UNIVERSITY OF TARTU
Faculty of Science and Technology
Institute of Computer Science
Software Engineering Curriculum

Helen Tera

Analytics and Decision Support for Public Bicycle Sharing in Tartu

Master's Thesis (30 ECTS)

Supervisors: Amnir Hadachi, PhD
Mozhgan Pourmoradnasseri, PhD

Tartu 2022

Analytics and Decision Support for Public Bicycle Sharing in Tartu

Abstract:

A public bicycle sharing system is a short-term rental service in which bicycles are made available for shared use. Due to the environmental, social, and health benefits of cycling, significant motivation exists for promoting cycling in a city, and public bicycle sharing is a promising approach for increasing the number of cyclists. In 2019 a public bicycle sharing program was launched in Tartu. Since then, it has experienced intense usage, but no thorough analysis has been conducted based on its novel data. This thesis aims to use this little-explored data in order to understand the mobility patterns of Tartu Smart Bike users and provide justified suggestions for improving both the existing bicycle sharing network and the cycling infrastructure of Tartu in general. As a result, a detailed analysis of the usage patterns of Tartu Smart Bike is carried out, giving a better understanding of the cyclists' behavior and motivation in the city of Tartu. A reproducible method for locating new bicycle share stations is presented, and its results based on the case of Tartu Smart Bike are discussed and assessed. Additionally, the road usage intensity of Tartu Smart Bike users is analyzed in order to identify areas that demonstrate a lack of designated cycling infrastructure. This thesis presents both findings specific to Tartu Smart Bike and methods that can easily be adjusted for use in other cities in the hope of contributing to the development of a more bicycle-friendly society.

Keywords: *public bicycle sharing, network analysis, spatiotemporal analysis, Tartu Smart Bike*

CERCS: P170 (Computer science, numerical analysis, systems, control)

Andmeanalüütika ja otsustustugi Tartu rattaringlusele

Lühikokkuvõte:

Rattaringlus on transporditeenus, mis pakub selle kasutajatele ühisrataste lühiajalise rentimise võimalust. Jalgrattaga liiklemisel on märkimisväärsed positiivsed mõjud nii keskkonnale kui ka jalgratturi tervisele ja heaolule ning rattaringlus on paljulubav strateegia jalgratturite arvu suurendamiseks linnapildis. 2019. aastal avati Tartus 750 jalgrattaga laenutusvõrk. Sellest ajast on Tartu rattaringlus leidnud palju kasutust ja vastukaja, kuid ühtegi süvaanalüüsi Tartu uudseid rattaringlusandmeid kasutades pole läbi viidud. Selle lõputöö eesmärgiks on kasutada neid vähe-uuritud andmeid, et aru saada Tartu jalgratturite liikuvusmustritest ning teha põhjendatud ettepanekuid Tartu rattaringluse ja jalgrattavõrgustiku parandamiseks. Lõputöö tulemusena valmis põhjalik analüüs Tartu rattaringluse kasutusmustritest, andes parema arusaama Tartu jalgratturite käitumisest ja eesmärkidest liikluses. Samuti on esitatud taaskasutatav mudel uute rattaringlusjaamade asukohtade hindamiseks ning selle mudeli tulemused Tartu rattaringluse näitel. Lisaks on analüüsitud Tartu rattaringluse kasutajate sõidutrajektoore, et leida puudujääke Tartu jalgrattateede võrgustikus. Antud lõputöö kirjeldab nii konkreetseid tulemusi Tartu rattaringluse näitel kui ka terviklikke taaskasutatavaid mudelid, mis on kohandatavad kasutuseks teistes linnades, lootuses panustada rattasõbralikuma ühiskonna arengusse.

Võtmesõnad: *rattaringlus, liiklusvõrgustike analüüs, aegruumiline analüüs, Tartu rattaringlus*

CERCS: P170 (Arvutiteadus, arvanalüüs, süsteemid, kontroll)

Acknowledgements

First, I would like to express my gratitude to the Tartu City Government for providing the data and the possibility to complete this research. I would also like to thank my supervisors, Dr. Amnir Hadachi and Dr. Mozhgan Pourmoradnasseri, for constantly pushing me to improve while being utterly understanding and supportive at the same time. Additionally, I would like to thank Artjom Lind for providing the necessary technical help and always being positive and enthusiastic about the challenges faced.

I hold special gratitude in my heart to my family for their unwavering belief in me and to Ralf, who can always put a smile on my face.

Contents

1 Introduction

As sustainability and public health are major objectives of modern city planning, public bicycle sharing (PBS) programs play an increasingly important role in urban transportation system design. Public bicycle sharing is a transport service in which bicycles are provided for individual short-term use. Reducing the number of personal motorized vehicles is one of the more common strategies for reducing carbon emissions and traffic congestion, and introducing PBS systems can play an important role in increasing bicycle modality share in cities. Some programs have reported an increase in modal share of bicycles up to 1.5% owing to the introduction of bicycle sharing [1]. Besides demobilization and decarbonization, other common benefits of PBS systems to its users include social interaction, accessibility, and low cost [2].

Along with the advancements and availability of GPS technologies, emerging PBS programs produce large amounts of data, which give insight into PBS usage and human mobility patterns. Analyzing this novel data can provide guidance for better bicycle scheduling, station planning, and road network design [3]. Using this data to advance existing PBS systems and cycling infrastructure can in turn convince more people to choose cycling as their primary means of getting around, benefiting both the individuals and the community.

In 2019, a public bicycle sharing program was established in Tartu. Since then, its benefits and problems have been widely debated in media and journalism, discussing pricing [4], safety [5], and accessibility [6]. Despite various problems, Tartu Smart Bike developed into a heavily used system - its 750 bicycles covered a distance of almost 2.5 million kilometers in total during their first year of use [7]. While bicycle sharing in Tartu has gotten a lot of attention since its launch, no publicly available comprehensive analysis of its usage has been conducted. This thesis aims to look into the mobility patterns of Tartu Smart Bike users and suggest improvements to the already popular PBS system and local cycling infrastructure, hoping to further promote sustainable mobility and healthy lifestyles.

1.1 Research goals

This thesis aims to make use of the relatively unexplored data from Tartu Smart Bike in order to understand the dynamics of Tartu Smart Bike and suggest improvements for both the PBS system and overall cycling infrastructure in Tartu. There are many tasks in public bicycle sharing, including bicycle reallocation, demand prediction, and station network design, that serve to improve bicycle sharing service. To limit its scope, this thesis focuses on the following tasks:

- **RG1.** Understanding the usage patterns and demand of Tartu Smart Bike from temporal and spatial aspects.

- **RG2.** Identifying optimal locations for new PBS stations in Tartu.
- **RG3.** Analyzing the bicycle road network in order to pinpoint deficiencies in the existing cycling infrastructure in Tartu.

While public bicycle sharing, in general, is becoming an increasingly more studied topic, the PBS system in Tartu has remained unexplored. This thesis hopes to provide decision support for city officials and policymakers in Tartu to develop bikeability of the city and improve the quality of life for its citizens.

1.2 Road map

The rest of this thesis is organized as follows:

- **Chapter 2 - Background.** This chapter presents the reader with the necessary context and background to understand the topics and terminology of the thesis. In addition, a short overview of prior research on relevant methods is presented.
- **Chapter 3 - Data preparation.** In this chapter, the used datasets are introduced. Data is cleaned and prepared for use in the following analysis. A separate subsection covers a method for geolocation data cleaning.
- **Chapter 4 - Analysis.** This chapter comprises the core of the thesis. In this chapter, various analytical methods and models are employed to reach the research goals. The first subsection covers exploratory analysis; the second subsection presents a model for PBS station location, and the third subsection deals with road network analysis.
- **Chapter 5 - Discussion.** In this chapter, the findings of the analysis are discussed. Ideas on improvements and future work are included.

2 Background

This chapter provides background information on the topics relevant to this thesis. Section 2.1 presents an overview of PBS systems and their history, while section 2.2 introduces the public bike sharing system of Tartu - Tartu Smart Bike - in more detail. Relevant information about the demographics and geography of Tartu is available in section 2.3. Lastly, various related work in connection with all three research goals is explored in sections 2.4, 2.5 and 2.6.

2.1 Public bicycle sharing

A public bicycle sharing (PBS) system is a short-term rental service in which bicycles are made available for shared use. The history of PBS systems dates back to 1965, in Amsterdam, where the first PBS system named the White Bikes was introduced [8]. Ordinary bicycles, which were painted white for recognition, were left around the city for anyone to use. Due to customer anonymity, it was easy to steal or damage the bicycles, and thus, the program collapsed within days. While the program ended in failure, it already included some of the components of PBS systems we know today, such as distinguishing public bikes by design or color [8].

According to DeMaio [1], the first large-scale 2nd generation PBS system was launched in Copenhagen in 1991. The program introduced multiple improvements to 1st generation PBS systems, such as reinforced bicycles, fixed pick-up and drop-off locations, and a coin deposit system. While more organized than the previous programs, it similarly struggled with theft problems due to customer anonymity.

Five years later, in 1996, the advancements in information technologies allowed the forming of 3rd generation PBS systems. The 3rd generation of bicycle share programs is characterized by various technological improvements, such as electronically-locking racks, bicycle tracking, and user interfaces for locating and renting bicycles. New technologies provided a solution to the problem of user anonymity and impelled the popularity of PBS systems [8]. As of mid-2021, there were at most 1999 simultaneously operating PBS systems around the world. A major portion of world PBS programs was employed in Asia (44%) and Europe (41%) [9].

There are two types of public bicycle sharing systems: traditional station-based and the more recent free-floating systems. In station-based systems, bicycles have to be picked up and dropped off at designated stations, while free-floating systems provide an option of picking up and leaving the used bicycle anywhere within the region of operation [10]. While free-floating PBS systems provide more comfort and increase usability, they also pose multiple difficulties. Efficient re-balancing strategies need to be considered for free-floating systems to avoid congregation of shared bicycles in high-interest regions or scattering of the bicycles in remote areas. Station-based systems also provide an option for charging electric bicycles and thus minimize maintenance activities, and costs [11].

2.2 Tartu Smart Bike

The public bicycle sharing scheme in Tartu was launched on the 8th of June 2019 with 750 bikes in 69 different stations. By mid-2021, the number of stations had increased to 85, including multiple stations outside the bounds of the city [12]. In this thesis, the data from 90 stations that were or became operational during the year 2021 is used with the emphasis on the stations located within the city limits. Tartu Smart Bike stations are displayed on a map using red markers in Figure 1. Additionally, the system provides support for *virtual stations*, which are short-term docking locations without physically hard-to-install docks resembling free-floating PBS systems, that can be used during large-scale events [13].

