Service Accessibility Analysis
Documentation

Compiled by SPIN Unit

Input data

Data sources

Origin:

- Cadastral address points (ADS)
- Building address points (ADS)
- Cadastral parcels (Maa-amet)

Destination:

- Service POIs (points, Maa-amet)
- Service POIs, RMK hiking network (linestring, Maa-amet)
- Supplementary grocery data (OpenStreetMap)
- Public green areas (based on TAR, Tallinn Strategic Management Office)
- Bike network - the core network of cycle paths and the health network of cycle paths (Tallinn Strategic Management Office)
- Education (Maa-amet POIs, EHIS - Eesti Hariduse Infosüsteem)
- Building uses (EHR)

Network:

- OpenStreetMap (OSM) data in the Protocolbuffer Binary Format, covering all of Estonia, from 26/01/2024 (Geofabrik)
- OpenStreetMap (OSM) cycling/pedestrian network (OpenStreetMap Overpass API)
- GTFS feed from 16/11/2022

Data processing

Destination processing:

This process involves categorizing various points of interest (POIs) and other relevant locations based on their type and significance. The categorization is done using a predefined scheme that classifies each destination into categories such as education, healthcare, nutrition, and leisure, among others.

The methodology involves several key steps:

- **The categorization of destinations:** each destination point is categorized based on specific attributes from the source datasets. For instance, educational institutions are identified and categorised based on their type (e.g., primary school, vocational school, university) using attributes from the input education data.
- **The processing of irregular destinations:** this includes destinations that do not fit into standard point formats, such as public green areas, the bike network, and the hiking network. These are processed differently to ensure they are accurately represented in the analysis. See the subsequent sub-section, *"Processing irregular destinations (public green areas, bike network, hiking network)"*, for more information.
- **Integration and finalisation:** All processed data, including standard and irregular destinations, are integrated into a comprehensive dataset. The outcome of this process is a detailed and categorized dataset of destination points, which forms the basis for subsequent accessibility analyses. This dataset not only includes point POIs but also a detailed representation of green spaces and the bike and hiking networks, providing a holistic view of the urban landscape in terms of service accessibility.

Below you will find the destination categorisation scheme used for this analysis. The categorisation scheme is based on the clsassification developed by the City of Tallinn, in keeping with their strategic planning documents. The source dataset, the categorisation column in that source dataset that is used, and the values within that categorisation column that map to the given category are listed.

Destination categorisation scheme:

- **Primary**
  - **Education** (refer to the appendix for more pre-processing information)
    - Childcare
      - Input dataset: education data (Maa-amet POIs, EHIS - Eesti Hariduse Infosüsteem)
      - Column in the input dataset: "Haridusteenuse kategooria"
      - Values of the column that map: 'Esmane teenus_Lastehoid'
    - Kingergarten (municipal, private)
      - Input dataset: education data (Maa-amet POIs, EHIS - Eesti Hariduse Infosüsteem)
      - Column in the input dataset: "Haridusteenuse kategooria"
      - Values of the column that map: 'Esmane teenus_Lasteaed_eraomand', 'Esmane'teenus_Lasteaed_munitsipaal'
    - Primary (municipal, state, private)
      - Input dataset: education data (Maa-amet POIs, EHIS - Eesti Hariduse Infosüsteem)
      - Column in the input dataset: "Haridusteenuse kategooria"
      - Values of the column that map: 'Esmane teenus_Põhikool_eraomand', 'Esmane teenus_Põhikool_munitsipaal', 'Esmane teenus_Põhikool_riigiomandus'
  - **Mobility**
    - Bus/tram stop
      - Input dataset: service POIs (Maa-amet)
      - Column in the input dataset: "alamgrupp"
      - Values of the category that map: 'bussipeatus', 'määratlemata', 'trammipeatus', 'bussijaam', 'trollipeatus'
    - Train stop
      - Input dataset: service POIs (Maa-amet)
      - Column in the input dataset: "alamgrupp"
      - Values of the column that map: 'rongipeatus'
  - **Nutrition**
    - Grocery
      - Input dataset: supplementary grocery data (OpenStreetMap)
      - Column in the input dataset: "Type"

- Values of the column that map: 'supermarket', 'greengrocer', 'grocery', 'convenience'

- **Taking care**
  - Pharmacy
    - Input dataset: service POIs (Maa-amet)
    - Column in the input dataset: "alamgrupp"
    - Values of the column that map: 'apteek'
  - Other
    - Input dataset: service POIs (Maa-amet)
    - Column in the input dataset: "alamgrupp"
    - Values of the column that map: 'regionaalhaigla', 'keskhaigla', 'muu haigla, määratlemata', 'perearstikeskus', 'perearsti vastuvõtt', 'kiirabibaas'
- **Public green areas**
  - *See the subsequent subsection, "Processing irregular destinations (public green areas, bike network, hiking network)"*
- **Bike network**
  - *See the subsequent subsection, "Processing irregular destinations (public green areas, bike network, hiking network)"*

- **Secondary**
  - **Education** (refer to the appendix for pre-processing information)
    - Gymnasium (municipal, state, private)
      - Input dataset: education data (Maa-amet POIs, EHIS - Eesti Hariduse Infosüsteem)
      - Column in the input dataset: "Haridusteenuse kategooria"
      - Values of the column that map: 'Teisejärguline teenus_Gümnaasium_eraomand', 'Teisejärguline teenus_Gümnaasium_munitsipaal', 'Teisejärguline teenus_Gümnaasium_riigi'omandus'
    - Recreational sports school (municipal, private)
      - Input dataset: education data (Maa-amet POIs, EHIS - Eesti Hariduse Infosüsteem)
      - Column in the input dataset: "Haridusteenuse kategooria"
      - Values of the column that map: 'Teisejärguline teenus_Huvikool või spordikool_eraomand', 'Teisejärguline teenus_Huvikool või spordikool_munitsipaal'
    - Vocational school (municipal, state, private)

- Input dataset: education data (Maa-amet POIs, EHIS - Eesti Hariduse Infosüsteem)
- Column in the input dataset: “Haridusteenuse kategooria”
- Values of the column that map: ‘Teisejärguline teenus_Kutsekool_eraomand’, ‘Teisejärguline teenus_Kutsekool_munitsipaal’, ‘Teisejärguline teenus_Kutsekool_riigiomandus’

