Tampere University

Shamsur Raza Chowdhury

PUBLIC TRANSPORT ACCESSIBILITY IN TAMPERE

Bachelor's thesis
Faculty of Built Environment
Dr. Anna-Kaisa Viitanen
Jonathon Taylor
May 2023

ABSTRACT

Shamsur Raza Chowdhury: Public Transport Accessibility in Tampere
Bachelor Thesis
Tampere University
Bachelor of Science (Technology) in Sustainable Urban Development
May 2023

Public transport is becoming an integral part of sustainable city planning and development. Promoting accessible public transport for the residents of a city can encourage faster adoption. As one of the largest and growing cities in Finland, Tampere's public transport is undergoing a phase of new demand and expansion. The accessibility of the public transport network across different regions of Tampere was thus studied with the contemporary definition of accessibility within public transportation discourse. The research used publicly available General Transit Feed Specification (GTFS) data along with GIS software to analyse the data. The analysis's outcome suggests a substantial disparity in accessibility among different areas of Tampere. To continue the sustainable growth of Tampere, the authority needs to address these shortcomings of the public transport network.

Keywords: GTFS, Public Transport Accessibility, GIS, Sustainable Transport, Tampere Public Transport

PREFACE

I am sincerely grateful to all the teachers from the Sustainable Urban Development program for their unwavering support and continuous guidance throughout the whole semester, starting with brainstorming ideas to finalizing everything for the thesis submission. In particular, I would like to thank Dr. Anna-Kaisa Viitanen for supervising me through the process of writing the thesis and reviewing the work. Finally, I would like to end by thanking my mother for checking in on my thesis from time to time and providing mental support to complete my studies.

Tampere, 28th of May 2023

Shamsur Raza Chowdhury

CONTENTS

1. INTRODUCTION 1
2. BACKGROUND 3
  2.1 Public Transport Accessibility 3
  2.2 General Transit Feed Specification (GTFS) 4
  2.3 Tampere and its public transport 5
3. METHODOLOGY 7
  3.1 Data Source 7
  3.2 Analysis 8
    3.2.1 Programming Analysis 8
    3.2.2 GIS Analysis 9
4. RESULTS 11
5. DISCUSSION 16
  5.1 Implication for the Commuters and Authority 16
  5.2 Limitations of the Study 17
6. CONCLUSIONS 18
REFERENCES 19
APPENDIX A 22

LIST OF SYMBOLS AND ABBREVIATIONS

GTFS	General Transit Feed Specification
GIS	Geographic Information System
ESRI	Environmental Systems Research Institute, Inc.
AVL	Automated Vehicle Location
CSV	Comma Separated Values
QGIS	Quantum Geographical Information System

1. INTRODUCTION

In Tampere, the second-largest city in Finland, the public transportation system plays a vital role in connecting residents to jobs, education, and other essential services. As a sprawling city with an increasing population, Tampere's demand for transportation is also growing (City of Tampere, 2023). Considering this rising demand for public transportation, it is worth investigating the performance of the public transportation system in terms of accessibility. Transport accessibility is a critical issue in urban planning, as it affects the mobility and quality of life of residents. With climate change becoming ever more imminent, it is important to cut down on carbon emissions from transportation. Mass adoption of public transport is going to help in achieving that goal.

The General Transit Feed Specification (GTFS) is a widely used public transportation data format that provides a standardized way to describe transit schedules, stops, and routes (Google Developers, 2023). Despite primarily being used for transit schedules, GTFS data contains a rich amount of information about the transport services that can be used to identify accessibility. This will give us insight into the area of the city where access to public transport is low and thus can be improved.

I aim to analyse public transport accessibility in Tampere with the help of publicly available GTFS data. The GTFS data, along with Tampere city's geographical data, will be analysed using programming and Geographical Information System (GIS) to evaluate the performance of the public transportation system in terms of accessibility, identify areas lacking sufficient transport access, and discuss the implication of transport accessibility in Tampere. The research question I will answer is how accessible public transport is across different areas of Tampere.

In this thesis, I provide a literature review of the research area, including the definition of public transport accessibility in the context of this research, an overview of General Transit Feed Specification data, and a brief description of Tampere and its existing transport system in the second chapter. In the third chapter, I describe the research methodology, including the GTFS data collection and their subsequent analysis using Python programming language and QGIS software. In the fourth chapter, I will present the results of the analysis, including the service level of the public transportation system and accessibility to the service. Finally, in Chapter 5, I will discuss the implications of the research for the residents of the city and the transport authority in Tampere.

2. BACKGROUND

In this section of the thesis, I will review the relevant literature about public transport accessibility (sub-chapter 2.1), as well as delve further into the details of GTFS (sub-chapter 2.2). This section also contains a short overview of Tampere as a city and its existing public transport infrastructure (sub-chapter 2.3).

2.1 Public Transport Accessibility

The concept of accessibility refers to the ease with which individuals can reach their intended destinations (Wachs & Kumagai, 1973). Public transport accessibility has been studied in many studies over the years to great lengths. Over the past four decades of transportation research, related but distinct terms have been employed (Gould, 1969; Makr et al., 1999). These terms, such as 'access,' 'accessibility,' 'availability,' and 'proximity,' have been used interchangeably in different studies, albeit sharing similar meanings (Church & Marston, 2003; Kawabata, 2003). There hasn’t been any consensus reached on the definition of “accessibility” in the context of public transport. Numerous metrics for assessing accessibility have been proposed (Church & Marston, 2003; Kawabata, 2003) but their measurements vary depending on the researchers' perspectives and objectives. Recent studies highlight a shift in the perspective of transportation accessibility research. While earlier studies primarily emphasized transit as a mode of reaching different urban services and activity areas, contemporary research now recognizes public transport as a crucial urban service that must be accessed (Daniels & Mulley, 2013). As a result, in this thesis, the concept of accessibility is studied from this recent perspective of accessibility *to* public transport, rather than focusing on the accessibility of public transport.