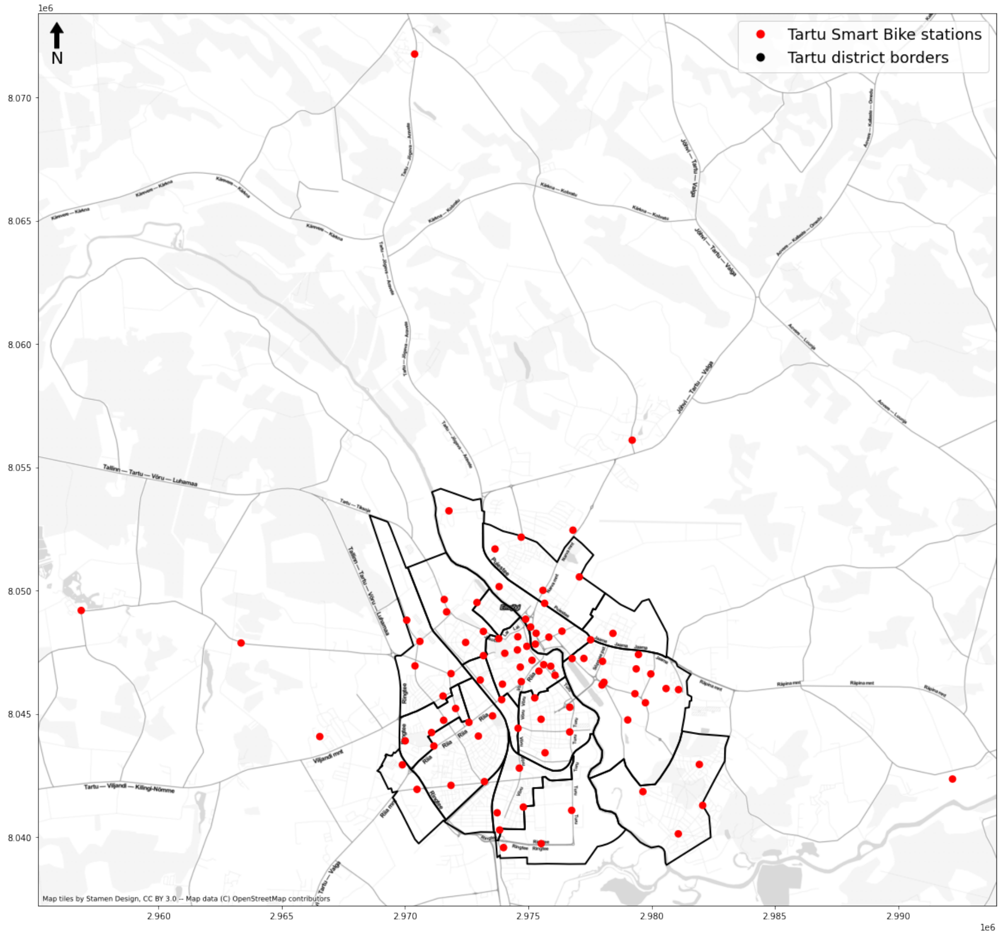


Figure 1. Tartu city districts and Tartu Smart Bike stations

There are two types of bicycles in Tartu Smart Bike system: electric-assist bikes and conventional 8-speed bikes. There were 500 electric-assist bikes and 250 regular bikes in circulation in 2021. The electric-assist bikes operate just like classic bikes but provide background power when accelerating or riding up an incline. In addition, all bikes are equipped with an electric horn, automatic front and back lights, kickstand, front and back fenders, seat height adjustment, and a front basket [12].

All bicycles are also equipped with a GPS module and can communicate with the back-end using 3G network connection. This allows for the bicycles to broadcast their location on a regular basis and save the origin and destination of each trip, as well as other helpful trip metrics such as duration and distance of the trip [13].

2.3 Study area

With a population of 91 thousand people, Tartu is the second largest city in Estonia. It is historically known as the intellectual capital of Estonia, as the oldest university in Estonia - University of Tartu - is located there. Students of higher educational institutions make up more than a fifth of the population, and almost half of the city's budget is spent on education. This makes Tartu a very youthful and innovation-prone city [14].

Tartu is situated along the banks of Emajõgi, which splits the city into two parts. Seven bridges connect these two parts, including two bridges that are designed solely for pedestrians and cyclists. While Tartu is officially divided into 17 districts, they carry no administrative purposes. The district borders are visualized in Figure 1. Some of the districts - such as Annelinna, Tammelinna, Veeriku - are considered primarily residential, while others, including Ränilinna and Ropka tööstuse, are more industrial. The most populated district of Tartu is Annelinna, which makes up for 27% of the city's population [15].

2.4 Features of public bicycle sharing

PBS data has many different features that can provide value for city administrations and urban scientists. The aim of this section is to collect different features used for the analysis and visualization of PBS data in order to give a better understanding of PBS usage and mobility patterns.

Station-level demand describes the number of bicycles rented from a given station in a specific time frame, either per hour, day, month, or even year. Being able to ensure that a user can rent a bicycle from a desired station and return it at a station matching their destination, is of crucial importance to provide the expected flexibility and service quality of PBS [16]. Understanding station-level demand plays a major part in solving this problem. Both Ram *et al.* [17] and Ying *et al.* [18] design a visualization tool to assist city administrators with the PBS system's decision making, and in both designs, station-level demand maps play a central role. Various other papers that aim to understand city dynamics based on PBS data make use of station-level demand analysis [19, 20].

Another common strategy for analyzing PBS data is concentrating on its usage from temporal aspects. By grouping the number of trips either by hour, day, or month, different patterns can be discovered, giving insight into the motivations and problems of PBS users. One of the most predominant ways of describing PBS data is through the distribution of trips by the time of day [18, 21, 22, 23]. Understanding public bicycle demand changes in the span of one day can provide decision support for maintenance planning or bandwidth management, for example, to avoid the evident disadvantages of scheduling station maintenance at a time when public bicycle demand is at its peak. Some studies also analyze PBS trips throughout a longer period of time, for example, using monthly distribution [24], or present the changes in trip counts as a line chart over a longer span of time - often a year - aggregating trips either daily or weekly [17, 3, 25].

Temporal analysis of PBS data often also includes the comparison of trip rates during weekdays and weekends [17, 20, 25]. Significant differences in the number of trips on weekdays and weekends can point to different motivations for PBS use. If the total number of PBS trips is much higher during weekdays, it could indicate that public bicycles are primarily used for utilitarian purposes - commuting to work or school - while a high number of PBS trips during weekdays could suggest that a given PBS program is also often used for recreational purposes. A similar idea is used in a study by Guo *et al.* [22] for classifying clusters based on their functional profiles.

Studying the PBS usage behaviors by different user groups can provide policymakers insight into the commuting behavior of various demographic groups. This insight can help to encourage underrepresented groups to partake in active modes of transport. One common technique that can provide a better understanding of PBS users' mobility patterns is grouping PBS trips by users' gender or age. In a paper by Zhang *et al.* [26], PBS users' gender is one of the metrics used to predict the potential destination station and arrival time of PBS user at the start of their trip, while Wood *et al.* [27] analyze the PBS trips by gender and discuss the reasons for differing travel patterns among users of different gender. It is revealed that women are underrepresented among the PBS users in London, and different travel patterns are discovered for men and women: male cyclists use the system predominantly for commuting, while the trip patterns of female PBS users demonstrate a dominant leisure function.

2.5 Station location selection in bicycle sharing

Station location selection problem in public bicycle sharing resides in identifying optimal locations for PBS stations. Various models and techniques have been proposed to solve the issue. Some studies concentrate on designing new PBS networks from scratch [28, 29], while others focus on improving existing PBS systems by proposing locations for new stations [30, 31, 32]. This thesis aims to improve an existing PBS system in Tartu, and thus the emphasis is on the latter.

To evaluate possible candidate locations for new PBS stations, different criteria can be considered. The difficulty of the station location selection problem is in choosing a good set of criteria or appropriate weights for these criteria. Models that aim to find the best option from a set of candidate alternatives with respect to multiple criteria are called multi-criteria decision-making (MCDM) [33]. MCDM has been efficiently used for PBS station location selection and is often combined with GIS (geographic information system) tools, as the surrounding infrastructure and built environment can be used to create meaningful criteria [31, 34, 35]. For example, Banerjee *et al.* [30] use a combination of GIS techniques and a modified Huff's model in order to find optimal station locations for PBS in Baltimore. GIS tools are used to analyze PBS trip data collected over a period of four months in order to assess road usage intensity among PBS users. This information is used to calculate the correlation between road usage intensity and various criteria, such as proximity to pubs and restaurants, transit stops, and attraction sites. Pearson's correlation test is used to select criteria with the strongest negative correlation, and a suitability score based on the selected criteria is calculated for each candidate location using Huff's gravity model. Kabak *et al.* [31] use MULTIMOORA model in order to rank the candidate locations based on 12 conflicting criteria, including proximity to sports centers, shopping malls, public transport stations, and population density. A free, open-source software - SuperDecisions - with analytic hierarchy process (AHP) in the background is used to select appropriate weights for the criteria. In the study by Guler *et al.* [32] similar criteria are used, but the emphasis is mainly on weight selection for the criteria. An expert in the field uses the best worst method (BWM) to assign weights to the criteria, while another MCDM model called TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) is used for the ranking of the candidates. All of these studies follow a similar approach of selecting meaningful criteria, calculating appropriate weights for these criteria, and then ranking them using a model of choice. The weight selection is often conducted with the help of a domain expert or using existing software.

2.6 Bicycle road network analysis

The road network analysis task encapsulates the exploration of a road network as a graph in order to accomplish various goals, such as finding the shortest paths, evaluating road usage intensity, or finding coverage. The earlier works on bicycle road network analysis are largely based on the stated preference (SP) surveys, in which participants are asked about their decisions in hypothetical situations [36]. Advancements in GPS technology and its availability have given rise to studies using revealed preference (RP) data. According to Pritchard [37], the majority of studies on route choice behavior of bicycle users rely on GPS data. For example, Rupi *et al.* [38] use a combination of GPS traces and traditional counting methods in order to identify weaknesses in a bicycle road network. The GPS traces from volunteers' smartphones are map-matched to the road network of Bologna using SUMO (Simulation of Urban Mobility). The bicycle counts acquired from traditional counting methods were used to discover a significant correlation between the GPS dataset and traditional counts. This allowed for the counts to be used for scaling the map-matched cyclists' volumes to estimate the actual road usage intensity.

With the development of modern PBS systems, a novel RP data source of GPS traces from public bicycles has emerged. This data provides useful input for both cycling infrastructure and bicycle share program improvements. Scott *et al.* [36] use data from a public bicycle sharing program in Hamilton, Ontario, to study the route choice behavior of PBS users. PBS trips are map-matched to the local road network in order to find alternative routes between a pair of PBS stations. For each origin-destination pair, the Gini index is calculated to assess the dispersion in the route choice and identify bias towards a specific route choice. As a result, meaningful patterns are found in the behavior of PBS users, such as users' tendency to choose a longer route to their destinations if it avoids steep slopes or high traffic volumes.

As an alternative approach, Wysling and Ross [39] use only road network data to evaluate bicycle suitability along different road segments in Paris. Based on the available infrastructure and speed limit, a suitability rating is assigned for each road segment. Along with the slope, these two multipliers are used to calculate the perceived distance for each street segment. Next, the study area is split into a grid of 250×250 meter cells, each of which is considered as a trip origin location, while destinations are various points of interest all over the city. The bikeability of each cell is defined as the average of perceived shortest paths to all destination locations. Thus, a complete bicycle suitability map for the city is acquired.

3 Data preparation

This thesis is based on various data from Tartu Smart Bike, provided by the city of Tartu. Three major datasets are used for the analysis:

- **List of Tartu Smart Bike stations.** This dataset consists of 85 records, representing all stations in Tartu Smart Bike network, with their features such as *station_id*, *station_name*, *latitude* and *longitude*. Additionally, to support district-level analysis, two features were added: *district_id* and *district_name*.
- **Tartu Smart Bike trip records.** In this dataset, each record depicts a single trip made via Tartu Smart Bike system. The records include following features: *bike_id*, *origin_name*, *destination_name*, *unlock_date_time*, *lock_date_time*, *length*, *id_code* and *year_of_birth*. Another attribute - *duration* - was added by comparing *unlock_date_time* and *lock_date_time*, that depicts the duration of trips in minutes.
- **Tartu Smart Bike bicycles' geolocation data.** This dataset is the most immense in its magnitude. Each record denotes a single location of a PBS bike during its trip. Each location is described using features *route_code*, *bike_id*, *latitude*, *longitude* and *date_time*.

The used data covers the period from the 1st of January 2021 until the 31st of December 2021. Before the data analysis stage, an exploratory approach was used to understand the data, and a simple data cleaning procedure was applied to remove noisy records and improve the quality of the available data.

Additional smaller datasets, such as Tartu district borders and population data, were acquired from the city of Tartu. Geographic data needed for this thesis was downloaded from OpenStreetMap[1] (OSM), including the road network of the study area and locations of schools, universities, and shopping malls. OSM is a free geographic database whose data has been widely used in various scientific research.

3.1 Cleaning trip records

Tartu Smart Bike trip records initially consisted of **838709** trips that took place throughout the year 2021. The extracted *duration* attribute was used to reduce the number of trips by removing trips whose duration was less than one minute, as trips with such a short duration would not provide value for this thesis and were likely the result of a user getting a damaged or a low-battery bike, which was immediately returned to the same station.