- University (state, public, private)
  - Input dataset: education data (Maa-amet POIs, EHIS - Eesti Hariduse Infosüsteem)
  - Column in the input dataset: “Haridusteenuse kategooria”
  - Values of the column that map: ‘Teisejärguline teenus_Kõrgkool_avalik-õiguslik’, ‘Teisejärguline teenus_Kõrgkool_eraomand’, ‘Teisejärguline teenus_Kõrgkool_riigiomandus’

- **Civic**
  - Input dataset: service POIs (Maa-amet)
  - Column in the input dataset: “alamgrupp”
  - Values of the column that map: 'riigiasutus', 'häirekeskus', 'avalik varjumiskoht', 'piirkonnapolitseinik', 'reisisadam', 'politseiamet' 'prefektuur', 'päästekomando', 'väikesadam', 'sadam', 'mitteametlik sadam', 'riiklik lennujaam', 'rongijaam', 'politseiteenindus', 'sõjaväelennuväli', 'riiklik lennuväli'
- **Income**
  - Input dataset: service POIs (Maa-amet)
  - Column in the input dataset: “alamgrupp”
  - Values of the column that map: 'pangakontor', 'ATM sisse välja', 'ATM välja', 'pangapunkt'
- **Leisure**
  - Input dataset: service POIs (Maa-amet)
  - Column in the input dataset: “alamgrupp”
  - Values of the column that map: 'välitiir,' 'välisujula', 'tehnikaspordi püsirada, v.a veemotorada', 'ööklubi', 'baar', 'velodroom', 'noortekeskus', 'ratsaspordi rajatis', 'kultuuri- ja huvikeskus', 'mäesuusa- või lumelauasõidu püsirajatis', 'sportliku liikumise püsirada', 'vabaõhulava', 'mõis'
- **Personal care**
  - Input dataset: service POIs (Maa-amet)
  - Column in the input dataset: “alamgrupp”
  - Values of the column that map: 'palvepink', 'kirik'

    - **Social**
        - Input dataset: service POIs (Maa-amet)
        - Column in the input dataset: “alamgrupp”
        - Values of the column that map: 'üldhooldusteenus', 'erihoolekandeteenus', 'asendushooldusteenus', 'rehabilitatsiooniteenus'
    - **Trade**
        - Input dataset: service POIs (Maa-amet)
        - Column in the input dataset: “alamgrupp”
        - Values of the column that map: 'toidukäitlemine', 'tervisepood'
    - **Hiking network**
        - *See the subsequent subsection, “Processing irregular destinations (public green areas, bike network, hiking network)”*
- **Misc.**
    - **Building uses** (EHR)

Primary schools, kindergartens	Hospitals, doctors, pharmacies	Bus, tram, train stops	Grocery stores, supermarkets…	Office buildings, banks…	Parks, forests	Core & health network of bike paths
Education	Taking care	Mobility	Nutrition	Income	Green areas	Bike network
P R I M A R Y DESTINATIONS (POIs)						
S E C O N D A R Y DESTINATIONS (POIs)						
Education	Civic	Leasure	Trade	Social	Personal care	Hiking network
Universities, Vocational schools	Post offices, services, parking	Bars, Restaurants, Cultural institutions	Food based commerce, health produces	Social services	Places of prayer	Hiking paths

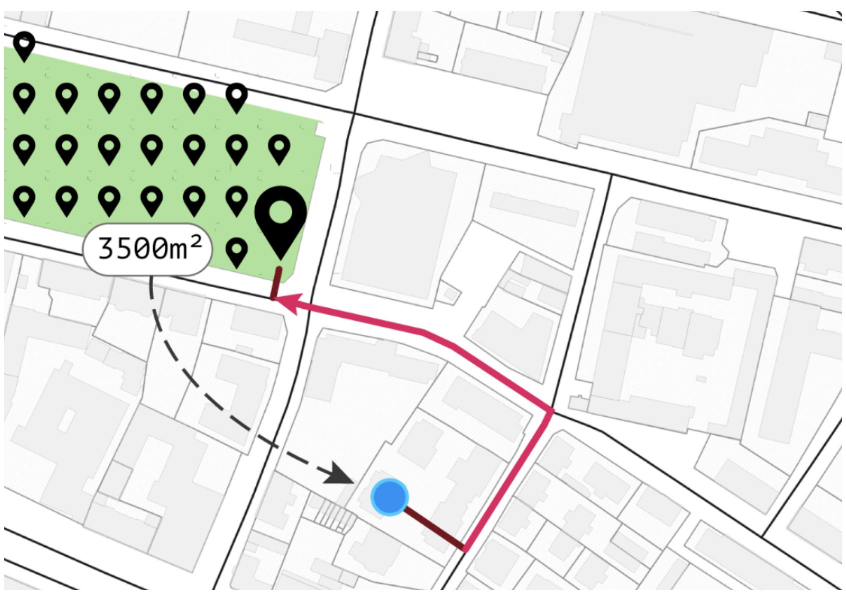
Categorisation scheme, simplified.

Processing irregular destinations (public green areas, bike network, hiking network):

The processing of irregular destinations such as public green areas, the cycling network, and hiking trails is a specialised part of this methodology. These destinations require unique processing techniques due to their non-point geometries.

The destinations that required a unique approach to processing were:

- **Public green areas:** the public green area data (TAR, Tallinn Strategic Management Office) is processed by creating a dense grid of points (50m) within each green area's polygon. This method ensures that the spatial characteristics of these areas are accurately captured, approximating their true size and shape. The public green area data (TAR, Tallinn Strategic Management Office) is gridded at 50m intervals, and the centroids of those grid cells become points. *Regarding their use in the analysis: the gridded points representing the green area are snapped to the network, like any other destination point in the analysis. Importantly, if multiple points belonging to a single green area are reached from an origin, that green area is only counted once.*
- **The bike and hiking networks:** The bike and hiking networks (Tallinn Strategic Management Office and Maa-amet, respectively) are processed by generating points at regular intervals (50m) along the paths. This approach provides a detailed and continuous representation of the networks. *Regarding their use in the analysis: the points representing each of these networks are snapped to the network, like any other destination point in the analysis; the networks are treated as one entity, so the maximum count from an origin to either of them is one.*



The image above describes how public green areas are treated in the analysis: even if the origin reaches multiple points representing the same green area, the green area's area is only passed to the origin once.

These methods and special considerations ensure that these non-point destinations can be treated as destinations and therefore effectively incorporated into the analysis.

Origin & parcel processing:

This process involves the detailed handling of building address points and cadastral parcels to ensure accurate representation of origins for accessibility analysis.

