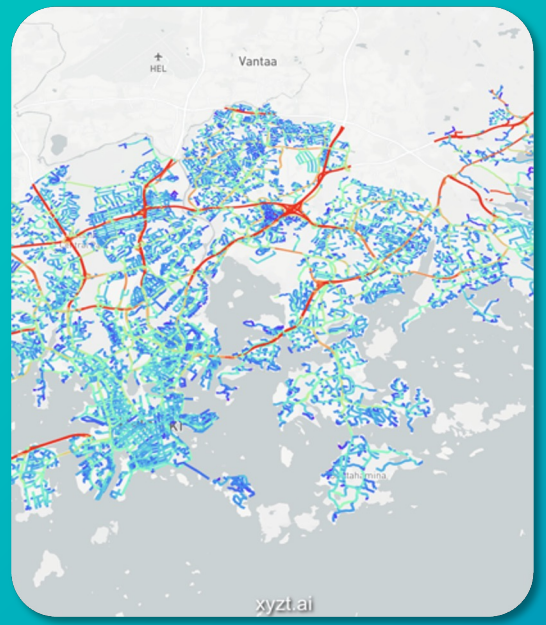
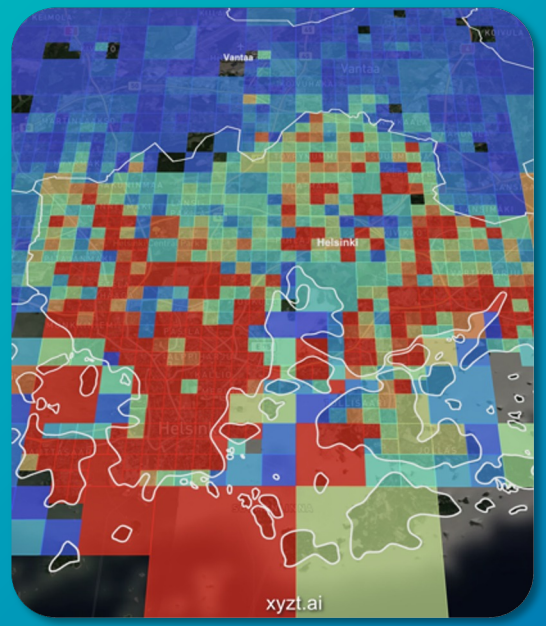
Data Driven Mobility Analysis

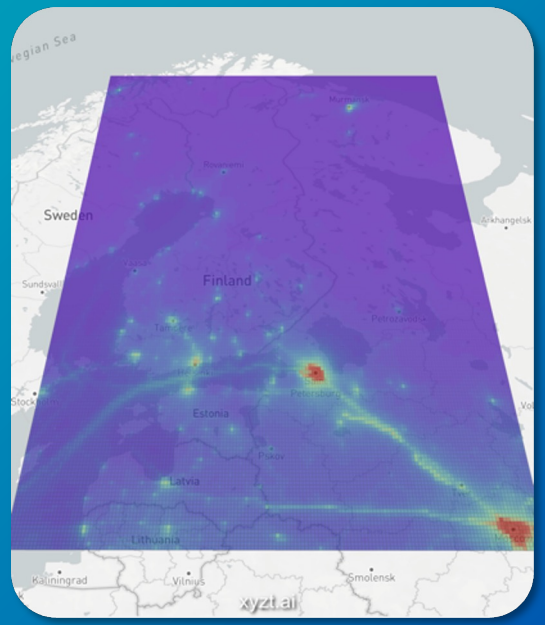
Use multi-source data and space-time visual analytics for smarter mobility insights



floating vehicle data



telecom data



emission data

xyzt.ai
location analytics at scale

GEO MOBILITY

Content

1. Goal(s) of the pilot
2. Description of the solution
3. Implementation
4. Results
5. Key findings
6. Next steps
7. Contacts

1. Goal(s) of the pilot

- Validation & comparison of different mobility data sources
- Communicating the potential of data for different use cases
- Keep data sets for future use

2. Description of the solution

- xyzt.ai platform
  - 10 user licenses
  - 500 Gb storage
  - 1 billion records per data set

- Multiple spatio-temporal mobility data sources
  - HERE (aggregate FVD)
  - TomTom (aggregate FVD)
  - ODIQ (Google, aggregate routing)
  - Traffic stats (street counters)
  - Telia (aggregate telco)
  - Copernicus (air quality)