[1] https://wiki.openstreetmap.org

After removing these rows, **828390** trip records remained and were used for further analysis. A similar procedure based on attribute *length* was also considered, but comparing the lengths of various trips to their duration and the actual trajectory suggested that the provided *length* attribute was not precise and could not be relied on. Additionally, all the records with unusable or irrelevant origin or destination stations were also removed. This included removing records where either the origin or destination station was missing or undetermined, reducing the number or trips further down to **821112**, and trips, which started or ended in Tartu Smart Bike repair workshop (recorded as station "Töökoda") making the final number of records **820932**.

3.2 Cleaning geolocation data

The next goal in data cleaning was to reduce the volume of the geolocation data. Geolocation data generated by Tartu city bikes had an immense magnitude - even December, which was the least busy month in regards to cycling, had almost a quarter a million data points, while the busiest month of June had over 12 million data points. Based on the available data, a uniform pattern could be detected in how Tartu Smart bikes pinged their location. A given bicycle would send its location every 5 seconds 5 times, after which a longer pause of 40 seconds would follow. This pattern would then repeat for the duration of the trip, although there were also some fluctuations to this pattern. Hence, a single bicycle would normally generate 5 data points in a minute. With an average trip duration of **17.6** minutes, this makes for around **90** points per trip.

3.2.1 Method description

In order to reduce the magnitude of geolocation points, a data cleaning algorithm designed specifically for geographical data was applied. The algorithm is based on the work of Yang *et al.* [40], which relies on the assumption of movement consistency. Since most real-world trips consist of numerous sequences of moving straight along a road alternating with turning at intersections, any trip can be partitioned into a set of sub-trajectories consisting of these straight lines split by intersections. By doing so, a significant amount of points along these lines can be omitted.

This is done by calculating the distance offset and angle offset from the previous movement trajectory, which is illustrated in Figure 2. For every point p_i in trajectory $T = (p_1, p_2, \ldots, p_n)$ where $i > 2$, the distance offset and the angle offset are calculated from the previous movement trajectory. The distance offset is the distance between the point p_i and its projection on the vector (p_{i-1}, p_{i-2}), while the angle offset is the angle between the vectors (p_{i-1}, p_{i-2}) and (p_i, p_{i-1}). If either of these offsets exceeds the corresponding threshold a_1 or a_2, a change of direction or a significant deviation from the movement trajectory is detected, and the previous point is added to the cleaned trajectory T', denoting the end of the previous sub-trajectory.

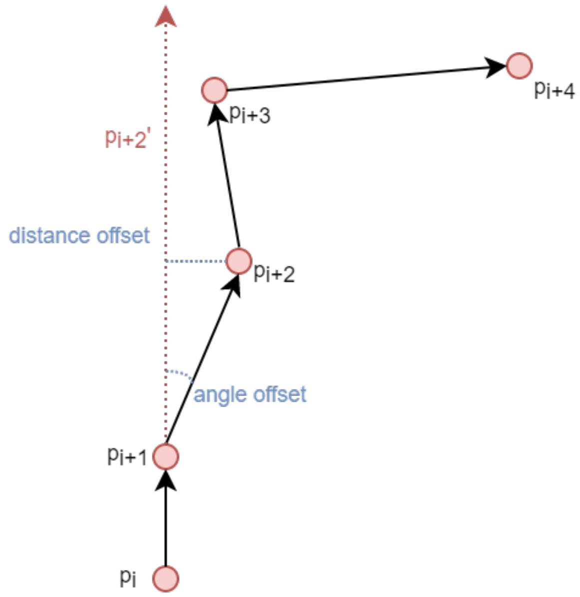


Figure 2. Using distance and angle offset to partition the movement trajectory.

If a point is discarded, the next point is compared to the previous movement trajectory rather than the one containing the discarded point. Thus, all the points in the trajectory are triple-wise compared, and distance and angle offsets are calculated. Based on these values, points are either kept or discarded, ideally leaving a trajectory with a minimal amount of points marking only significant turning points in the route.

3.2.2 Threshold determination

To decide if a point belongs to a movement trajectory, it is compared to a certain threshold. Threshold determination is essential since these thresholds should accurately identify all turning points while not being too sensitive to noise in geolocation data. Thresholds are based on two major factors - knowledge of the infrastructure of the city and the shape complexity of a trajectory. Distance and angle thresholds are determined as follows:

$$a_1 = \lambda_1 + \log_a(\beta_1)$$
$$a_2 = \lambda_2 + \log_a(\beta_2)$$

The first parameters λ_1 and λ_2 are features of infrastructure, where λ_1 signifies the maximum width of the road in a city and λ_2 denotes the smallest angle of intersection in a city. $\beta_1$and β_2 are the standard deviations of the distance offset and angle offset in a trajectory, respectively. In addition the tuning factor $0 < a < 1$ is used. The data cleaning threshold parameters selected for this thesis were based on the road characteristics of Tartu and were as follows:

$$\begin{aligned} \lambda_1 &= 20 \\ \lambda_2 &= 30 \\ a &= 0.25 \end{aligned}$$

3.2.3 Results

As a result of running the data cleaning method, the initial dataset consisting of 836265 trajectories and over 67 million GPS points was cut down to 747653 trips and a little over 14 million points. The trips lost during the cleaning were trips with two or fewer different points in their trajectory, which would not allow using the cleaning method described above. Furthermore, these trips would not provide any value to this analysis, as two or three different points cannot compose a meaningful trip.

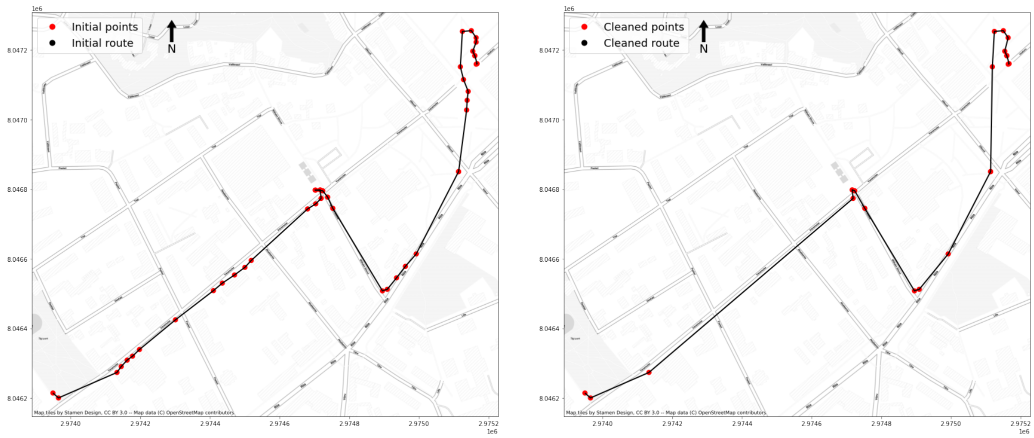


Figure 3. Comparison of a trajectory before and after the cleaning process.

To demonstrate the capabilities of the data cleaning method described above, an example based on a single trajectory is presented in Figure 3. The initial trajectory with 43 points is reduced to 21 points while retaining most of its original shape. It is visible that redundant points that are located on the same straight line of motion are removed. A detailed summary of the results of the geolocation data cleaning process by month is presented in Table 1.

Table 1. Results of geolocation data cleaning.

	Original nr. of trips	**Cleaned nr. of trips**	**Original nr. of points**	**Cleaned nr. of points**
January	6570	6519	443350	96587
February	4994	4937	324401	69986
March	10052	9966	838856	169445
April	63194	62073	5977021	1365391
May	109111	107378	9557246	2208448
June	140907	138077	12251211	2767218
July	128665	98831	11273862	1996976
August	110203	84851	8961266	1657692
September	112079	86318	7762143	1520129
October	92997	91926	6218019	1584544
November	53981	53311	3315475	855588
December	3512	3466	215225	48276
Total	**836265**	**747653**	**67138075**	**14340280**

4 Analysis

The analysis of the available data is split into three blocks, each addressing one of the research goals. Exploratory analysis (section 4.1) aims to give a general understanding of PBS usage patterns in the city of Tartu and cover RG1. In section 4.2 optimal locations for new potential PBS stations are proposed using the MULTIMOORA method as per RG2. The last section (section 4.3) follows RG3 and analyzes the road usage intensity in Tartu, specifically by Tartu Smart Bike users, trying to pinpoint road segments that could benefit from separate bicycle lanes.

4.1 Exploratory analysis

In this subsection, the usage patterns of Tartu Smart Bike are explored. Temporal analysis is performed in order to understand how Tartu's PBS system is used throughout the year. Next, users' sociodemographic characteristics such as gender and age are analyzed. Lastly, a small subsection describing station-level demand follows.

4.1.1 Temporal analysis

For the temporal analysis, the data is aggregated and analyzed in various time frames. For example, daily trip count in the span of the whole year is used to provide an overview of the travel patterns throughout the year, while smaller time-frames are used for more comprehensive analysis.

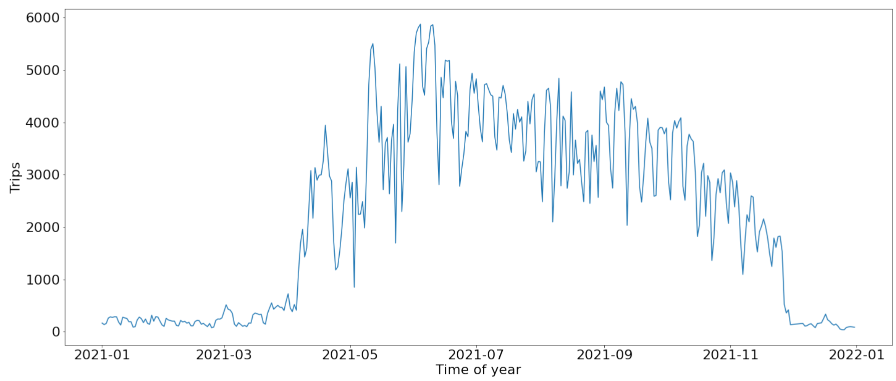


Figure 4. Yearly distribution of Tartu Smart Bike trips by day.

Figure 4 illustrates that the daily number of trips has a cyclic pattern throughout the year. This recurrent pattern can point to differing travel patterns during weekdays and weekends, which in turn suggest different intentions of PBS usage. In addition, the daily trip counts can also be affected by other parameters, such as weather or public holidays

[25]. For example, a pretty significant plummet in bike trips can be observed at the beginning of May. Further analysis reveals that this plummet occurred on the 3rd of May and was most likely caused by weather conditions, as the locals of Tartu experienced low temperatures and heavy rain (16.9 mm) that day [41].

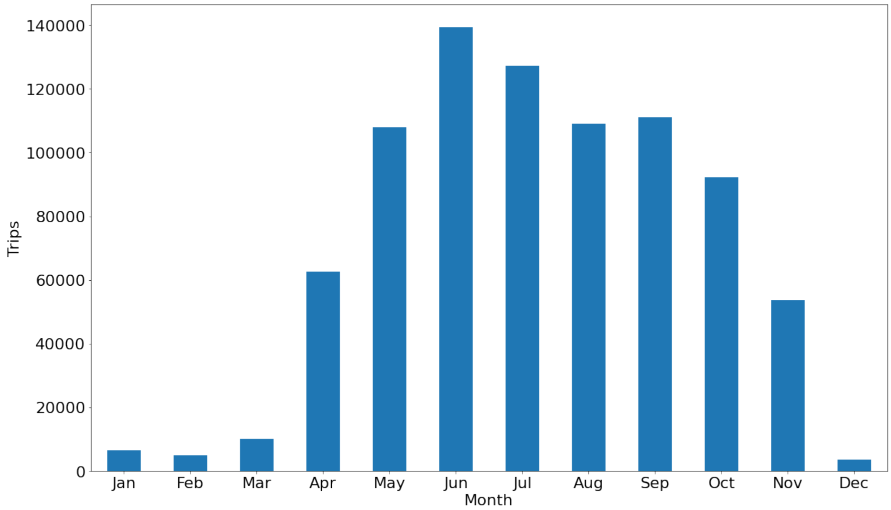


Figure 5. Monthly Tartu Smart Bike trip counts.