Origin processing:

I. **Filtering building address points:** the dataset of building address points is filtered to separate residential (marked as 'EE' in the ADS dataset) and non-residential (marked as 'ME' in the ADS dataset) buildings. Additionally, cadastral address points (CU), used when there is no ME or EE points in a parcel, are identified separately.
II. **Processing parcels:** the parcels dataset is divided into 'Transpordimaa' (transport land) and non-Transpordimaa parcels.
III. **Joining parcels with address points:** non-Transpordimaa parcels are then spatially joined with EE and ME address points to identify parcels containing address points of this type. Parcels without EE or ME address points are also identified for further processing.
IV. **Generating new address points:** for non-Transpordimaa parcels without EE or ME points, new address points are generated near the edges of the parcel, along an offset internal geometry, at an interval determined by the parcel's area. This step is essential for parcels that lack EE and ME address points but are still relevant for accessibility analysis.
V. **Buffering and joining with Transpordimaa parcels:** a buffer, greater than the offset distance of the internal geometries of non-Transpordimaa parcels without EE or ME points along which the points were created, is created around Transpordimaa parcels to capture the nearby newly generated points.
VI. **Combining address points:** the original EE and ME points are combined with the newly generated points to form a comprehensive set of address points. CU points, additionally, are used for parcels that, after this process, remain without any EE, ME, or custom-generated address points, which can happen for one of two reasons: because the sizes of the parcel's edges are smaller than the minimum interval or because the parcel does not border a Transpordimaa parcel.
VII. **Saving relevant parcels:** the parcels that these address points are within – effectively all of the parcels in the input data, excluding Transpordimaa parcels without EE or ME points – are saved.

The outcome of this process is two-fold: a detailed dataset of address points representing origins for the accessibility analyses, and a refined parcel dataset that the results of the analyses will be aggregated to.

Network processing:

This involves downloading and preparing OSM network data, which is used to analyze connectivity and accessibility within the study area.

Network data acquisition:

I. **Downloading OSM data:** the network data for pedestrian and cyclist analyses is sourced from OpenStreetMap, and was downloaded on 28/01/2024. This data is comprehensive and includes various types of pathways and roads suitable for walking and cycling.

II. **Custom filter application:** a crucial aspect of this process is the application of a custom filter to the OSM data. This filter is designed to exclude certain types of roads and paths that are not relevant or suitable for pedestrian and cyclist use. Specifically, the filter excludes:

    A. Roads marked as 'abandoned', 'bus_guideway', 'construction', 'motor', 'no', 'planned', 'platform', 'proposed', 'raceway', 'razed', 'primary', 'secondary', 'trunk', 'primary_link', 'secondary_link', 'trunk_link'.

III. **Combining the network data:** the network data for different land masses (Harju Maakond and Aegna) is downloaded separately, taking the largest weakly connected component for each and then combining them into a single dataset. This ensures comprehensive and connected coverage of the full study area.

IV. **Network data processing:** the filtered and combined OSM network data is then processed to create a spatial network. This network is composed of nodes (representing intersections, endpoints, etc.) and edges (representing the pathways, roads, etc. between nodes). This network is saved for further processing within the analysis before calculating the distance matrix between origins and destinations.

For the joint pedestrian-public transportation analysis as well as the motorist analysis, network construction is handled by r5py (henceforth referred to as R5), a Python wrapper of the Conveyal-developed R5 routing engine, licenesed under the MIT License. OSM network data in the Protocol Binary Format from 26/01/2024, downloaded from Geofabrik, serves as input for these analyses.

Processed data

The input data is transformed using the methodologies described in the "Data Processing" section into processed data, the structure of which is described below. This data is used as input for the analysis, described in the subsequent section, "Analysis Methodology".

Processed destination points table structure:

Count: 182,780

source_dataset	The dataset from which the destination point originates.
source_dataset_service_classification_column	The column that the mapping to service_subcategory is based on, described in section III, under the "Destination processing" section.
source_dataset_service_classification_value	The value of the point's service classification in the original dataset.
source_dataset_service_classification_value_en	The value in english, for reference.
source_name_column	The source dataset's name column - if there are multiple, one is selected.
source_name	The value of the point's name column in the original dataset.
source_id_column	The source dataset's id column - if there are multiple, one is selected.
source_id	The value of the point's id column in the original dataset.
unique_id	A unique string combining the source_dataset, the source_name, and the source_id values.
primary_secondary	Whether the point is categorised as primary or secondary, as described under the "Destination processing" section.
service_subcategory	The point's service subcategory, as described under the "Destination processing" section.
SUM_[**EHR building use typ**e]	If the point originates from the building uses / EHR dataset, the amount of area, in sqm., of each building use type within the building.
suurus_m2	If the point originates from the public green areas dataset, the amount of area, in sqm., of the parent green area, as described under the "Processing irregular destinations (public green areas, bike network, hiking network)"section.
processing_id	A unique processing_id, used for building the distance / travel-time matrices.

Filtered address/origin points table structure:

Count: 61,947

ADOB_ID	The address point's original ADOB_ID value.
ADOB_LIIK	The address point's original ADOB_LIIK value. In addition to potentially taking on 'ME', 'EE', and 'CU' values, which exist in the input dataset, ADOB_LIIK is set to 'CUSTOM-EDGE' for points generated along the Transpordimaa-adjacent edges of parcels without 'ME' or 'EE' points, as described under the "Origin & parcel processing" section.
ADS_OID	The address point's original ADS_OID value.
ADR_ID	The address point's original ADR_ID value.
TUNNUS	The TUNNUS value of the parcel that contains the address / origin point.
processing_id	A unique processing_id, used for building the distance / travel-time matrices.

Filtered parcels:

Count: 35,156

The parcels retain all of their original columns but, as described under the "Origin & parcel processing" section, are limited to Transpordimaa parcels that contain 'ME' and 'EE' address points and all of the non-Transpordimaa parcels.

Networks:

The network used for the pedestrian and cyclist analyses was downloaded on 28/01/2024 and is, as implied in the *"Network processing"* section, built using the following query to the OpenStreetMap Overpass API:

Unset

```
'[\"highway\"][\"area\"!~\"yes\"]'
'[\"highway\"!~\"abandoned|bus_guideway|construction|motor|no|planned|platform'
 'proposed|raceway|razed|primary|secondary|trunk|'
 'primary_link|secondary_link|trunk_link\"]'
```

The pedestrian and motorist networks used for both the pedestrian + public transportation analysis and the motorist analysis, respectively, are embedded in OpenStreetMap data in the Protocol Buffer Binary Format downloaded on 26/01/2024, covering all of Estonia. These networks are parsed according to Conveyal's specification for each network type: 'WALK',' CAR'.