In this study, public transport accessibility is assessed primarily based on spatial and temporal dimensions. The spatial aspect involves considering the distance from each access point to public transport like a bus or tram stop. Physical proximity to those points is a fundamental element of accessibility, as it ensures that public transport is physically reachable for a significant portion of the population (Azar et al., 1994). However, proximity alone is not sufficient to define accessibility. It is also crucial to consider when and how frequently vehicles can be accessed. To that end in assessing accessibility, it is important to consider the service frequency of public transport. This is because the availability of public transport may not be perceived as satisfactory by users of public transport if they are required to wait beyond a certain threshold level (Polzin et al., 2002).

2.2 General Transit Feed Specification (GTFS)

General Transit Feed Specification (GTFS) is a widely used standardized format for public transportation schedules and associated geographic information across the globe. GTFS allows transit agencies to share their data with external developers, who can then use the data to create a myriad of transport-related applications, commonly used to offer transit directions, real-time updates, and other services. According to the GTFS specification documentation, there are two types of GTFS feeds: static and real-time. While Static GTFS is a "common format for public transportation schedules and associated geographic information", Real-time GTFS is an “extension” of the static GTFS and provides "real-time updates about public transport fleet to application developers” (Google Developers, 2023). In this thesis, GTFS always refers to static GTFS feed unless otherwise mentioned.

GTFS consists of a set of files as described in Table 2-A, typically stored in a compressed ZIP archive, that provides information about the transit agency and its routes, stops, and schedules. The files are in a CSV-like format, which can be read by software applications (Google Developers, 2023).

File Name	Description
agency.txt	Contains information about the transit agency, including its name, website, and time zone
stops.txt	Lists all the stops on each transit route, including their names, locations, and other details
routes.txt	Describes the routes taken by transit vehicles, including their names, colours, and other attributes
trips.txt	Specifies the trips made by transit vehicles on each route, including their start and end times, and other details
stop_times.txt	Provides the arrival and departure times for each stop on a given trip

calendar.txt	Describes the dates when the transit system is operating, including holidays and other exceptions
calendar_dates.txt	Specifies the dates when the transit system is not operating, including service cancellations and other changes
fare_attributes.txt	Describes the fares charged for transit service, including their prices, restrictions, and other details
fare_rules.txt	Defines the relationship between fares and transit routes, including which routes are covered by which fares

Table 2-A Description of GTFS files. Modified from (Google Developers, 2023).

GTFS data is an incredibly valuable tool for gaining deep and meaningful insights into the transportation system and improving its efficiency. GTFS data can be used to optimize routes, schedules, and services based on usage patterns and rider demand (GTFS community, 2023). With the help of data analysis, GTFS data provide useful metrics about transit demand and level of access, helping to reduce wait times and improving service for commuters (Bok & Kwon, 2016). In addition, GTFS data in combination with other data like Automated Vehicle Location (AVL) data, ambient population, etc. can be utilized to evaluate the overall performance and impacts of changes to the transit system (Bok & Kwon, 2016). For instance, the performance of service in a route or the expansion of service to new areas can be measured with these powerful data, which will allow the transport authority to allocate resources and improve their services based on concrete data (Prommaharaj et al., 2020). Using GTFS as a tool for data-driven decision-making, transport authorities can ensure that they are investing in the right areas and commuter satisfaction (Prommaharaj et al., 2020). In a nutshell, GTFS provides a comprehensive and nuanced view of the transportation system that can help transit agencies optimize their services and make public transit a more accessible and attractive option.

2.3 Tampere and its public transport

Tampere is the third biggest city in Finland and the second biggest outside of the capital region in terms of land area (Britannica, 2023). At the moment, Tampere city spans 689.6 square kilometres of area. As of early 2023, population of Tampere is 249,060 inhabitants with a population density of 465,28 persons/km^2 and a big part of the population is students due to the Tampere Universities (City of Tampere, 2023). Tampere is the biggest city in the Pirkanmaa administrative region. It borders a number of small to medium cities like Ylöjärvi, Nokia, Pirkkala, Lempäälä, Kangasala etc. Tampere is also well connected with the rest of Finland by a robust rail and road network. Outside of the capital region, Tampere is one of the few cities in Finland that is experiencing net positive migration and projecting population growth in the future (City of Tampere, 2023). Therefore, it is fair to say that Tampere as a growing city bears a lot of significance.

In Tampere, the authority responsible for managing the public transport network is Tampere Regional Transport Authority (Finnish: *Tampereen seudun joukkoliikenne*), which goes by the name Nysse. While Nysse is responsible for Tampere's public transport, it also manages public transport in some of the neighbouring cities (Nysse, 2023). In Tampere, the public transport network primarily consists of bus service and light-rail-based tram service. The latter has been introduced to Tampere fairly recently in 2021. Currently, there are only two lines of tram service active within the city with a plan of rapid expansion well underway (Tampereen Ratikka, 2023). While Nysse has multiple zones for the transport network, Tampere City falls within the A and B zone of the network. In this thesis, only the buses and the tram lines that operate within Tampere are considered for the discussion.

3. METHODOLOGY

In this study, I analysed GTFS data using Python programming language and QGIS software to evaluate the accessibility of public transportation in Tampere, Finland. In order to answer the research question, which was how accessible public transportation in different areas of Tampere, I have used a combination of quantitative and spatial technique. This approach provides a comprehensive perspective on the state of public transport accessibility with both statistical and visual data. In doing the analysis, at first, I extracted the necessary data from the GTFS feed using Python scripts, which allowed me to obtain information about the location, frequency, and headway of services. Next, in QGIS, I have visualized and analysed the spatial distribution of transit stops, their service level, and their density within geographic boundaries.

3.1 Data Source

The City of Tampere publishes their GTFS data regularly through the ITS Service Factory (Tampereen seudun joukkoliikenne, 2023). The GTFS data was collected from that archive and saved for analysis. In this analysis, GTFS data from 17.05.2023 was used given that it is a usual working day with no exception i.e., holiday, strike, extreme weather etc. The GTFS data contains 5250 trips originating from and arriving at 1293 stops located within Tampere and scheduled between 00:00 to 23:59 on the day in the entire transport network.