The distribution of Tartu Smart Bike trips by month (Fig. 5) is subject to the expected spread, with the majority of trips happening during the warmer months of summer and autumn. As the average temperature in Estonia between December and March is below freezing [42], the PBS demand is much lower during the winter season. The significantly lower number of trips during the winter period is also attributable to the fact that two-thirds of the PBS bicycles in Tartu are removed in the winter [12]. The highest number of trips took place during June, with a total of 137718 trips, which makes up for 16.8% of all trips that year.

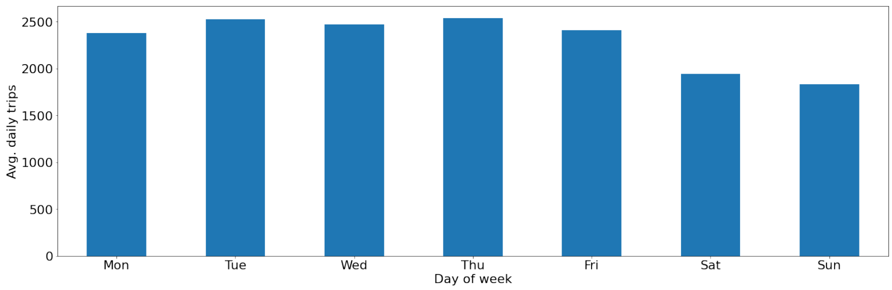
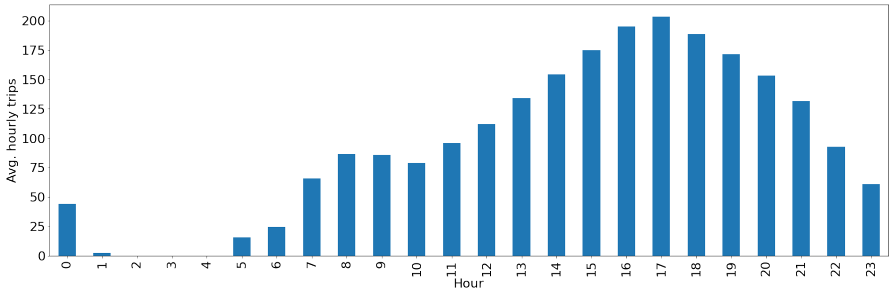
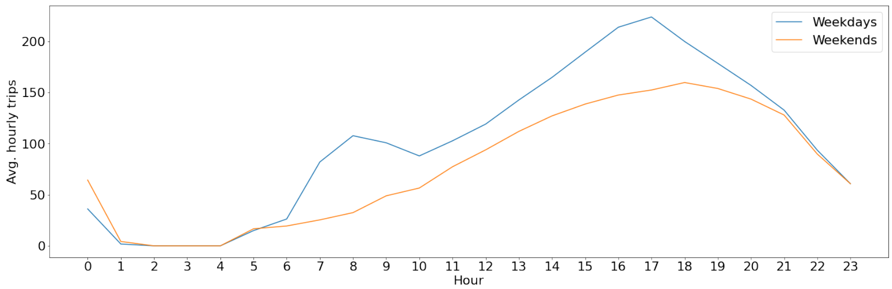


Figure 6. Daily average Tartu Smart Bike trips by weekday.

Figure 6 shows a significant difference between the number of trips during weekdays and weekends, which can account for the oscillating yearly distribution in Figure 4. The average number of trips during a weekday is 2464, while the same indicator for the weekends is 1889. In other words, the average daily trip count during weekdays is 30.4% higher than that on the weekends. This difference suggests that Tartu Smart Bike share is used primarily for commuting and less for leisure purposes.



(a)



(b)

Figure 7. (a) Average number of Tartu Smart Bike trips per hour. (b) Comparison of hourly trip counts on weekdays and weekends.

The hourly distribution of trips (Fig. 7a) shows a predictable pattern, with most trips happening during the daytime and significantly fewer trips taking place during the night. There are no trips from 2 a.m. until 5 a.m. as the PBS system is closed for customers during these hours to allow the electric-assist bicycles to recharge [12]. Slight peaks can be observed around 8 a.m. and 5 p.m., pointing to utilitarian usage of the bike share system. The average number of trips per hour is 93, while the most trips on average happen between 5 p.m. and 6 p.m. (201 trips on average).

The utilitarian usage pattern is more evident from the comparison of average hourly trips during weekdays and weekends (Fig. 7b). The peaks at 8 a.m. and 5 p.m. are clearly visible during the working days. In comparison, the distribution of trips during the weekends is more even, with significantly fewer trips taking place during the morning hours.

4.1.2 User analysis

To have a better understanding of the users of Tartu Smart Bike, age and gender of the users is analyzed briefly. The youngest user according to the data was 13, while the oldest registered user's age was 103. It is noteworthy that Tartu Smart Bike system does not officially allow users younger than 14 years of age to register for the bicycle sharing system [12], yet there was a small number of trips made by users aged 12-13. This indicates that the age data might not be entirely correct. Figure 8 displays the distribution of total trips made by users of different age groups. As could be expected, the majority of trips were made by teenagers (aged 14-18) and young adults (aged 19-26), accounting in total for 50% of all trips made during 2021. The former age group is chosen to mostly encompass students finishing their basic and secondary education, while the latter age group is meant to contain a large number of students of higher education institutions [43]. The number of trips drops off significantly for users aged 27 and older.

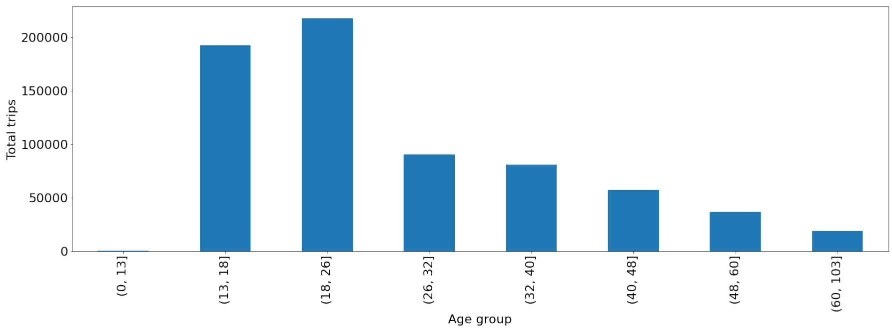
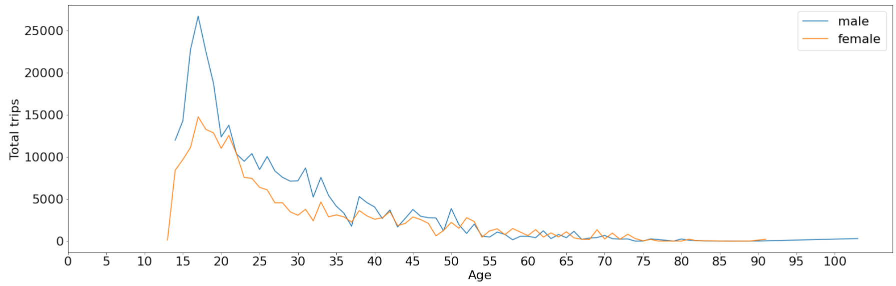
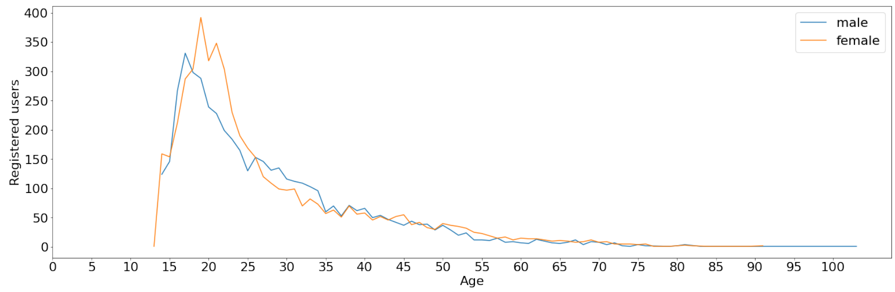


Figure 8. Age distribution of Tartu Smart Bike users.

For the gender analysis, the gender attribute was derived from the existing *id_code* attribute, using the first digit of the users' Estonian national identification number. Since the identification number was not present for all of the users, the gender analysis is based on approximately 66% of the total number of trips. Male users account for 58.6% of all the trips, while trips made by female users form 41.4% of the total trips. While there is a considerable difference in trips made by male and female users, the difference is relatively small compared to many other PBS systems around the world, as the gender gap in bicycle sharing is a substantial issue worldwide. For example, women account for only a quarter of all PBS trips in multiple major cities in the USA, such as New York, Chicago, and Boston [44]. A similar trend can be observed in Europe, as the female users account for 30% of all PBS users in Paris [45] and 27% of all PBS users in London [46].



(a)



(b)

Figure 9. (a) Age distribution of Tartu Smart Bike trips by gender. (b) Age distribution of Tartu Smart Bike registrations by gender.

The most significant difference in trip rates between male and female users can be observed in youngsters between 14 and 20 years of age (Fig. 9a). While the gap in trips made by female users and male users is sizeable, the split between the number of registered users by gender is almost even, with 48% of the registered users being women (Fig. 9b). Thus, while having almost the same amount of male and female users, Tartu Smart Bike is used considerably more by its male users.

4.1.3 Station-level demand

A total of 90 stations were in use during the year 2021. While 8 of these stations are located beyond the city limits, this analysis concentrates on the 82 stations that are located in the city. Figure 10 shows the locations of these 82 stations along with their monthly production. Monthly production and attraction of trips was calculated for all of the stations using the trip records from 2021. As expected, the stations with the highest demand are mainly situated in the city center and the most population-dense district of Annelinna.

The ten most popular stations sorted by their demand can be found in Table 2. The station with the highest monthly production is Uueturu station, which produces over a thousand trips more than any other station per month. This is probably due to its very central location and proximity to shopping centers and study buildings. To support this

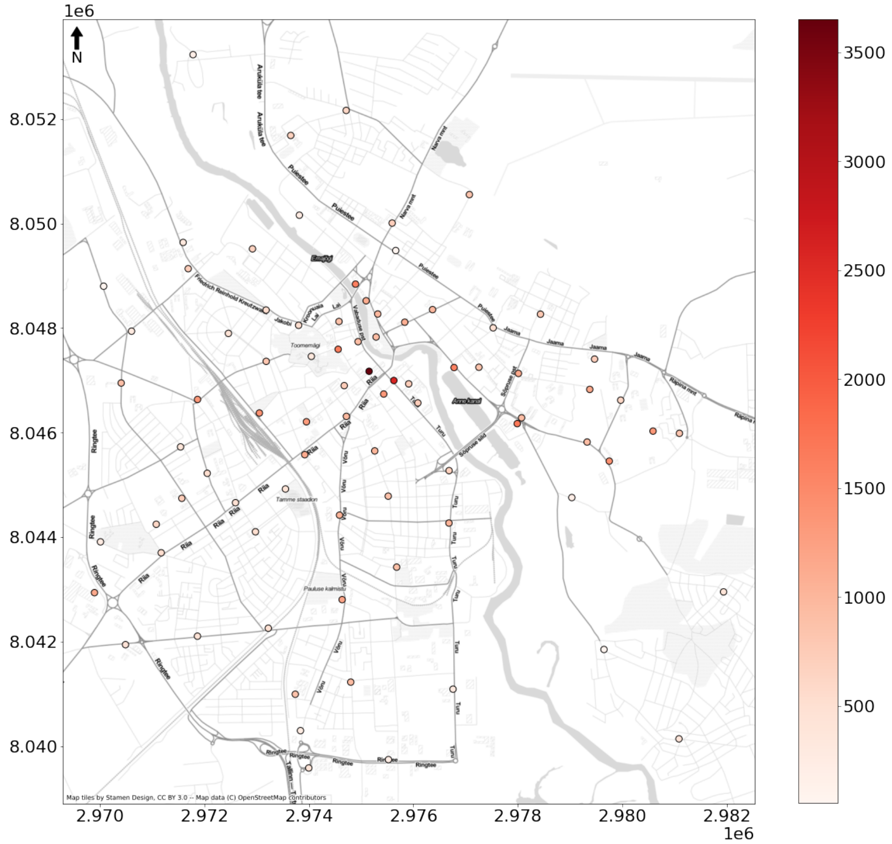


Figure 10. Monthly production of Tartu Smart Bike stations in Tartu.

claim, all the shopping malls, schools, and university buildings were downloaded using Overpass, and the distance from each station to each point of interest was calculated. As can be seen in Table 2, all of the most utilized stations are within 600 meters of the nearest school building, which also corresponds well to the high number of teenagers using Tartu Smart Bike share, as stated in the previous subsection. Additionally, the three most popular stations are also located next to shopping malls, and all ten stations are closer or equally close to shopping centers than the average of 82 stations which is 0.87 km. While the stations located in Annelinna residential district are located further from university buildings, the rest of the stations also demonstrate a tendency to be situated close to university buildings.