GTFS feeds:

A GTFS feed from 16/11/2022, downloaded from TransitFeeds, is used for the analysis. *Note that, on request, an earlier date for the GTFS feed, 16/11/2022, was selected due to the relative normalcy of the public transportation service level during that month and the comparative lack of service, due to road work in 2023, since.*

Analysis methodology

Pedestrian and cyclist distance matrix:

Network construction:

I. **Initialisation of the network graph:** the network data, described previously as being derived from OpenStreetMap and processed to include only relevant paths for pedestrians and cyclists, is used to construct a network graph: a primal, undirected multigraph.

II. **Interpolation of network edges:** the network graph's edges are interpolated at regular intervals (25m) while preserving their attributes, reconstructing the network and creating more nodes to which the origins and destinations will later snap.

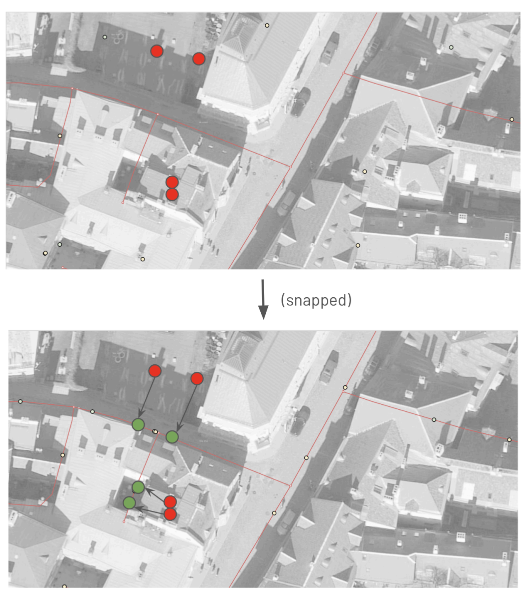
III. **Creation of the network graph**: an operable NetworkX graph is created from the interpolated network edges. This graph forms the basis for subsequent distance calculations.

Snapping of address and destination points:

I. **Nearest node identification:** for each address and destination point, the nearest node on the network graph is identified. This process involves finding the closest network node within a specified snapping search radius (400m) – origins and destinations for which no nodes are within a 400m radius from their original geometry are dropped.

II. **Snapped points generation:** the original geometries of the address and destination points are replaced with the geometries of their nearest nodes.

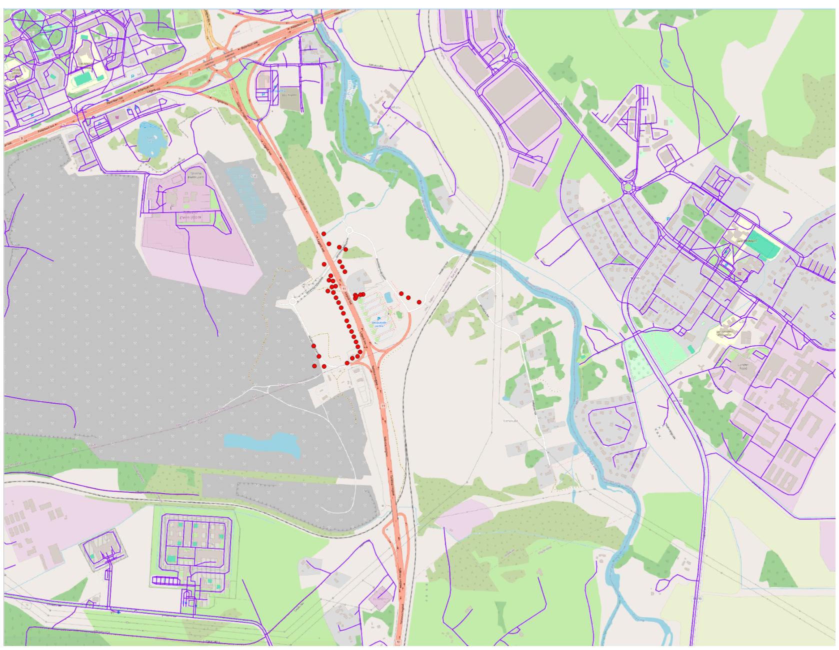
III. **Data prepreration for distance calculation:** the snapped points are prepared with necessary attributes (like node IDs) for the distance matrix calculation.



How points are snapped – to the closest node.

In the case of the pedestrian/cyclist network, 61903 of 61947 origins / address points are snapped to the network. The other 44 points are not snapped to the network due to the network being beyond the threshold of 400m.

The points that are not snapped are distributed like so, and excluded from the analysis.



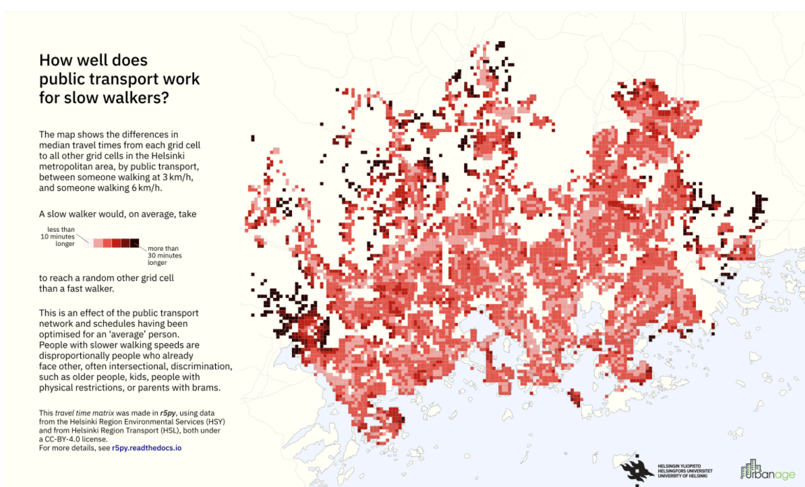
None of the points within the cluster of 43 points at the southeast extent of the points, in Veneküla, reach the network within a 400m threshold.