In addition to the GTFS data, a vector layer of the Statistical Metropolitan Area of Tampere (in Finnish: *Tampereen Suuralueet*) is collected from the City of Tampere's Open Data Service portal for identifying the regional boundaries of Tampere. The statistical regions, as seen in Figure 3-I, correspond to the "municipal subdivision" (in Finnish: *kunnan osa-alue*) classification used by Statistics Finland (City of Tampere, 2012).

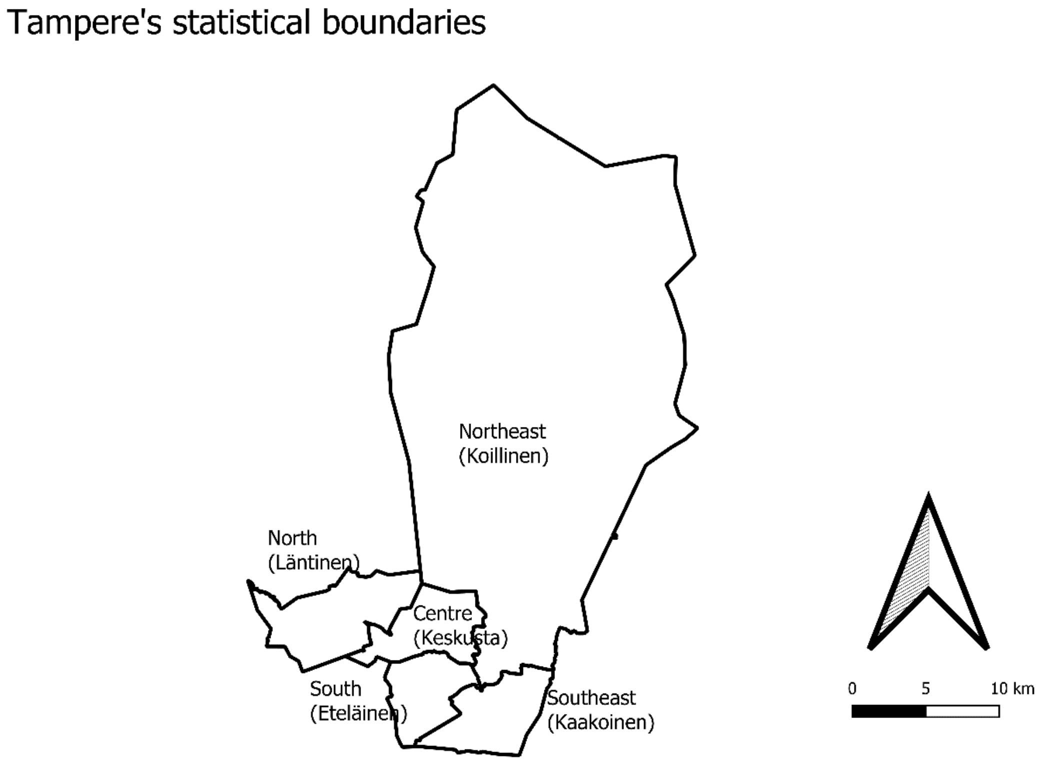


Figure 3-I Statistical areas/region of Tampere.

3.2 Analysis

The data analysis was done in two parts: using Python programming language and open-source GIS software QGIS. In 3.2.1, the GTFS dataset provides the average departure per hour and headway for every stop on the specified date. Then, in the last step, GIS software is used to form statistical and map-based results in combination with the data from the earlier step.

3.2.1 Programming Analysis

For the programming analysis, Python 3.10.2 is used. The script I wrote for the analysis can be found in Appendix A. The GTFS data is imported in Python using Pandas' library as a Pandas data frame (McKinney, 2010; Pandas development team, 2020). From *calender_dates.txt* of the GTFS data, all the *service_id* for the specific date 17.05.2023 is first collected and then all the corresponding *trip_id* representing each induvial trip scheduled is collected. Then, from *stop_times*, only the rows with *trip_id* that are scheduled for the specific date are separated. In the next step, a new data frame is created by aggregating the *stop_times* data frame based on *stop_id*, thus giving a total number of departures from every stop. From that, the average per-hour departure number is calculated. Then, the headway – which is the time between two consecutive departures from a stop – is calculated from the average per-hour departure and rounding it to the closest integer. Based on the average headway, each stop is classified into three service levels as shown in Table 3-A as a study suggest high-frequency service is assumed to have a headway of less than or equal to 10 minutes, and those services with a headway greater than 10 minutes are considered lower frequency service (Ansari Esfeh et al., 2021)

Headway (in minutes)	**Service level**
Less than 10	High
10 or more but less than 30	Medium
30 or more	Low

Table 3-A Service levels based on headway in minutes.

Finally, *stops* are joined with the *stop_times* data-frame where *stop_id* matches to create the final data-frame which contains average departure per hour and average headway for every stop along with all information about the stop from the original *stops.txt* GTFS file.

As the last step, using the Python Geo Pandas library, the Pandas data frame is converted into a geo data-frame (Jordahl et al., 2020), which is then used to create an ESRI shape file of point layers representing the stop and all the columns as attribute (See *Appendix A*). This shape file is then later used in GIS software for further analysis.

3.2.2 GIS Analysis

The Statistical Metropolitan Areas of Tampere from the City of Tampere comprise five regions, namely: Northeast (*Koillinen*), South (*Eteläinen*), Centre (*Keskusta*), Southeast (*Kaakkoinen*), and West (*Läntinen*). At first, by using QGIS 3.28.3 LTS version, the total area of each of these regions is calculated using the field calculator and *area* function.

All the stops with the calculated frequency, headway, and service level are imported as another vector layer. The stops that are contained within the different regions are separated into their layers. This is done using the extract by location processing tool in QGIS with points that are within the polygon layer set as the parameter. In this step, all the stops that are not part of these regions of Tampere are excluded.