Table 2. Ten most used Tartu Smart Bike stations.

Name	**Monthly production**	**Monthly attraction**	**Closest university**	**Closest school**	**Closest mall**
Uueturu	3652	3762	0.35 km	0.45 km	0.07 km
Soola	2583	2709	0.6 km	0.48 km	0.05 km
Eeden	1694	1732	1.27 km	0.42 km	0.09 km
Pirogovi plats	1645	1741	0.07 km	0.24 km	0.31 km
Vabadussild	1625	1700	0.09 km	0.24 km	0.87 km
Turusild	1585	1597	1 km	0.28 km	0.5 km
Annelinna keskus	1408	1429	2.18 km	0.58 km	0.18 km
Raudteejaam	1393	1228	0.44 km	0.28 km	0.65 km
Annelinna kiir	1347	1340	2.61 km	0.35 km	0.69 km
Veeriku	1337	1265	0.6 km	0.59 km	0.08 km

4.2 Station location selection

To make suggestions for potential new PBS stations for Tartu Smart Bike, multi-criteria decision-making (MCDM) approach is used. MCDM is a technique that aims to find the best option from a set of candidate alternatives with respect to multiple criteria. Various MCDM models have been proposed for PBS station location selection, including Analytic Hierarchy Process (AHP) [47] and TOPSIS [48]. In this thesis, MULTIMOORA (Multi-Objective Optimization on the basis of a Ratio Analysis plus the full MULTIplicative form) is used to rank potential locations for new PBS stations. MULTIMOORA yields an aggregated ranking from three ranking methods - Ratio System, Reference Point Approach, and Full Multiplicative Form - by minimizing and maximizing multiple relevant criteria [33]. A similar approach was used in a study by Kabak *et al.* [31], which has been partly the basis of station location selection in this thesis.

Five criteria are used to assess the suitability of potential station locations in this thesis:

1. **Cell population.** This criterion is chosen based on a simple assumption that stations that are located in areas with higher population density are accessible to more people. Population data with a precision of 100×100 meters is used to accurately determine the population of each candidate region.

2. **Demand of the closest existing PBS station.** To assess the demand for existing stations, average monthly trip production was calculated for each station based on the trip level data from the year 2021. The high demand of the closest station signifies if that region is already popular among PBS users.

3. **Distance to the closest existing station.** As one of the main objectives of public bicycle sharing is accessibility, it is rational to spread stations over the study area. Thus potential locations further from existing stations are considered better. Additionally, this criterion is also later used as a threshold for location selection, as it is unreasonable to locate multiple stations next to each other.

4. **Distance to the closest study building.** Since exploratory analysis reveals that the majority of Tartu Smart Bike users are teenagers and young adults likely to still be studying (Fig. 8), distance to the closest study building is used as one of the criteria. All the study buildings - schools and higher education buildings - were fetched using the Overpass API.

5. **Distance to the closest shopping mall.** As per the analysis of station-level demand, almost all of the most popular Tartu Smart Bike stations are located near shopping malls. Thus, distance to the closest shopping mall is chosen as one of the criteria. This also helps to shift the focus from school and university students towards a wider target group.

To find the optimal location, criteria 1-3 are maximized and criteria 4-5 are minimized.

Candidate locations are selected by creating a grid covering the city of Tartu by generating points evenly 300 meters apart from each other. Next, newly created cell centroids are mapped to the closest 100×100 meter population grid cell centroid in order to be able to use the accurate population data. Every cell of this slightly shifted grid is considered a candidate location. Figure 11 presents the 100×100 meter population grid as a heat map and the centroids of candidate locations using blue dots. 430 potential locations were generated and used in the MULTIMOORA model.

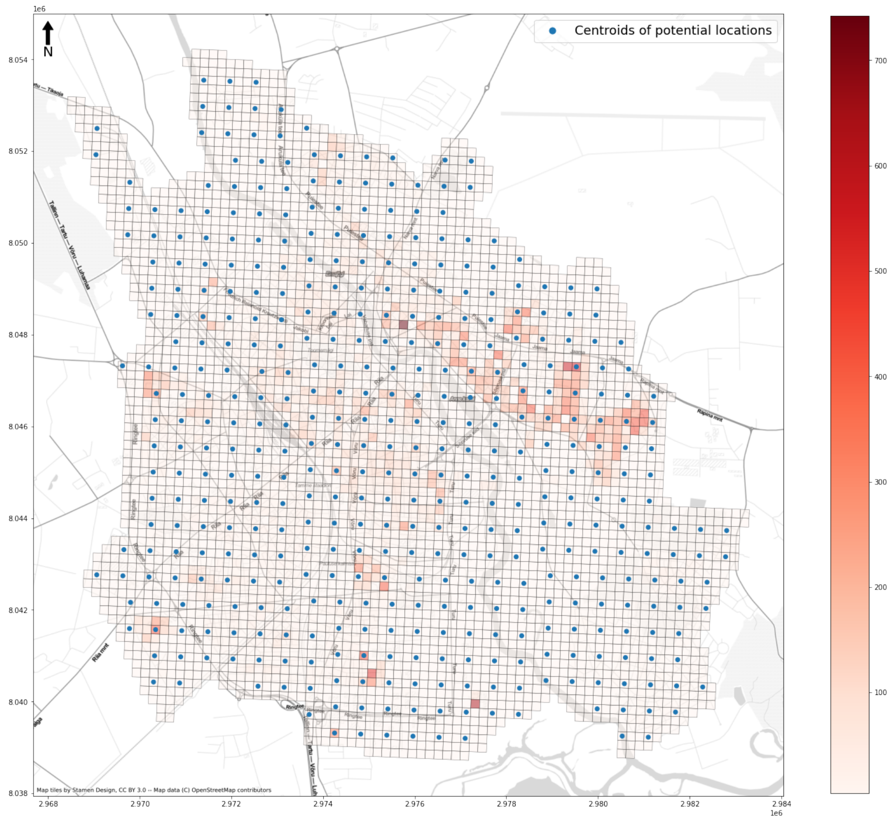


Figure 11. Population heat map with candidate location centroids for station location selection.

For each of the candidate locations all the aforementioned criteria are calculated and stored in a decision matrix X:

$$X = \begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{m,1} & \cdots & x_{m,n} \end{bmatrix},$$

where m is the number of candidate locations, n is the number of criteria and $x_{i,j}$ is the response of candidate j on criterion i, where $j \in [1, 2, \ldots, m]$ and $i \in [1, 2, \ldots, n]$.

Next, the values $x_{i,j}$ are normalized:

$$x^*_{i,j} = \frac{x_{i,j}}{\sqrt{\sum_{j=1}^{m} x^2_{i,j}}}$$

The first ranking system used in MULTIMOORA is **Ratio System**. In the Ratio System approach, the suitability of each candidate is calculated by adding the normalized responses of beneficial criteria and deducting the normalized responses of non-beneficial criteria:

$$y^*_j = \sum_{i=1}^{g} x^*_{i,j} - \sum_{i=g+1}^{n} x^*_{i,j},$$

where g is the number of criteria to be maximized, and $n - g$ is the number of criteria to be minimized. Based on the calculated scores, the alternatives are ranked, and the one with the highest value is deemed to be best alternative.

Reference Point approach ranks the alternatives based on their worst performance in any of the criteria. For each criterion, the maximal value among all candidates is chosen as a reference point (r_i) for criteria to be maximized, and the minimum value among all candidates is chosen as a reference point for criteria to be minimized. Then the distance of each alternative to the reference point in terms of each criterion is calculated, and the final ranking is obtained by choosing the best (minimal) value from these distances:

$$\min_{(j)}\{\max_{(i)} |r_i - x^*_{i,j}|\}$$

The final part of the MULTIMOORA model is the **Full Multiplicative Form**. In Full Multiplicative Form, the criteria to be maximized of each candidate is multiplied and divided by the product of criteria to be minimized of each candidate:

$$u_j = \prod_{i=1}^{g} x^*_{i,j} \; / \prod_{i=g+1}^{n} x^*_{i,j}$$

All three ranking methods are combined in order to determine the final ranking of the candidate locations. A heat map representing the suitability of each candidate location is presented in Figure 12, where darker areas represent candidate locations with higher suitability ranks. As a general rule, the best candidate locations were situated in the residential regions with the highest population density and the city center, which holds the most points of interest, such as schools, university buildings, and malls.

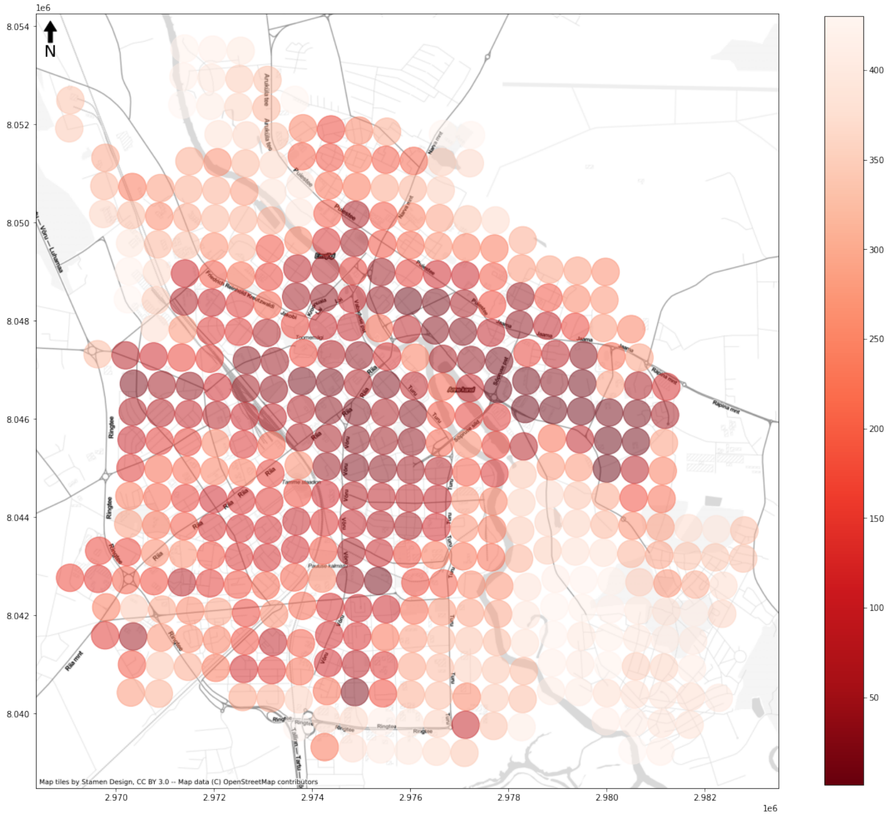
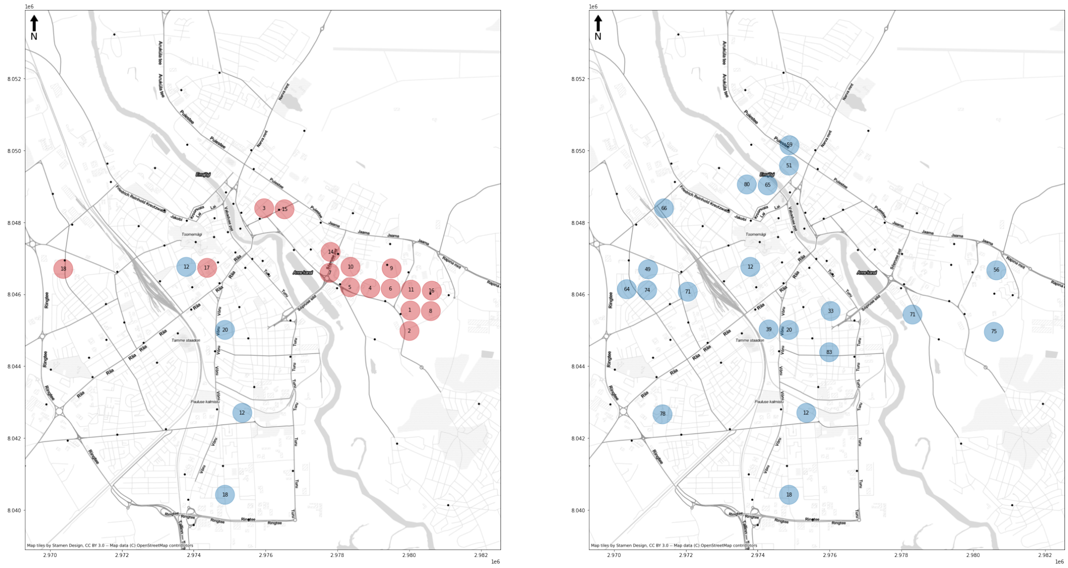


Figure 12. Results of the MULTIMOORA ranking model.