- Self-service visual analysis, dashboarding, and reporting

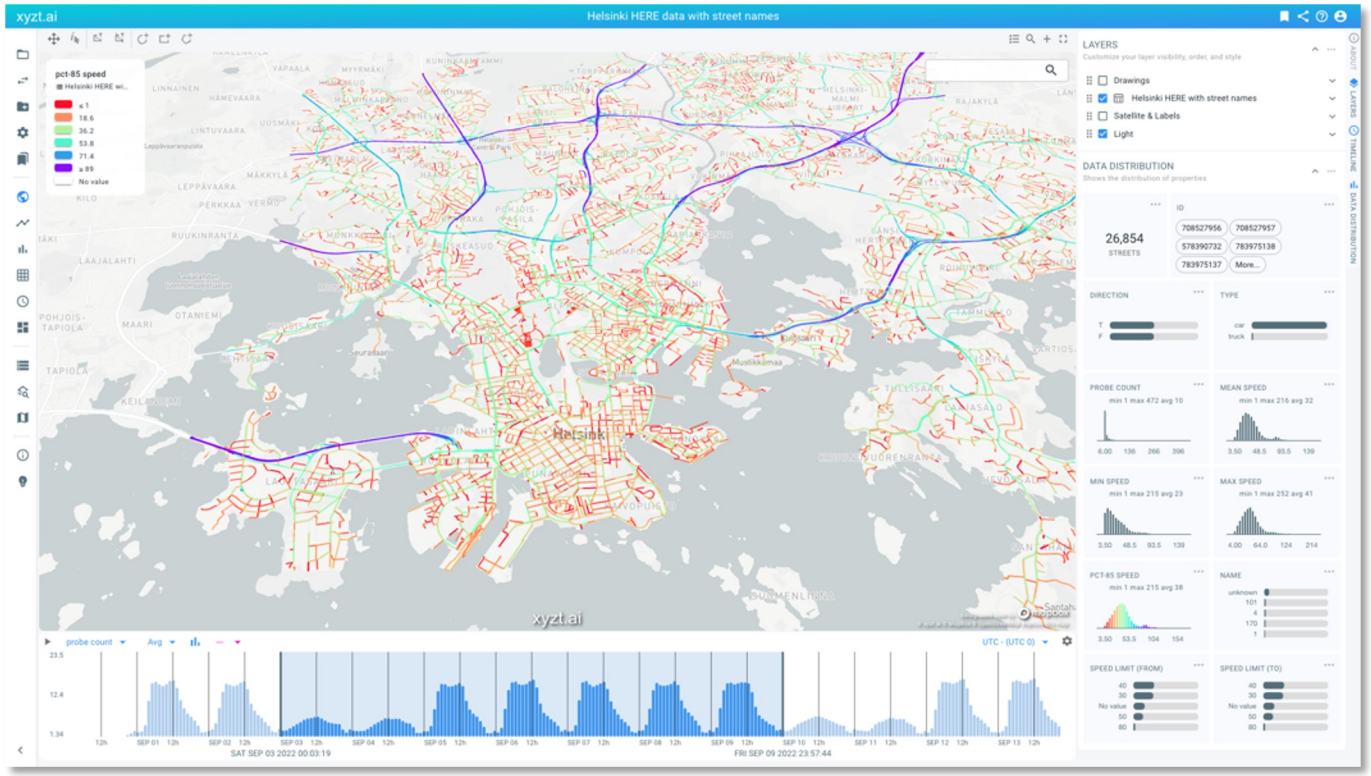


Figure: xyzt.ai visual analytics platform with HERE road data for September 2022

3. Implementation

- **Collect one month of mobility related data sources for the Helsinki region (September 2022)**
  - HERE:
    - hourly road statistics based on floating vehicle data, including sample count, mean and 85-percentile velocity,…
  - TomTom:
    - monthly summarized road statistics for specific periods during the day and week, including commute hours,…
    - Monthly summarized origin-destination data between different regions in Helsinki for specific periods during the day and week, including commute hours
  - Google:
    - hourly route delay statistics for main routes into Helsinki city center
  - Traffic counters:
    - (mostly) hourly stats by counting traffic and average speed at fixed locations in Helsinki. Data provided by Helsinki.
  - Telia:
    - hourly origin-destination people flow data derived from telecom data between different municipalities/regions with destination or origin Helsinki
  - Copernicus
    - hourly NO2/CO/… emission data measured by the European Copernicus satellite program
- **Make data available in xyzt.ai platform for visual analysis**
- **Train Helsinki mobility and traffic analysts on the data and the platform**
- **Self-serve analysis**
  - Visually analyze data coverage and quality
  - Identification of use-cases that can be solved with the data and the xyzt.ai platform

4. Results

- Data sets
- Visual analysis in the xyzt.ai platform
- Data coverage and quality overview
- Applicable use-cases overview
- Helsinki specific use cases

Data Sets

Overview of the different mobility data sets used in the project.