In the next step to find the area served by the stops, I used the Valhalla plugin in QGIS which uses the Open Street Map service's road network to find points reachable within 5 minutes of walking (Nolde, 2020). Due to the computationally intensive nature of this task, all the stops within 100 meters of each other are clustered using the DBSCAN clustering tool and then using the aggregate tool assigned the service level of the clustered stop based on the majority service level among all the stops in the cluster (QGIS Documentation, 2020). To find the geometrical centre of the clustered stop, I used the centroid tool in QGIS. Then, with the clusters of stops, using the Isochrone Pedestrian tool of the Valhalla plugin, the service area – which is the area reachable within 5 minutes of walking - for each of the clustered stops is obtained. The resulting polygon layer is then joined with the clustered stop layer to find the service level corresponding to the polygon.

Finally, all the areas are individually analysed to find the density of stops within every region and the total area of each isochrone map is calculated and exported to an Excel spreadsheet. Then the results are presented in the form of tables and maps in the Result sections.

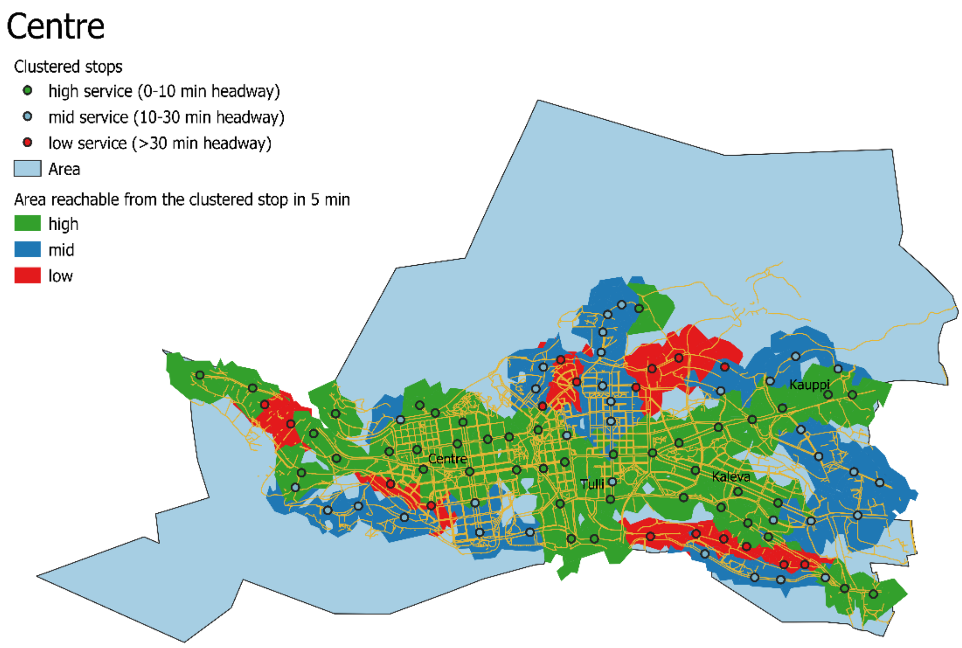
4. RESULTS

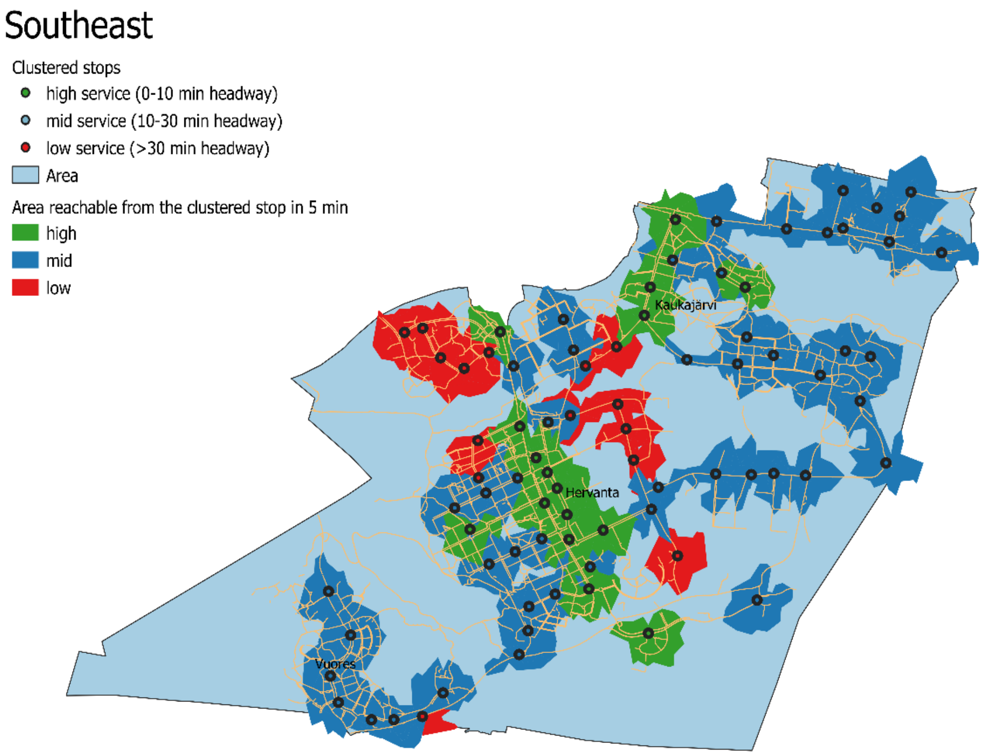
Based on the analysis, I have obtained valuable data that provides insight into the condition of the public transport network's accessibility in the five regions of Tampere City. In Table 4-A, all the different regions with their total land area, the total number of bus stops, and the density of stops are presented. The centre area has the highest density of stops, while the Northeast area has the lowest. In terms of land area, the Northeast is the biggest while the South region is the smallest.