Figure 13 displays more detailed results of MULTIMOORA ranking model. Red circles denote the candidate locations that are within 300 meters of the nearest existing PBS station (Fig. 13a), while the blue circles denote the candidate locations that are further than 300 meters from existing PBS stations. The numbers within the circles signify the candidate locations' final ranking according to the MULTIMOORA model, and black dots indicate existing Tartu Smart Bike stations. Since the population has a significant weight on the final ranking, the most population-dense regions in Annelinna district get the highest suitability scores. Removing the candidate locations that are within 300 meters of existing PBS stations (Fig. 13b) yields more serviceable results that are spread more evenly along the city.



(a) Best candidates including locations within 300 meters of an existing PBS station (red).

(b) Best candidate locations further than 300 meters from the closest existing station (blue).

Figure 13. Best candidate locations according to the MULTIMOORA model.

4.3 Road network analysis

Next, the cleaned geolocation data is used to conduct road network analysis and pinpoint the road segments that are most utilized by Tartu Smart Bike users. To achieve that, the PBS trip trajectories are mapped to the road network of Tartu, and the road usage intensity for each road segment is calculated. Then the most utilized road segments can be extracted and analyzed in order to locate those segments that do not enclose a separate bicycle lane or path.

The full road network covering the city of Tartu and its surrounding areas is downloaded from OpenStreetMap via Overpass API[2]. Additionally, the cycling network is downloaded separately, and these two networks are merged, adding the Boolean *bike_path* attribute to all the road segments so that the bike paths would later be easily distinguishable in the road network. The resulting full network is converted to a graph. Then all the cleaned trajectories are map-matched to this road network one by one, matching the segments in a given trajectory to Tartu road segments. The graph is used to connect the matched road segments into a single connected line. This process is visualized in Figure 15. The black line represents the initial cleaned PBS trajectory, while the red lines denote the results of map-matching. This method is somewhat sensitive to

[2] https://wiki.openstreetmap.org/wiki/Overpass_API

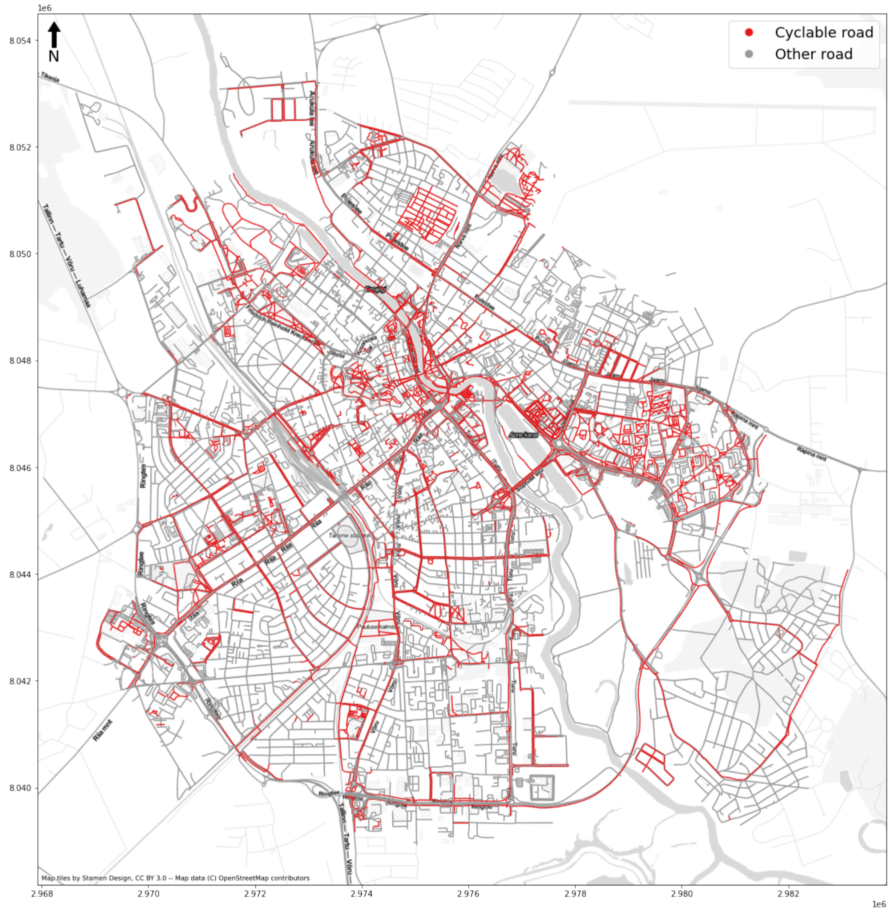
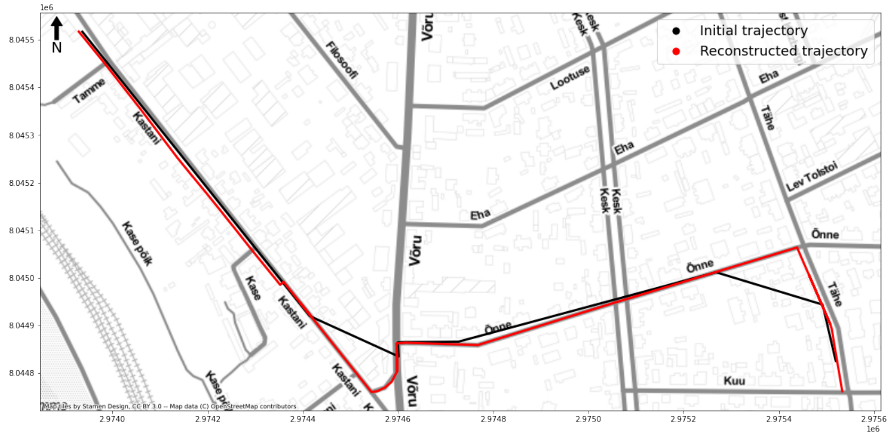


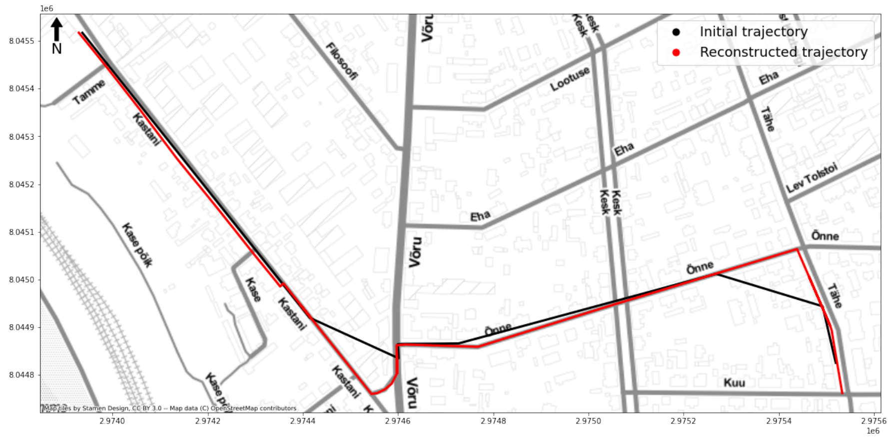
Figure 14. Cycling network of Tartu.

the error in GPS location, as spatial join is used to map-match the trajectory segments to the closest road network segments. Since a road can encompass multiple lanes with different purposes (pedestrian roads, bicycle paths, car lanes, etc.), which are treated as different roads by OpenStreetMap, a small fluctuation in the GPS coordinates can lead to the trajectory shifting between different purpose lanes along one road. On the other hand, it provides a solution to another issue. Tartu Smart Bikes report their location in a periodic manner every five seconds with an occasional pause of 40 seconds. If a user makes a turn during such a pause, this turn can be lost in their trajectory and results in a user seemingly cutting diagonally, possibly through buildings and other infrastructure. Since graph-based map-matching is used to connect the matched segments, the final trajectory only contains actual road segments, as can be seen in Figure 15.

Using this method, all of the Tartu Smart Bike trips from 2021 are mapped to the full road network, and the monthly usage intensity for each road segment is calculated.



(a)



(b)

Figure 15. (a) Mapping a PBS trajectory to road network segments. (b) Constructing the final trajectory using a graph based on the matched segments.

The results can be found in Figure 16. As could be expected, the areas in the city center experience the highest usage intensity. Bridges form bottlenecks with especially high intensity since Tartu is located on both banks of river Emajõgi and crossing the river is only possible by one of the seven bridges.

In order to pinpoint the road segments that could benefit from a separate bicycle path, only the road segments that do not include a bicycle path are considered. The results are illustrated in Figure 17, which depicts all those road segments that do not include a bicycle path and experience more than 1000 trips monthly.

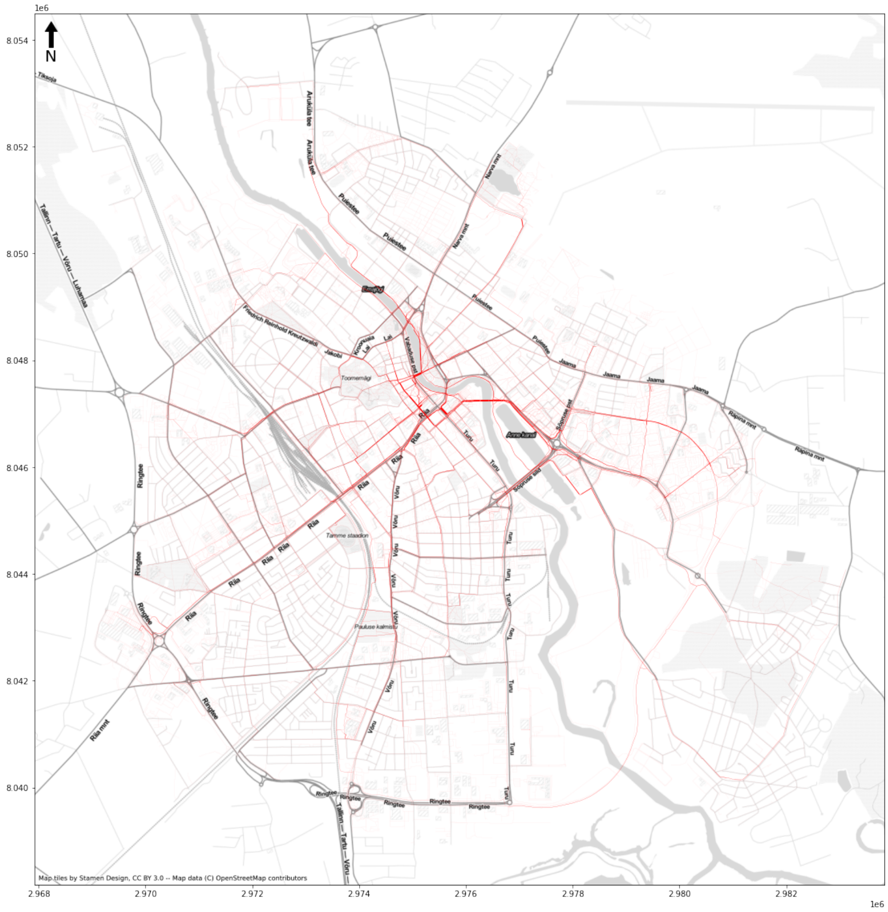


Figure 16. Monthly average road usage intensity of PBS users in Tartu. The thickness of the line signifies the intensity on the given road segment.