Distance matrix calculation:

- **Subset division for parallel processing:** the snapped origin points are divided into subsets to facilitate parallel processing.
- **Distance calculation:** using the *igraph library*, the shortest path distances between each origin and nearby destinations within a maximum radius of 5000m from them are calculated using Dijkstra's algorithm. 5000m is selected as the radius due to it being the largest distance of interest for the pedestrian and cyclist analyses.
- **Distance matrix assembly:** the distances calculated for each subset are combined to form a comprehensive distance matrix. This matrix represents the distances between all origin and destination pairs in the study area.
- **Debugging and validation:** for debugging purposes, the process includes the creation of files that allow for the inspection of the network graph, snapped points, and the distance matrix.
- **Output generation:** the final distance matrix is saved in long-format in the CSV file format, providing a detailed representation of the connectivity between origins and destinations within the network and used, in later steps, for accessibility counts and aggregation to parcels.

Use of R5 for the pedestrian + public transportation and motorist travel-time matrix calculations:

A Python wrapper for the Conveyal-developed R5 routing engine, r5py, licensed under the MIT License, is used to compute the pedestrian + public transportation and motorist travel-time matrices. Its usage is described in the following two sub-sections.



Fink, C., Klumpenhouwer, W., Saraiva, M., Pereira, R., & Tenkanen, H., 2022: DOI:10.5281/zenodo.7060437

Pedestrian + public transportation distance matrix:

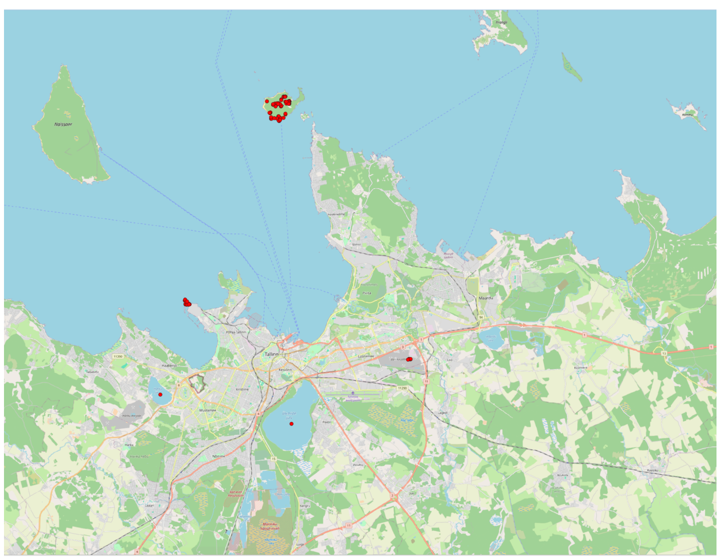
Network construction:

- **Transport network setup:** R5 constructs a transport network combining OSM network data in the Protocolbuffer Binary Format, parsed per Conveyal's specification for 'WALK', from 26/01/2024 and a GTFS feed from 16/11/2022.
- **Network building:** the network is built for a specific date (16/11/2022) and time (a 30-minute window from 9:00), capturing the state of the pedestrian network and public transportation services.

Snapping of address and destination points:

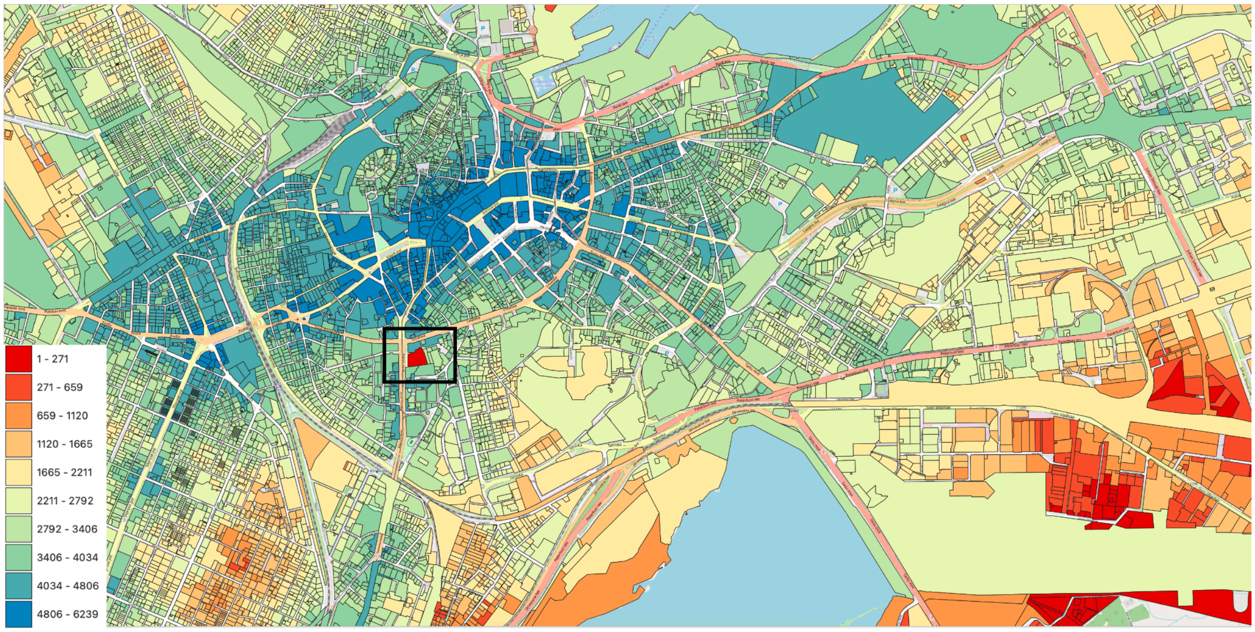
- **Snapping process:** origin, destination, and public transportation stop points are snapped to Conveyal's specification for a 'WALK' network using a specified search radius of 400m, thereby ensuring accurate placement on the pedestrian network for travel-time calculations – note that, like for the other modes, origins and destinations for which no nodes are within a 400m radius from their original geometry are dropped from the analysis.
- **Data preparation:** The snapped points are prepared with the necessary attributes for the calculation of the travel-time matrix, ensuring correct identification and processing.

In the case of the pedestrian + public transportation analysis, 61853 of 61947 origins / address points are snapped to Conveyal's parsing of the 'WALK' network, which is structurally different to the pedestrian and cyclist network defined in the previous sub-section. The other 94 points are not snapped to the network due to the network being beyond the threshold of 400m from each.



The points that are not snapped are distributed like so, and are excluded from the analysis.