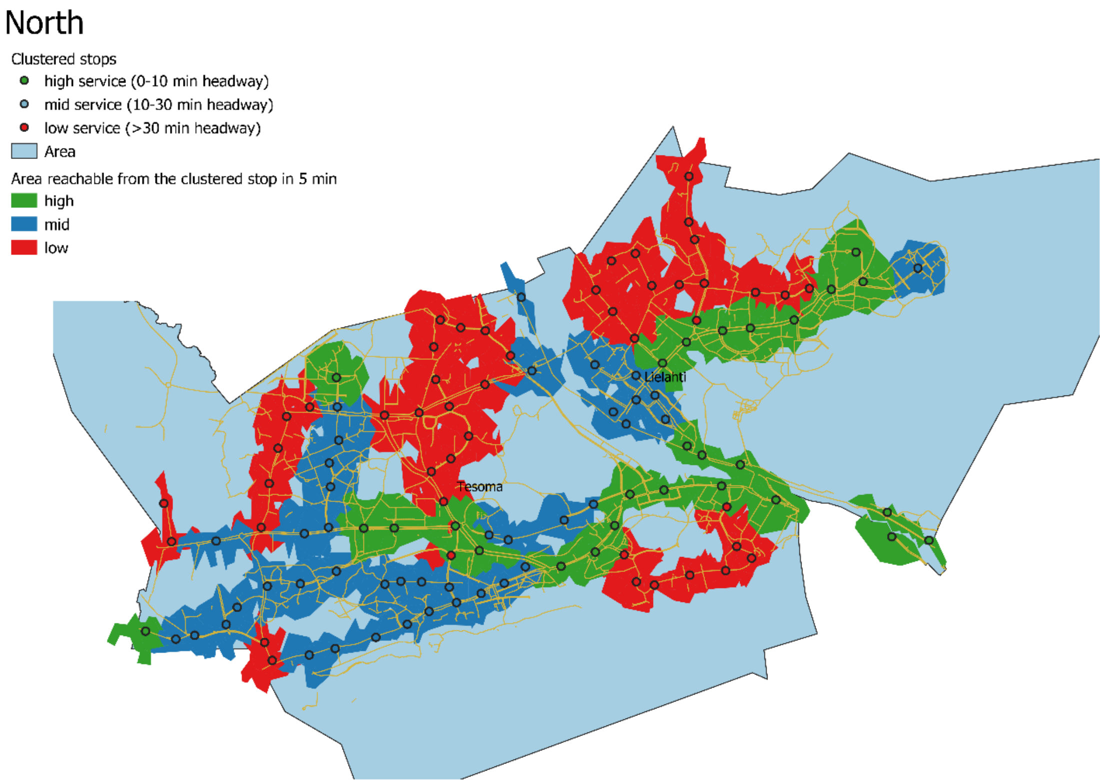
Region	**Total Land Area (km^2)**	**Total Bus and Tram Stop**	**Stops/km^2**
Southeast (Kaakoinen)	32.04	190	5.93
South (Eteläinen)	25.71	194	7.55
Centre (Keskusta)	29.53	272	9.21
North (Läntinen)	43.64	257	5.89
Northeast (Koillinen)	558.67	385	0.69
Total	689.59	1298	

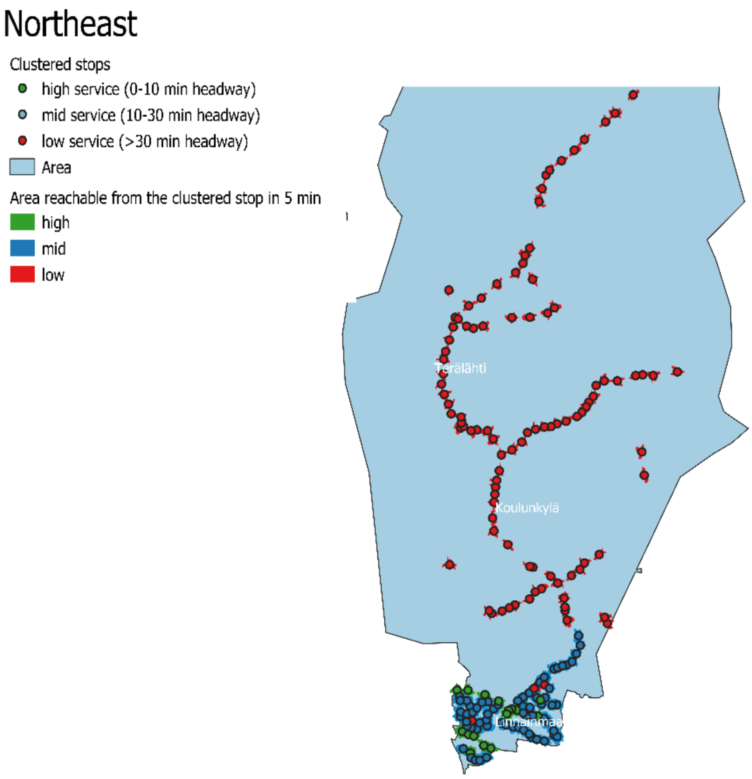
Table 4-A Different regions in Tampere and their stops.

From the GIS analysis, I have produced the following map representation of all the regions with clusters of stops and their corresponding service area which is the area within reach by a 5 minute of walking. These maps are seen in Figure 4-I below:


Centre
Clustered stops
high service (0-10 min headway)
mid service (10-30 min headway)
low service (>30 min headway)
Area
Area reachable from the clustered stop in 5 min
high
mid
low
Centre
Tulli
Kaleva
Kauppi


Southeast
Clustered stops
high service (0-10 min headway)
mid service (10-30 min headway)
low service (>30 min headway)
Area
Area reachable from the clustered stop in 5 min
high
mid
low
Kaukajärvi
Hervanta
Vuores


North
Clustered stops
high service (0-10 min headway)
mid service (10-30 min headway)
low service (>30 min headway)
Area
Area reachable from the clustered stop in 5 min
high
mid
low
Tesoma


Northeast
Clustered stops
high service (0-10 min headway)
mid service (10-30 min headway)
low service (>30 min headway)
Area
Area reachable from the clustered stop in 5 min
high
mid
low
Koulunkylä