The results contain a significant amount of noise, such as tiny segments scattered around the city that get matched a lot as parts of various different longer trajectories. Additionally, some error is present due to obsolete data from OpenStreetMap. For example, parts of Vanemuise street have been reconstructed recently, including adding separate bicycle lanes [49], but this change is not yet present in OpenStreetMap, rendering the intensity along Vanemuise street incorrect. Despite those shortcomings, some longer road sections are identified that experience a significant amount of trips monthly yet do not include a designated lane for bicycles. Such sections include the majority of Lai street, Aleksandri street, and parts of Võru street, among others. The identified sections should be further analyzed by a domain expert in order to assess the need and suitability for designated cycling infrastructure.

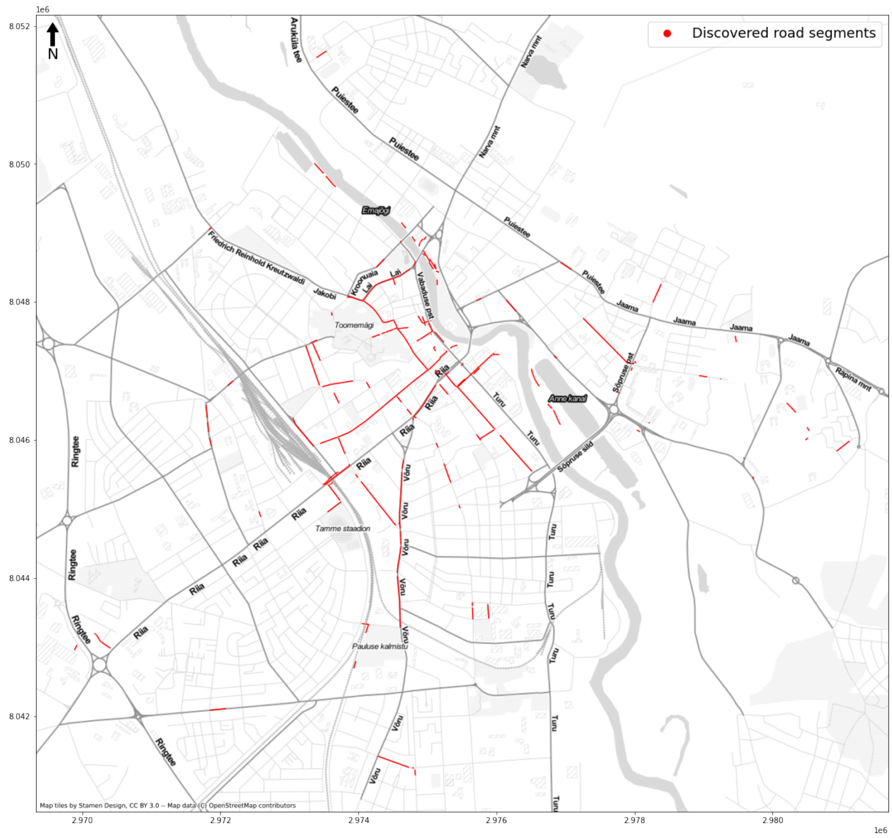


Figure 17. Road segments without dedicated bicycle path and with more than 1000 monthly trips on average.

5 Discussion

In this chapter, the results of the completed analysis and further opportunities are discussed. In addition, the challenges faced, and possible future work is presented.

5.1 Tartu cycling strategy and action plan

In 2021, a strategic action plan for improving bikeability in Tartu [50] was created by the city of Tartu in cooperation with HeiVäl Consulting[3]. The final goal of this plan is to increase the share of bicycle trips in the city by one percent per year until the year 2040, achieving the modality share of 26% for cyclists. This action plan consists of clearly defined goals, aiming to improve bikeability in Tartu. The primary goal is to decrease the number of personal vehicles in daily traffic. The analysis presented in this thesis could provide valuable support for achieving these goals. One of the first steps in the action plan is understanding and mapping the motivations of pedestrians and cyclists in the city (strategic goal #12), which includes Tartu Smart Bike users. Another goal described in the action plan is increasing the overall length of designated cycling lanes in Tartu (strategic goal #11), which is strongly related to the road network analysis conducted in this thesis. Multiple other goals, including increasing the number of cyclists (strategic goal #3) and decreasing CO emissions and noise (strategic goal #2), are intertwined with public bicycle sharing and are considered some of its main advantages. Thus, PBS data and the analysis presented in this thesis could be used to further improve the action plan and achieve its goals.

5.2 Assessing the station location selection model

Assessing the results of a PBS station location selection model is not an easy task. Most of the related works do not include a specific method for evaluating the results of a station location selection model. In the case of Tartu Smart Bike, some confidence in the results can be bestowed by looking into the actual decisions made by city officials in regard to PBS station location selection.

In early 2022, five new PBS stations were opened in Tartu, one of them located outside of the city bounds [51]. Figure 18 displays four new stations established in Tartu in 2022, along with the already existing PBS stations and the proposed locations based on the MULTIMOORA model. Two of the new stations (Kiigeplatsi and Savi) are located within a proposed candidate area that is among the 20 best candidate locations further than 200 meters from an existing PBS station. Another new station (Nõlvaku) is on the border of such a candidate location. Thus, the results of the automated MULTIMOORA model

[3] https://heival.ee/

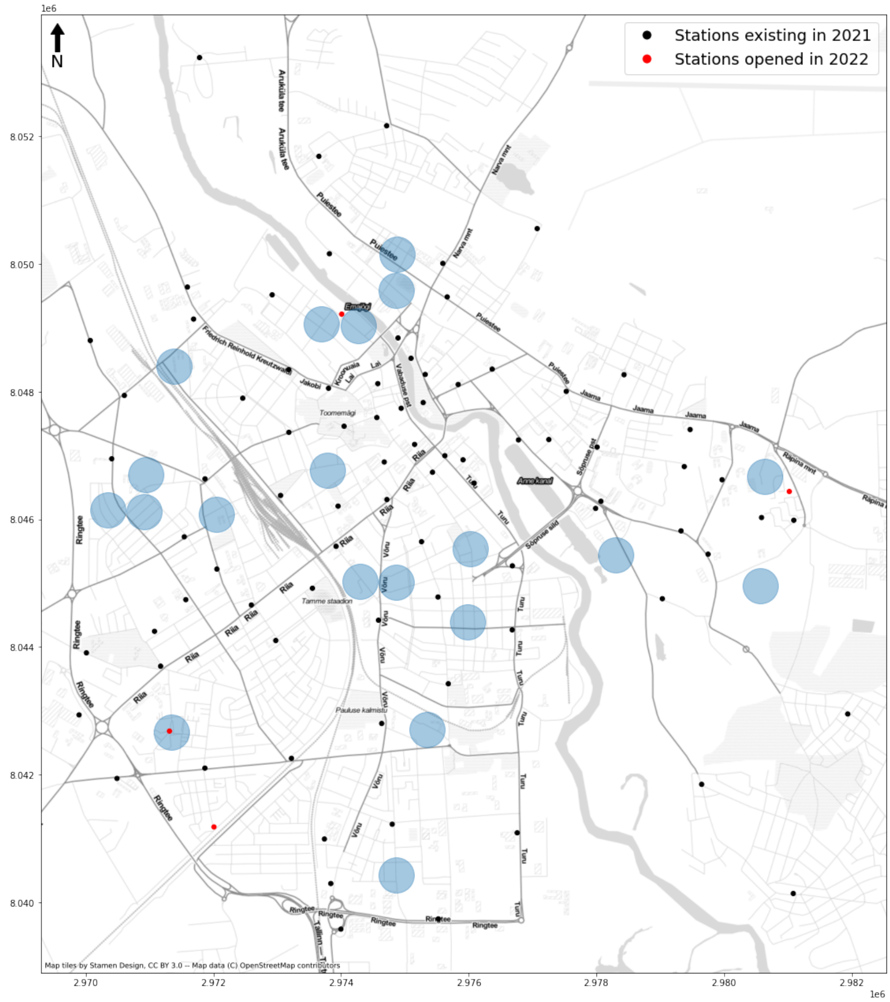


Figure 18. Locations of existing and new PBS station in regards to proposed locations. Blue areas denote the 20 best locations further than 200 meters from an existing station.

seem to somewhat correspond to the beliefs and decisions of the Tartu city government, giving a reason to have confidence in the results provided.

The proposed MULTIMOORA model could be further improved by exploring other relevant criteria and using more criteria in total. Other commonly used criteria include distance to pubs and restaurants, the proximity of housing facilities, and surrounding cycling infrastructure. In addition, MULTIMOORA model supports adding weights to the criteria, making the model even more customizable and adaptable. In this thesis, criteria weights were not added, as selecting relevant weights is a significant task on its own, often requiring the help of domain experts.

5.3 Improvements to the road network analysis

The proposed map-matching method for quantifying the road usage intensity on various road segments provides promising results that should be further analyzed by a local city official for best results. Still, the method could benefit from some technical improvements. Currently, the whole road network of Tartu is used in the map-matching process, where roads of all types, including driveways, cycling paths, and pedestrian roads, are present. In other words, if a street is comprised of all of these road types, they are also all separately represented in OSM. That, in turn, means that due to fluctuations in GPS accuracy, some trajectories can be matched to the wrong part of the street. A better solution would be to assume that a cycling lane is used when applicable, even if a driveway or a pedestrian road is closer to the points in the trajectory.

While this thesis concentrates on finding road segments that experience high usage intensity by PBS users, another common approach for providing decision support for cycling infrastructure improvements is identifying the most avoided road segments. Lu *et al.* [52] compare the routes taken between two PBS stations to the shortest path connecting them. Assuming users normally prefer the shortest way to reach their destination, analyzing the avoided road segments along the shortest paths can reveal limitations in the cycling infrastructure.

Additionally, if the speed of a PBS bicycle at each point is available in the GPS data, the average speed of PBS users along different segments could be analyzed. Pinpointing the road segments where the average speed is significantly below average could reveal more about cyclists' mobility patterns and help identify problematic road segments for cyclists.

6 Conclusion

This thesis aimed to analyze the novel data from Tartu Smart Bike in order to get a better understanding of the mobility patterns of cyclists in Tartu and provide decision support for improving the existing PBS system and cycling infrastructure in the city.

Firstly a short overview of the background and related works was presented to provide context and understanding of the topic. Then the data was cleaned and prepared to use for the analysis. A data cleaning algorithm was employed in order to shrink the volume of geolocation data, reducing the number of points by almost 80% percent. The cleaned data was then used to achieve the research goals. A thorough analysis of Tartu Smart Bike usage patterns was conducted, highlighting the utilitarian use of the PBS station, as well as identifying high-demand stations and providing a better understanding of Tartu Smart Bike users. Next, a model based on MULTIMOORA multi-criteria decision-making technique was presented to identify possible locations for new PBS stations. The criteria for assessing the suitability of candidate locations were carefully selected based on the previous analysis. By generating candidate locations uniformly as centroids in a grid overlaying the study area, the resulting model is easily adaptable for use in other cities. Lastly, the routes taken by Tartu Smart Bike users were analyzed in order to pinpoint road segments that exhibit a need for designated cycling infrastructure. The cleaned trajectories were map-matched to the road network of Tartu, counting road usage intensity for each road segment. The road segments with high usage intensity, but no cycling infrastructure, were identified.

The results of this thesis provide a better understanding of Tartu Smart Bike users' usage behavior and mobility patterns. This work could provide decision support for improving Tartu Smart Bike system and the cycling infrastructure of Tartu, as well as help future researchers conduct similar analysis in other areas of interest.

