representation of these variations across various categories, with a focus on census tract 36055005900 as the area most profoundly impacted. In

this chart, the use of darker shades indicates higher percentages of change, illustrating the effects on numerous points of interest.



In transitioning to the geographical map, Census Tract 59, recognized as a Low-Income Community Opportunity Zone in Rochester, New York, exposes distinct characteristics contributing to its singular accessibility challenges. The median household income, approximating \$31,000, conspicuously lags by 57% in comparison to the state average, thereby accentuating socio-economic disparities that exert influence on accessibility. These findings underscore the exigency for targeted interventions and community-centric enhancements, particularly within census tract 36055005900.

The imperative of effective collaboration between local authorities and community organizations becomes evident in addressing the distinct challenges confronted by low-income communities. Such collaborative endeavors are instrumental in ensuring the implementation of equitable measures, thereby fostering heightened accessibility to essential resources.



## 4 Limitations

### 4.1 Data Limitations:

The research primarily relies on General Transit Feed Specification (GTFS) data and demographic information from the American Community Survey (ACS). While these sources are comprehensive, they may not fully encapsulate the dynamic nature of transit operations and the diverse experiences of transit users. For instance, the GTFS data used does not capture real-time delays or disruptions in transit services, which can significantly affect accessibility. Moreover, the ACS data, while detailed, might not reflect the most current socio-economic conditions, given its collection and publication cycle.

### 4.2 Scope of Data:

The study's use of static GTFS data, as opposed to real-time data, presents a limitation in capturing the day-to-day variability in transit services. Static data provides a snapshot of scheduled services but does not account for on-the-ground realities like service interruptions, temporary route changes, or seasonal variations in transit provision. This gap could lead to an overestimation or underestimation of actual transit accessibility.

### 4.3 Temporal Boundaries:

The analysis is limited to predetermined time windows, which might not fully represent the variability in transit accessibility throughout the day or across different days of the week. For example, transit accessibility during peak hours, late-night hours, or during special events could differ significantly from the study's findings. This limitation restricts the ability to generalize the results to all times and days, potentially missing critical insights into transit accessibility.

## Work cited

- Handley, J. C., Fu, L., & Tupper, L. L. (2019). A case study in spatial-temporal accessibility for a transit system. *Journal of Transport Geography*, 75, 25–36. <https://doi.org/10.1016/j.jtrangeo.2019.01.005>
- Higgins, C., Palm, M., DeJohn, A., Xi, L., Vaughan, J., Farber, S., Widener, M., & Miller, E. (2022). Calculating Place-based transit accessibility: Methods, tools and algorithmic dependence. *Journal of Transport and Land Use*, 15(1). <https://doi.org/10.5198/jtlu.2022.2012>
- Fortin, P., Morency, C., & Trépanier, M. (2016). Innovative GTFS Data Application for transit network analysis using a graph-oriented method. *Journal of Public Transportation*, 19(4), 18–37. <https://doi.org/10.5038/2375-0901.19.4.2>
- Google. (n.d.). GTFS static overview | static transit | google for developers. Google. <https://developers.google.com/transit/gtfs>
- What is the arcgis network analyst extension?. What is the ArcGIS Network Analyst extension?-ArcGIS Pro | Documentation. (n.d.). <https://pro.arcgis.com/en/pro-app/3.1/help/analysis/networks/what-is-network-analyst-.html>
- Network analyst solvers. Network Analyst solvers-ArcGIS Pro | Documentation. (n.d.). <https://pro.arcgis.com/en/pro-app/3.1/help/analysis/networks/network-analyst-solver-types.htm>
- Morang, M. (n.d.). ESRI/Public-transit-tools: Tools for working with GTFS public transit data in arcgis. GitHub. <https://github.com/Esri/public-transit-tools>
- Transit Network Analysis Tools User's Guide. GitHub. (n.d.). <https://github.com/Esri/public-transit-tools/blob/master/transit-network-analysis-tools/UsersGuide.md>
- Benson, C., Bishaw, A., & Glassman, B. (2023). Persistent Poverty in Counties and Census Tracts. U.S. Census Bureau. Retrieved from <https://www.census.gov/library/publications/2023/demo/persistent-poverty-in-counties-and-census-tracts.html>
- Chen, J., Yang, S. T., Li, H. W., Zhang, B., & Lv, J. R. (2013). Research on Geographical Environment Unit Division Based on the Method of Natural Breaks (Jenks). *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. DOI:10.5194/isprsarchives-XL-4-W3-47-2013