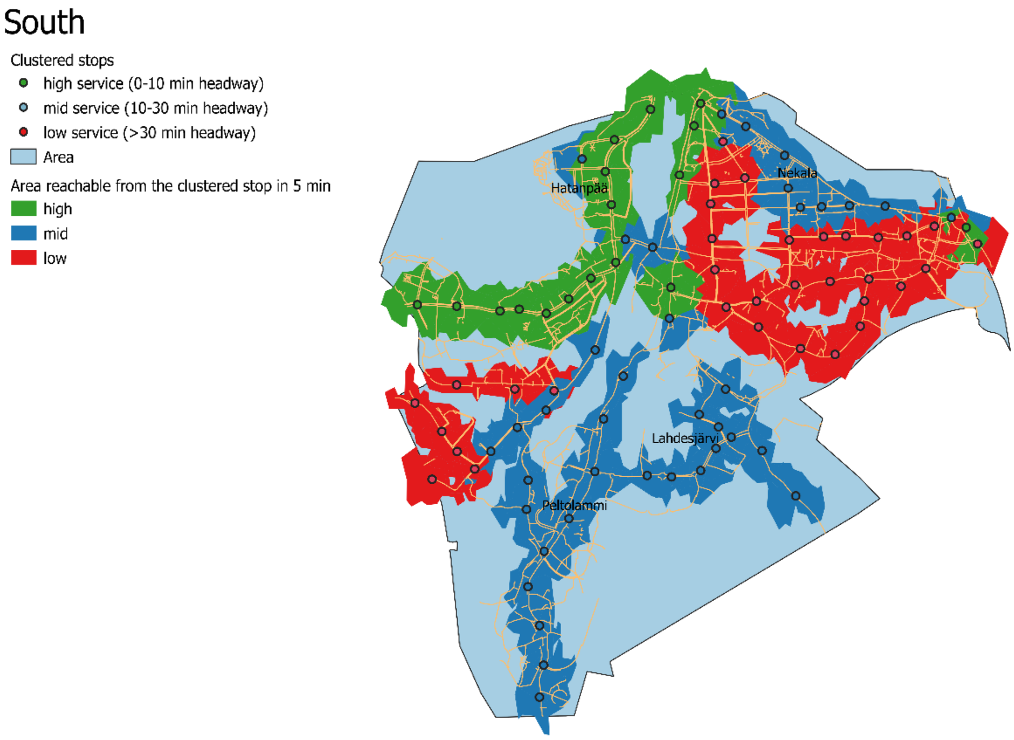


Figure 4-I Maps of all five regions (centre, southeast, south, north, and north-east) with a cluster of stops and their corresponding service areas.

The data from the map is further shown in Table 4-B and Table 4-C with the number of stops that fall within the three different categories and the percentage of the service area that correspond to those stops based on their category across all five regions of Tampere.

Category of stop	**North**	**Northeast**	**South**	**Southeast**	**Centre**
High	49 (19%)	33 (9%)	28 (14%)	39 (21%)	117 (43%)
Mid	94 (37%)	141 (37%)	88 (45%)	35 (18%)	103 (38%)
Low	114 (44%)	211 (55%)	78 (40%)	116 (61%)	52 (19%)
Total	257	385	194	190	272

Table 4-B Number of different categories of stops in five regions.

Category of the corresponding service area	North	Northeast	South	South-east	Centre
High	4.89 km^2 (29%)	3.16 km^2 (11%)	3.39 km^2 (24%)	3.10 km^2 (22%)	7.04 km^2 (54%)
Mid	5.65 km^2 (33%)	10.92 km^2 (39%)	5.88 km^2 (42%)	9.00 km^2 (63%)	4.45 km^2 (34%)
Low	6.48 km^2 (38%)	13.69 km^2 (49%)	4.84 km^2 (34%)	2.09 km^2 (15%)	1.56 km^2 (12%)

Table 4-C Total area and percentage of service area corresponding to different categories of stops in five regions.

In combination with the data from the aforementioned table, the maps provide a robust visual and statistical view of the state of public transport accessibility from a temporal and spatial perspective in different parts of the five regions in Tampere.

5. DISCUSSION

In this section, a further look into the outcome of the study and its limitations is presented. First the implications of the findings of the study are discussed in sub-chapter 5.1 and then the constraints and limitation of the study is discussed in sub-chapter 5.2.

5.1 Implication for the Commuters and Authority

The output of the study shows that different regions of Tampere have vastly different levels of accessibility due to having varying levels of frequency service stops within a quick walk reach. In general, apart from the centre of the city, most regions have somewhat worse access to public transport compared to centre of the city. This shows that the transport planning authority focuses a lot on providing high and easy access to public transport in the centre of the city. While the centre is possibly the most important part of the city due to its economic and functional significance, poor access to public transport in other parts of the city can have negative implications for the overall public transport network of the city. It might be easy to reach other parts of the city from the centre within a reasonable distance to the stop and waiting time, but if the trip from other parts of the city to the centre or other areas is not relatively up to the level it may discourage commuter from relying on public transport as their primary mode of transport.

From the spatial aspect of public transport accessibility, a lot of regions in Tampere have very few stops within reach that have a high frequency of service. A lack of easily reachable bus or tram stops makes it more difficult for commuters to access public transportation. This can be especially problematic for individuals with mobility issues, the elderly, or those carrying heavy loads. It may force them to walk long distances or find alternative modes of transportation, leading to inconvenience and decreased accessibility (Lunke, 2020). In addition to that, due to longer waiting times and lack of easily reachable stops, commuters can experience increased unpredictability in commuting with public transport. Faced with longer waiting times and limited accessibility, commuters may experience higher levels of stress, frustration, and dissatisfaction with the public transportation system. This can have negative effects on their overall well-being and quality of life (Rezapour & Richard Ferraro, 2021).

Tampere's Regional Transport Authority can consider these results for evaluating the accessibility and connectivity of their overall public transport network in Tampere. Based on the data from this study the authority allocates their limited resources to cover more areas and serve more commuters. In addition, the result can also be used to create more stops in underserved areas where there may be more demand.