Due to the nature of the OSM Protocolbuffer Binary Format data downloaded on 26/01/2024 and Conveyal's definition of the 'WALK' network, certain islands exist. This, for example, affects TUNNUS:78401:101:4967.

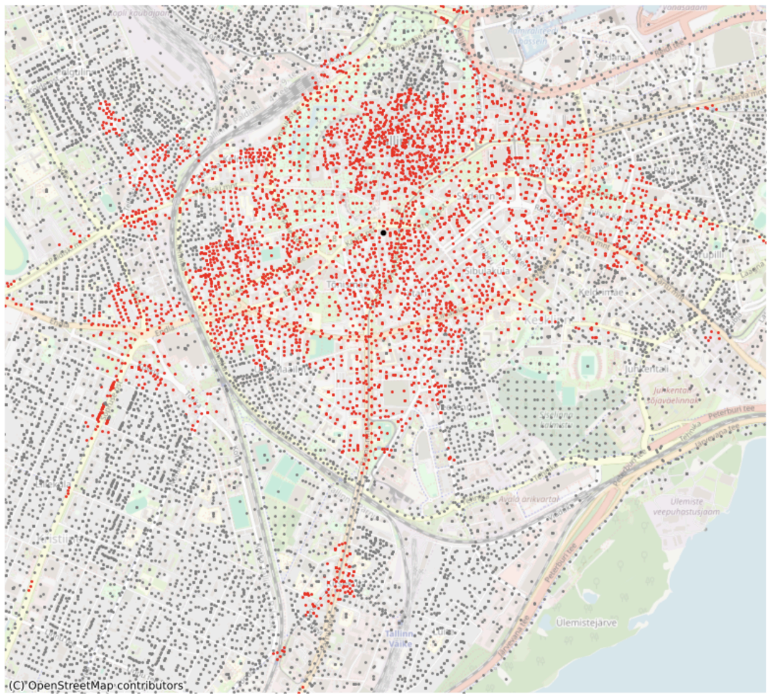


Map showing the total count of services, excluding building uses (EHR), within a 30m travel-time.
TUNNUS:78401:101:4967 highlighted.

Travel-time matrix calculation:

- **Travel-time computation:** R5's TravelTimeMatrix computer is used to compute the travel-time matrix, calculating travel-times between each pair of origin and destination points along the 'WALK' and public transportation networks.
- **Parameter setting:** parameters such as transport modes (walking and public transportation), walking speed, departure time, and time windows are set to reflect realistic pedestrian and public transportation conditions.
    - Walking speed: the average walking speed is set to 3.6 km/h. In the case of the analysis, this affects the pedestrian coverage in accessing and egressing public transportation, from and to origins and destinations.
    - Transport modes: the analysis is limited to specific transport modes, specifically walking and public transportation ('WALK', 'TRANSIT').
    - Access and egress modes: the analysis is limited to walking for both access to and egress from public transportation stops, meaning that individuals, in the analysis, can only access and depart from public transportation points via walking.
    - Departure time and departure time window: the departure time is set for 9 AM on 16/11/2022 with a 30-minute departure time window, meaning that, in the analysis, departures are searched for every minute between 9 AM and 9:30.
    - Maximum travel-times: A maximum 30-minute limit is set for the maxiximum travel-time across access, transport, and egress, to avoid any unnecessary computations given that a travel-time of 30 minutes is the maximum time that the study is interested in.

- **Output generation:** as is the case for other modes, the final distance matrix is saved in long-format in the CSV file format, providing a detailed representation of the connectivity between origins and destinations along the pedestrian and public transportation networks and used, in later steps, for accessibility counts and aggregation to parcels. R5 is set such that the median travel-time between each origin and destination is written to the travel-time matrix.



Points accessible within a 15-minute travel-time from an origin using a combination of walking and public transportation.

Motorist distance matrix:

Network construction:

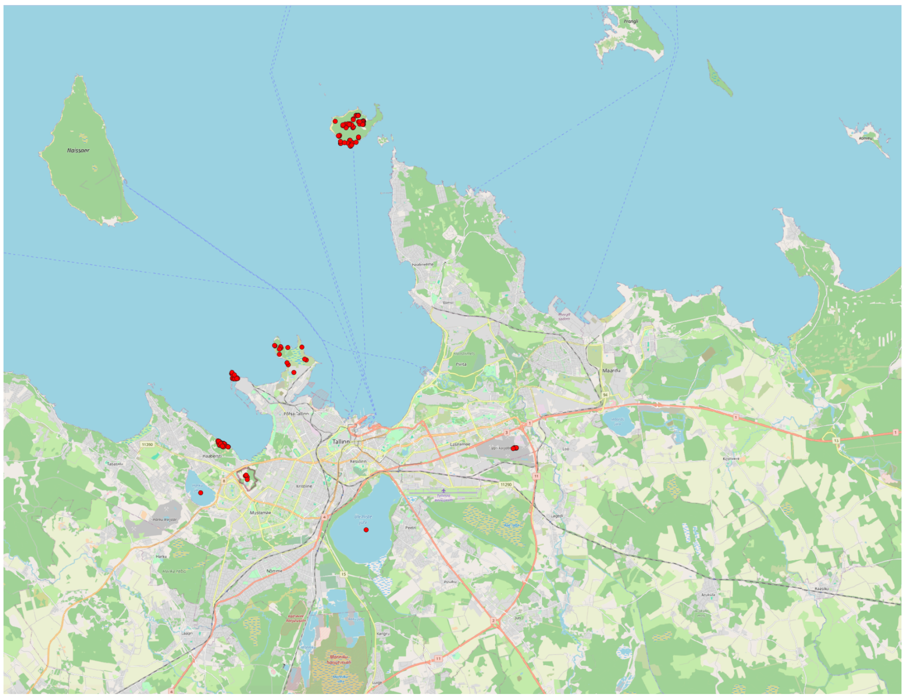
- **Transport network setup:** R5 constructs a transport network combining OSM network data in the Protocolbuffer Binary Format, parsed per Conveyal's specification for 'CAR', from 26/01/2024.
- **Network building:** the network is built for a specific date (16/11/2022) in a simplified way – congestion not factored in, for example.

Snapping of address and destination points:

- **Snapping process:** origin, destination, and public transportation stop points are snapped to Conveyal's specification for a 'CAR' network using a specified search radius of 400m, thereby ensuring accurate placement on the motorist network for travel-time calculations – note that, like for the other modes, origins and destinations for which no nodes are within a 400m radius from their original geometry are dropped from the analysis.