References

[1] P. DeMaio, "Bike-sharing: History, Impacts, Models of Provision, and Future," *The Journal of Public Transportation*, vol. 12, p. 3, 2009.

[2] Y. M. K. Sya'Bana and K. H. Sanjaya, "The Applicability of Sustainable Design Values on Electric Bike Sharing Concept in Indonesia," in *2019 International Conference on Sustainable Energy Engineering and Application (ICSEEA)*, pp. 125–130, 2019.

[3] X. Shi, Z. Yu, and X. Qu, "Touching PBSData — An exploratory study of visual analytics for public bicycle system supporting multi-touch device," in *2016 International Conference on Progress in Informatics and Computing (PIC)*, pp. 390–394, 2016.

[4] Tartu Postimees, "Hannes Luts: hinnatõus, mitte transpordipoliitika," 2022. https://tartu.postimees.ee/7499314/hannes-luts-hinnatous-mitte-transpordipoliitika, (02.08.2022).

[5] Tartu Postimees, "KIRI: ohtlikud rendirattad," 2019. https://tartu.postimees.ee/6775722/kiri-ohtlikud-rendirattad, (02.08.2022).

[6] Tartu Postimees, "Rannar Raba: Tartu rattaringlus ei täida eesmärki. Ratas peab ootama mind, mitte mina ratast," 2019. https://tartu.postimees.ee/6712530/rannar-raba-tartu-rattaringlus-ei-taida-eesmarki-ratas-peab-ootama-mind-mitte-mina-ratast, (02.08.2022).

[7] Tartu Postimees, "Ringlusrattad sõitsid aastaga ligi 2,5 miljonit kilomeetrit," 2020. https://tartu.postimees.ee/6992000/ringlusrattad-soitsid-aastaga-ligi-2-5-miljonit-kilomeetrit, (02.08.2022).

[8] S. A. Shaheen, S. Guzman, and H. Zhang, "Bikesharing in Europe, the Americas, and Asia: Past, Present, and Future," *Transportation Research Record*, vol. 2143, no. 1, pp. 159–167, 2010.

[9] C. Yu, O. O'Brien, P. DeMaio, R. Rabello, S. Chou, and T. Benicchio, "The Meddin Bike-sharing World Map: Mid-2021 Report," tech. rep., PBSC Urban Solutions, 2021.

[10] A. Duz, M. Corno, and S. M. Savaresi, "Is Charge Sustaining Achievable in Electric Free-Floating Bicycle Sharing?," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–6, 2020.

[11] M. Corno, A. Duz, and S. M. Savaresi, “Design of a Charge-Sustaining Energy Management System for a Free-Floating Electric Shared Bicycle,” *IEEE Transactions on Control Systems Technology*, pp. 1–13, 2021.

[12] Tartu City Government, “Tartu Smart Bike,” 2021. https://ratas.tartu.ee/, (21.03.2022).

[13] R. Meeksa, “Presentation in data science seminar: Tartu smart bike share – how and what do we see.” https://www.uttv.ee/naita?id=28980, (27.03.2022).

[14] Tartu City Government, “Tartu in figures 2020/2021.” https://tartu.ee/sites/default/files/uploads/Tartu%20linn/Statistika/Tartu_figures_2021_ENG.pdf, (20.04.2022).

[15] Tartu City Government, “Tartu statistiline ülevaade 2017. Asend ja keskkond..” https://www.tartu.ee/sites/default/files/uploads/Statistika/2017/Tartu_stat_aastaraamat_veeb.pdf, (20.04.2022).

[16] C. Shui and W. Szeto, “A review of bicycle-sharing service planning problems,” *Transportation Research Part C: Emerging Technologies*, vol. 117, p. 102648, 2020.

[17] S. Ram, F. Dong, F. Currim, Y. Wang, E. Dantas, and L. A. Sabóia, “SMART-BIKE: Policy making and decision support for bike share systems,” in *2016 IEEE International Smart Cities Conference (ISC2)*, pp. 1–6, 2016.

[18] X. Shi, Z. Yu, H. Xu, and J. Chen, “PBikeVis: Applied Visual Analytics for Public Bicycle System,” in *2015 8th International Symposium on Computational Intelligence and Design (ISCID)*, vol. 1, pp. 490–493, 2015.

[19] X. Shi, Q. Zhou, X. Qu, G. Liu, and Z. Gong, “Understanding city dynamics based on public bicycle data: A case study in Hangzhou,” in *2016 10th International Conference on Software, Knowledge, Information Management Applications (SKIMA)*, pp. 146–150, 2016.

[20] X.-F. Xie and Z. J. Wang, “Examining travel patterns and characteristics in a bikesharing network and implications for data-driven decision supports: Case study in the Washington DC area.,” *Journal of Transport Geography*, vol. 71, pp. 84 – 102, 2018.

[21] Y. Zhang, H. Wen, F. Qiu, Z. Wang, and H. Abbas, “iBike: Intelligent public bicycle services assisted by data analytics.,” *Future Generation Computer Systems*, vol. 95, pp. 187 – 197, 2019.

[22] Y. Guo, X. Shen, Q. Ge, and L. Wang, "Station Function Discovery: Exploring Trip Records in Urban Public Bike-Sharing System," *IEEE Access*, vol. 6, pp. 71060–71068, 2018.

[23] Q. He, J. Liao, J. Wang, and Z. Bao, "Visual Analysis of Shared Bicycle Data Based on User's Spatiotemporal Behavior Characteristics," in *2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, pp. 2065–2071, 2020.

[24] A. Maulit, Y. Baiburin, M. Rakhymbek, G. Sadykova, and A. Nugumanova, "Statistical and Network Analysis of Shared Bikes – In the Case of Almaty Bike," in *2021 IEEE International Conference on Smart Information Systems and Technologies (SIST)*, pp. 1–5, 2021.

[25] K. Kim, "Investigation on the effects of weather and calendar events on bike-sharing according to the trip patterns of bike rentals of stations.," *Journal of Transport Geography*, vol. 66, pp. 309 – 320, 2018.

[26] J. Zhang, X. Pan, M. Li, and P. S. Yu, "Bicycle-Sharing System Analysis and Trip Prediction," in *2016 17th IEEE International Conference on Mobile Data Management (MDM)*, vol. 1, pp. 174–179, 2016.

[27] J. Wood, R. Beecham, and J. Dykes, "Moving beyond sequential design: Reflections on a rich multi-channel approach to data visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 2171–2180, 2014.

[28] L. Caggiani, A. Colovic, and M. Ottomanelli, "An equality-based model for bike-sharing stations location in bicycle-public transport multimodal mobility," *Transportation Research Part A: Policy and Practice*, vol. 140, pp. 251–265, 2020.

[29] J.-R. Lin, T.-H. Yang, and Y.-C. Chang, "A hub location inventory model for bicycle sharing system design: Formulation and solution," *Computers & Industrial Engineering*, vol. 65, p. 77–86, 05 2013.

[30] S. Banerjee, M. M. Kabir, N. K. Khadem, and C. Chavis, "Optimal locations for bikeshare stations: A new GIS based spatial approach," *Transportation Research Interdisciplinary Perspectives*, vol. 4, p. 100101, 2020.

[31] M. Kabak, M. Erbaş, C. Çetinkaya, and E. Özceylan, "A GIS-based MCDM approach for the evaluation of bike-share stations," *Journal of Cleaner Production*, vol. 201, pp. 49–60, 2018.

[32] D. Guler and T. Yomralioglu, "Location Evaluation of Bicycle Sharing System Stations and Cycling Infrastructures with Best Worst Method Using GIS," *The Professional Geographer*, vol. 73, pp. 535–552, 04 2021.

[33] A. Hafezalkotob, A. Hafezalkotob, H. Liao, and F. Herrera, "An overview of MULTIMOORA for multi-criteria decision-making: Theory, developments, applications, and challenges," *Information Fusion*, vol. 51, pp. 145–177, 2019.

[34] A. Liu, R. Wang, J. Fowler, and X. Ji, "Improving bicycle sharing operations: A multi-criteria decision-making approach," *Journal of Cleaner Production*, vol. 297, p. 126581, 2021.

[35] E. Eren and B. Y. Katanalp, "Fuzzy-based GIS approach with new MCDM method for bike-sharing station site selection according to land-use types," *Sustainable Cities and Society*, vol. 76, p. 103434, 2022.

[36] D. Scott, W. Lu, and M. Brown, "Route choice of bike share users: Leveraging GPS data to derive choice sets," *Journal of Transport Geography*, vol. 90, p. 102903, 01 2021.

[37] R. Pritchard, "Revealed Preference Methods for Studying Bicycle Route Choice—A Systematic Review," *International Journal of Environmental Research and Public Health*, vol. 15, no. 3, 2018.

[38] F. Rupi, C. Poliziani, and J. Schweizer, "Data-driven Bicycle Network Analysis Based on Traditional Counting Methods and GPS Traces from Smartphone," *ISPRS International Journal of Geo-Information*, vol. 8, p. 322, 07 2019.

[39] L. Wysling and R. S. Purves, "Where to improve cycling infrastructure? Assessing bicycle suitability and bikeability with open data in the city of Paris," *Transportation Research Interdisciplinary Perspectives*, vol. 15, p. 100648, 2022.

[40] X. Yang, L. Tang, and Q. Li, "A Data Cleaning Method for Big Trace Data Using Movement Consistency," *Sensors*, vol. 18, 03 2018.

[41] Estonian Environment Agency, "Weather." `https://www.ilmateenistus.ee/`, (20.04.2022).

[42] Climates to travel: World climate guide, "Estonian climate." `https://www.climatestotravel.com/climate/estonia`, (14.04.2022).

[43] Estonian Public Broadcasting, "Statistics Estonia: 66 percent of young people studied in 2016," 2017. `https://news.err.ee/606031/statistics-estonia-66-percent-of-young-people-studied-in-2016`, (12.07.2022).

[44] K. Hosford and M. Winters, "Quantifying the Bicycle Share Gender Gap," *Transport Findings*, 2019.

[45] A. Gorrini, R. Choubassi, F. Messa, W. Saleh, A. Ababio-Donkor, M. Leva, L. D'Arcy, F. Fabbri, D. Laniado, and P. Aragón, "Unveiling Women's Needs and Expectations as Users of Bike Sharing Services: The H2020 DIAMOND Project," *Sustainability*, vol. 13, p. 5241, 05 2021.

[46] F. Ogilvie and A. Goodman, "Inequalities in usage of a public bicycle sharing scheme: Socio-demographic predictors of uptake and usage of the London (UK) cycle hire scheme," *Preventive Medicine*, vol. 55, no. 1, pp. 40–45, 2012.

[47] E. Eren and B. Y. Katanalp, "Fuzzy-based GIS approach with new MCDM method for bike-sharing station site selection according to land-use types," *Sustainable Cities and Society*, vol. 76, p. 103434, 2022.

[48] D. Guler and T. Yomralioglu, *Bicycle Station and Lane Location Selection Using Open Source GIS Technology*, pp. 9–36. Cham: Springer International Publishing, 2021.

[49] Tartu City Government, "Vanemuise tänava rekonstrueerimine," 2021. https://tartu.ee/et/vanemuise-tn, (25.07.2022).

[50] Tartu City Government: Department of Communal Services, "Tartu jalgrattaliikluse strateegiline tegevuskava," 2021. https://www.tartu.ee/sites/default/files/uploads/Linnavarad/SECAP/Tartu_jalgrattastateegia.pdf, (03.08.2022).

[51] Estonian Public Broadcasting, "Electric bikes returning to Tartu bikeshare, city to expand network," 2022. https://news.err.ee/1608561664/electric-bikes-returning-to-tartu-bikeshare-city-to-expand-network, (03.08.2022).

[52] W. Lu, D. M. Scott, and R. Dalumpines, "Understanding bike share cyclist route choice using GPS data: Comparing dominant routes and shortest paths," *Journal of Transport Geography*, vol. 71, pp. 172–181, 2018.

Appendix

I. Licence

Non-exclusive licence to reproduce thesis and make thesis public