5.2 Limitations of the Study

Although the study has been conducted to minimize any flaws and limitations, it is not completely free from constraints and weaknesses. One of the biggest constraints of the study is that the GTFS data has been studied only for a single day. Transport schedules can change over time, so basing the entire conclusion on a single date may have left a more accurate picture of the overall state of public transport in Tampere. Especially considering dates at different times of the year would have provided a more insightful and complete view of transport level around the year.

The study has also overlooked the population of the regions. The population of each of the regions would have provided a strong data point to justify the discrepancy in public transport accessibility across different regions and more strongly recommend areas where public transport can be improved. In addition, population density would have provided a better explanation of why the number of stops and their location has varied in different parts of these regions. By addressing these limitations of this research, future studies can make valuable contributions to enhancing the comprehensive understanding of the subject matter, while also facilitating the development of more accurate and reliable conclusions.

6. CONCLUSIONS

This thesis explored the public transport accessibility in different areas of Tampere and shed light on the disparities in accessibility across the city. This study redefined accessibility within the public transport discourse by embracing the contemporary perspective, which emphasizes access to public transport itself rather than solely considering it as a means to reach other urban areas. This new perspective offered a broad understanding of the challenges related to accessibility faced by residents in different areas of Tampere.

The analysis, based on GTFS data, revealed that the centre area of Tampere boasts the best accessibility to public transportation, characterized by a high frequency of stops located within a short distance. However, the findings also highlighted the relatively poorer public transit access in the rest of the city's areas. The disparities identified in this research can have significant implications for residents' mobility, quality of life, and opportunities for social and economic engagement.

Being one of the important and growing cities in Finland, addressing the disparities in public transport accessibility across Tampere is crucial for achieving a more equitable and sustainable city. Policy interventions, such as the expansion of public transportation networks, improvement of service frequencies, and optimization of stop locations in underserved areas, can play a pivotal role in enhancing accessibility for all residents. By acknowledging the variations in public transport accessibility and taking proactive measures to bridge the gaps, Tampere can strive towards a more a seamless and efficient public transportation experience for all.

Future research could delve deeper into understanding the underlying factors contributing to the observed differences in public transport accessibility. Additionally, exploring the impact of improved accessibility on various socio-economic aspects, such as employment opportunities and quality of life, could further inform decision-making processes and facilitate the development of sustainable and inclusive urban transportation network in Tampere.

REFERENCES

Ansari Esfeh, M., Wirasinghe, S. C., Saidi, S., & Kattan, L. (2021). Waiting time and headway modelling for urban transit systems – a critical review and proposed approach. *Transport Reviews*, *41*(2), 141–163. https://doi.org/10.1080/01441647.2020.1806942

Azar, K. T., Ferreira Jr, J., & Wiggins, L. (1994). Using GIS tools to improve transit ridership on routes serving large employment centers: The Boston South End Medical Area case study. *Computers, Environment and Urban Systems*, *18*(3), 205–231.

Bok, J., & Kwon, Y. (2016). Comparable Measures of Accessibility to Public Transport Using the General Transit Feed Specification. *Sustainability*, *8*(3), Article 3. https://doi.org/10.3390/su8030224

Britannica. (2023, May 14). *Tampere | Finland | Britannica*. https://www.britannica.com/place/Tampere

Church, R. L., & Marston, J. R. (2003). Measuring accessibility for people with a disability. *Geographical Analysis*, *35*(1), 83–96.

City of Tampere. (2012, December 14). *Statistical metropolitan areas in Tampere*. https://data.tampere.fi/data/en_GB/dataset/tampereen-suuralueet

City of Tampere. (2023, February 3). *The population of Tampere grew record-breakingly strong*. https://www.tampere.fi/en/current/2023/02/03/population-tampere-grew-record-breakingly-strong-growth-expected

Daniels, R., & Mulley, C. (2013). Explaining walking distance to public transport: The dominance of public transport supply. *Journal of Transport and Land Use*, *6*(2), 5–20.

Google Developers. (2023). *GTFS Static Overview*. https://developers.google.com/transit/gtfs

Gould, P. R. (1969). *Spatial Diffusion, Resource Paper No. 4.*

GTFS community. (2023). *Best Practices—General Transit Feed Specification*. https://gtfs.org/schedule/best-practices/

Jordahl, K., Bossche, J. V. den, Fleischmann, M., Wasserman, J., McBride, J., Gerard, J., Tratner, J., Perry, M., Badaracco, A. G., Farmer, C., Hjelle, G. A., Snow, A. D., Cochran, M., Gillies, S., Culbertson, L., Bartos, M., Eubank, N., maxalbert, Bilogur, A., … Leblanc, F. (2020). *geopandas/geopandas: V0.8.1* (v0.8.1). Zenodo. https://doi.org/10.5281/zenodo.3946761

Kawabata, M. (2003). Job access and employment among low-skilled autoless workers in US metropolitan areas. *Environment and Planning A*, *35*(9), 1651–1668.

Lunke, E. B. (2020). Commuters' satisfaction with public transport. *Journal of Transport & Health*, *16*, 100842. https://doi.org/10.1016/j.jth.2020.100842

Makr, M.-C., Folkesson, C., & others. (1999). Accessibility measures for analyses of land-use and travelling with geographical information systems. *Proceedings from the Annual Transport Conference at Aalborg University*, *6*(1).

McKinney, W. (2010). Data Structures for Statistical Computing in Python. In S. van der Walt & J. Millman (Eds.), *Proceedings of the 9th Python in Science Conference* (pp. 56–61). https://doi.org/10.25080/Majora-92bf1922-00a

Nolde, N. (2020, December 14). Valhalla QGIS Plugin. *GIS•OPS*. https://gis-ops.com/valhalla-qgis-plugin/

Nysse. (2023). *Nysse, Tampere Regional Transport Authority*. Nysse, Tampere Regional Transport. https://www.nysse.fi/en/front-page.html

Pandas development team. (2020). *pandas-dev/pandas: Pandas* (latest). Zenodo. https://doi.org/10.5281/zenodo.3509134

Polzin, S. E., Pendyala, R. M., & Navari, S. (2002). Development of time-of-day–based transit accessibility analysis tool. *Transportation Research Record*, *1799*(1), 35–41.