- **Data preparation:** The snapped points are prepared with the necessary attributes for the calculation of the travel-time matrix, ensuring correct identification and processing.

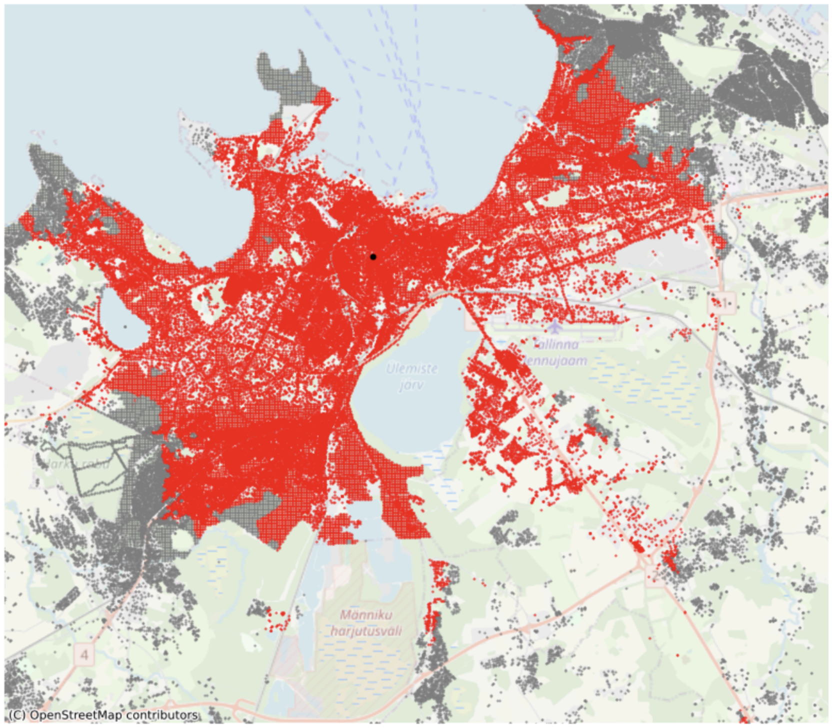
In the case of the motorist analysis, 61802 of 61947 origins / address points are snapped to Conveyal's parsing of the CAR network. The other 145 origin points are not snapped to the network due to the network being beyond the threshold of 400m from each.



The points that are not snapped are distributed like so, and are excluded from the analysis.

Travel-time matrix calculation:

- **Travel-time computation:** R5's TravelTimeMatrix computer is used to compute the travel-time matrix, calculating travel-times between each pair of origin and destination points along the 'CAR' network.
- **Parameter setting:** the transport mode is set to 'CAR' to reflect realistic motorist conditions.
- **Output generation:** as is the case for other modes, the final distance matrix is saved in long-format in the CSV file format, providing a detailed representation of the connectivity between origins and destinations along the motorist network and used, in later steps, for accessibility counts and aggregation to parcels. R5 is set such that the median travel-time between each origin and destination is written to the travel-time matrix.



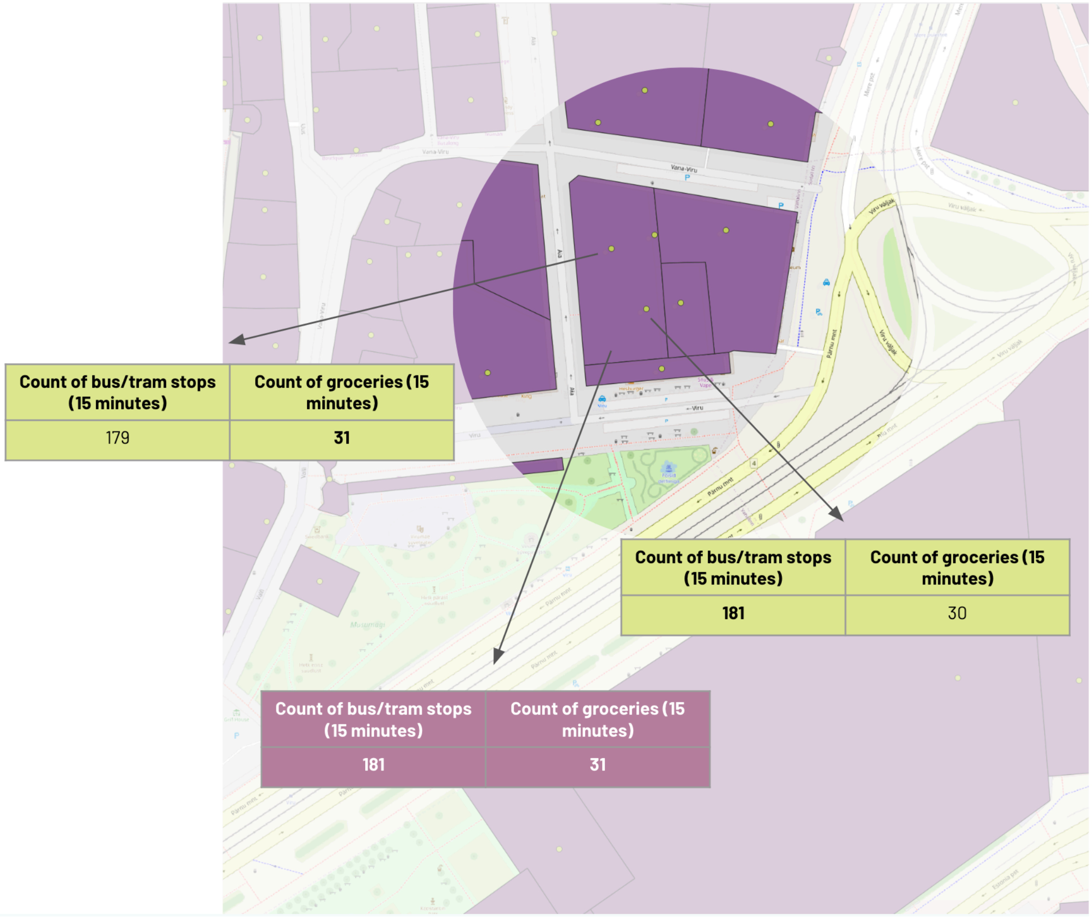
Points accessible within a 15-minute travel-time from an origin using a car.