Prommaharaj, P., Phithakkitnukoon, S., Demissie, M. G., Kattan, L., & Ratti, C. (2020). Visualizing public transit system operation with GTFS data: A case study of Calgary, Canada. *Heliyon*, *6*(4), e03729. https://doi.org/10.1016/j.heliyon.2020.e03729

QGIS Documentation. (2020, June 5). *Vector analysis*. https://docs.qgis.org/3.28/en/docs/user_manual/processing_algs/qgis/vectoranalysis.html#dbscan-clustering

Rezapour, M., & Richard Ferraro, F. (2021). The impact of commuters' psychological feelings due to delay on perceived quality of a rail transport. *Humanities and Social Sciences Communications*, *8*(1), Article 1. https://doi.org/10.1057/s41599-021-00865-z

Tampereen Ratikka. (2023). *Tampereen Ratikka*. Tampereen Ratikka. https://www.tampereenratikka.fi/en/

Tampereen seudun joukkoliikenne. (2023). *Tampere Public Transport timetables and routes (GTFS)*. https://data.tampere.fi/data/en_GB/dataset/tampereen-joukkoliikenteen-aikataulut-ja-reitit-gtfs

Wachs, M., & Kumagai, T. G. (1973). Physical accessibility as a social indicator. *Socio-Economic Planning Sciences*, *7*(5), 437–456. https://doi.org/10.1016/0038-0121(73)90041-4

APPENDIX A

main

May 23, 2023

0.0.1 Programmatic Analysis of GTFS Data

Importing GTFS files as pandas data-frames

[1]:
```
import pandas as pd
import geopandas as gpd

stops = pd.read_csv('feed\stops.txt')
stop_times = pd.read_csv('feed\stop_times.txt')
trips = pd.read_csv('feed/trips.txt')
calendar_dates = pd.read_csv('feed/calendar_dates.txt')
```

[2]:
```
date_str = '20230517' # date of interest
```

`service_id` for that date from `calender-dates.txt` and the corresponding trips from `trips.txt` are extracted.

[13]:
```
calendar_dates = calendar_dates[calendar_dates['date'] == int(date_str)]
service_ids = calendar_dates['service_id'].unique()

trips_for_date = trips[trips['service_id'].isin(service_ids)]
# TOTAL TRIPS FOR DATE
total_trips = trips_for_date['trip_id'].unique()
print('Total trips for date: ', len(total_trips))
trip_ids = trips_for_date['trip_id'].unique()
```

```
Total trips for date:  5250
```

Filter stop_times to only include trips on the specified date

[4]:
```
stop_times = stop_times[stop_times['trip_id'].isin(trip_ids)]
```

Aggregate by stop and then by hour to get the average number of departures per hour for each stop

[5]:
```
freq_by_stop = stop_times.groupby('stop_id').size().
  ↪reset_index(name='total_freq')
freq_by_stop['per_hour'] = freq_by_stop['total_freq'] / 24

freq_by_stop.head()
```

[5]:
```
   stop_id  total_freq   per_hour
0        1         125   5.208333
1        2         150   6.250000
2        5         105   4.375000
3        6         262  10.916667
4        7         107   4.458333
```

[6]:
```
freq_by_stop['headway'] = (60 / freq_by_stop['per_hour']).round()

freq_by_stop.head()
```

[6]:
```
   stop_id  total_freq   per_hour  headway
0        1         125   5.208333     12.0
1        2         150   6.250000     10.0
2        5         105   4.375000     14.0
3        6         262  10.916667      5.0
4        7         107   4.458333     13.0
```

Classify stops into categories based on the headway and name the column level: `low`, `medium`, `high`

[7]:
```
freq_by_stop['level'] = pd.cut(freq_by_stop['headway'], bins=[0, 10, 30,␣
↪float('inf')], labels=['high', 'mid', 'low']).astype(str)

freq_by_stop.head()
```

[7]:
```
   stop_id  total_freq   per_hour  headway level
0        1         125   5.208333     12.0   mid
1        2         150   6.250000     10.0  high
2        5         105   4.375000     14.0   mid
3        6         262  10.916667      5.0  high
4        7         107   4.458333     13.0   mid
```

Merge again with `stops` to get the stop information for each stop

[8]:
```
freq_by_stop = pd.merge(freq_by_stop, stops, on='stop_id')

freq_by_stop.head()
```

[8]:
```
   stop_id  total_freq   per_hour  headway level  stop_code     stop_name  \
0        1         125   5.208333     12.0   mid          1  Keskustori H
1        2         150   6.250000     10.0  high          2  Keskustori G
2        5         105   4.375000     14.0   mid          5  Keskustori K
3        6         262  10.916667      5.0  high          6     Paloasema
4        7         107   4.458333     13.0   mid          7  Keskustori J

    stop_lat   stop_lon zone_id  wheelchair_boarding  municipality_id
0  61.497538  23.761522       A                  NaN              837
1  61.497588  23.761488       A                  NaN              837
```

2	61.497340	23.761540	A	NaN	837
3	61.500918	23.767665	A	NaN	837
4	61.497397	23.761547	A	NaN	837

Convert to a GeoDataFrame and save as a ESRI Shapefile

[9]:
```
freq_by_stop = gpd.GeoDataFrame(freq_by_stop, geometry=gpd.
 ↪points_from_xy(freq_by_stop.stop_lon, freq_by_stop.stop_lat))

# finland crs
freq_by_stop.set_crs(epsg=2392, inplace=True)

freq_by_stop.to_file('freq_by_stop_hour.shp', driver='ESRI Shapefile')
```

```
C:\Users\shams\AppData\Local\Temp\ipykernel_19396\1614197127.py:6: UserWarning:
Column names longer than 10 characters will be truncated when saved to ESRI
Shapefile.
  freq_by_stop.to_file('freq_by_stop_hour.shp', driver='ESRI Shapefile')
```