Counts:

The analysis, utilising the calculated distance or travel-time matrix, involves calculating the count of accessible services and the total sum of public green areas and EHR building uses within specified distance or travel-time thresholds for each origin point. The counts reflects the number of reachable destinations along the network from each snapped, routable origin, such as different kinds of educational institutions, healthcare facilities, and public transportation stops, within a certain walking or transit time from each origin, and the sums represent the amount of, for example, public green space available within certain walking or transportation times from each origin. In short, the counts and sums are categorized based on service subcategories and distance / travel-time thresholds, providing a detailed view of accessibility.

Aggregation:

The accessibility data (counts and sums), initially calculated for individual origin points, is aggregated to the level of building parcels. This aggregation process involves spatially joining the accessibility results with building parcel geometries and summarizing the accessibility metrics at the parcel level, taking only the maximum value of each column. The result is a comprehensive parcel-level dataset that reflects the accessibility of various services and amenities for each building parcel, considering the cumulative accessibility from all origins within that parcel.



A diagram showing how the maximum value is aggregated to the parcel: for each count / sum individually.

Indices:

In addition to simple counts and sums, the analysis computes indices at the parcel-level that provide a more nuanced understanding of accessibility. These indices include:

- **Diversity / Shannon-Wiener index:** This index is used to measure the diversity of accessible services by service category and the diversity of building-use area by building-use category (EHR: Elamud, Büroo, Transport, and so on) within each distance threshold. It provides insight into the variety of services and amenities accessible from each origin or parcel, beyond just the quantity.
- **Weighted diversity / Shannon-Wiener index:** A weighted version of the aforementioned diversity / Shannon-Wiener index is also included. The index is weighted by the ratio between the parcel's access to services / usable building area and the maximum access to services / usable building area across all of the parcels.
- **Threshold-based indices:** for more specific analysis needs in line with strategic targets, custom accessibility metrics are calculated based on certain criteria, based on the presence of certain service categories within specific distances from the origin along the network. A cumulative and binary metric are built based on these criteria: the former getting at the number of criteria that are met, and the latter getting at whether all of the criteria are met. The criteria and two indices are described in the table and image below:

Threshold-based indices criteria	
Primary education place (municipal, state,private)	1200m
Kindergarten (municipal, private)	1200m
Train stop	1000m
Bus / tram stop	400m
Grocery store	800m
Pharmacy	800m
Green area	300m
Bike network	500m

CONDITIONS NOT MET ON PARCEL 7/8 conditions met (TUNNUS:78408:801:1580)	
Primary education place (municipal, state,private) within 1200m	2
Kindergarten (municipal, private) within 1200m	2
Train stop within 1000m	1
Bus / tram stop within 400m	1
Grocery within 800m	1
Pharmacy within 800m	1
Green area within 300m	0
Health network within 500m	1

CONDITIONS MET ON PARCEL 8/8 conditions met (TUNNUS: 78408:801:2560)	
Primary education place (municipal, state,private) within 1200m	2
Kindergarten (municipal, private) within 1200m	3
Train stop within 1000m	2
Bus / tram stop within 400m	2
Grocery within 800m	3
Pharmacy within 800m	3
Green area within 300m	1
Health network within 500m	1

Output guide

Parcel-level column guide:

TUNNUS	Object ID of the parcel.
SIHT1	Use of the land #1.
SIHT2	Use of the land #2.
SIHT3	Use of the land #3.
SO_PRTS1	Percentage of the parcel’s land area that is dedicated to SIHT1.
SO_PRTS2	Percentage of the parcel’s land area that is dedicated to SIHT2.
SO_PRTS3	Percentage of the parcel’s land area that is dedicated to SIHT3.
COUNT_{**SERVICE SUBCATEGORY**}_{**DISTANCE/TIME**}m	The number of {**SERVICE SUBCATEGORY**} POIs within {**DISTANCE/TIME**}m from the building / buildings within the parcel (maximum value)
TOTAL_COUNT_SERVICE_SUBCATEGORIES_EXCL_BLDG_USES_{**DISTANCE/TIME**}m	The total number of services, across all counted categories except for EHR buildings (not considered a service), within {**DISTANCE/TIME**}m from the buildings within the parcel (maximum value)
SUM_{**BUILDING USE TYPE**}_EHR_m2_{**DISTANCE/TIME**}m	The total number of square metres of {**BUILDING USE TYPE**} within {**DISTANCE/TIME**}m from the building / buildings within the parcel (maximum value)
TOTAL_EHR_TYPE_SQM_SUM_{**DISTANCE/TIME**}m	The total number of square meters, across all building use types, within {**DISTANCE/TIME**}m from the buildings within the parcel (maximum value)
SUM_GREEN_AREA_suurus_m2_{**DISTANCE/TIME**}m	The total number of square meters of public green area accessible within {**DISTANCE/TIME**}m from the buildings within the parcel (maximum value)
SWI_SERVICE_SUBCATEGORIES_EXCL_BLDG_USES_{**DISTANCE/TIME**}m	Shannon-Wiener Index (SWI) measuring the diversity of service subcategories, except EHR buildings, within {**DISTANCE/TIME**}m from the buildings within the parcel.
SWI_SERVICE_SUBCATEGORIES_EXCL_BLDG_USES_WEIGHTED_{**DISTANCE/TIME**}m	Weighted Shannon-Wiener Index for service subcategories, except EHR buildings, within {**DISTANCE/TIME**}m. The index is weighted by the total

	count of services across each category relative to other parcels, giving more weight to parcels with more services.
SWI_EHR_TYPE_SQM_SUM_{**DISTANCE/TIME**}m	Shannon-Wiener Index measuring the diversity of square meters of different building use types within {**DISTANCE/TIME**}m from the buildings within the parcel. It reflects the variety in building usage types in terms of their total area.
SWI_EHR_TYPE_SQM_SUM_WEIGHTED_{**DISTANCE/TIME**}m	Weighted Shannon-Wiener Index for building use types within {**DISTANCE/TIME**}m. This index is weighted by total access to square meters of each building use type relative to other parcels, giving weight to those parcels with access to more square meters.
THRESHOLD_ACCESSIBILITY_METRIC_CUM	A cumulative metric assessing the presence of various essential services and amenities (like education, mobility, nutrition, health, and green areas) within specific distances. The score increases with the number of criteria met, offering a comprehensive measure of accessibility.
THRESHOLD_ACCESSIBILITY_METRIC_BIN	A binary metric that evaluates whether a parcel meets all specified accessibility criteria (covering various services and amenities). It is set to 1 if all criteria are met, indicating high accessibility, or 0 otherwise.