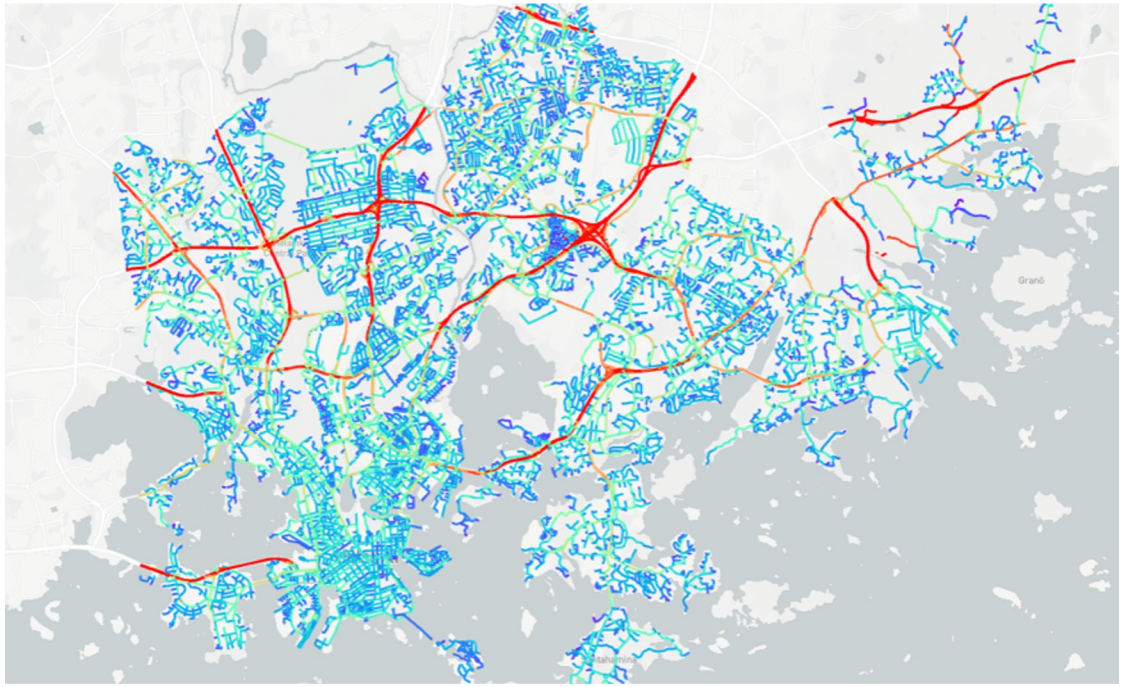
Data Sets Overview

Data Set	Summary	Origin
HERE road stats	For each road segment, contains information on number of vehicles, average speeds,… with hourly statistics.	Computed as aggregate floating vehicle probe data by HERE.
TomTom traffic stats	For each road segment, contains information on the number of vehicles, average speeds,… aggregated over a period of time (e.g., commute hours). There is no time series anymore in the resulting data set.	Downloaded from the TomTom platform after selecting a time period. Based on floating vehicle data.
TomTom origin destination	Trip counts and percentages between different regions. Statistics are aggregated data for periods selected by the user, e.g., morning/evening/entire day aggregated	Extracted from the TomTom platform after selecting an area and time period. Based on floating vehicle data.
ODIQ static link data	Fixed routes with information on how traffic flows (e.g., delays in seconds/km) along the routes during several days or weeks with an interval of between 1 and 15 minutes.	Obtained by querying the Google routing API over time for user selected origins and destinations.
Traffic counts	Contains (mostly) hourly statistics on traffic counts (number of vehicles) and average speeds at fixed locations.	Helsinki provided this information.
Telia origin destination	Contains movements (number of trips) between different areas in on grid cells or administrative boundaries.	Data is aggregate data obtained from cell phones on the Telia network.
Copernicus atmospheric monitoring data	Hourly measurements of different atmospheric particle concentrations such as NO2.	European Copernicus program, obtained through satellites.
Otonomo, Wejo, Bridgestone, INRIX,…	Raw floating vehicle data providers (GPS coordinates with associated velocities, accelerations, vehicle types, etc. of cars, trucks,…). This data was not used for the Finland region in the project, though there are sample projects with raw FVD data from INRIX that could be consulted.	From connected vehicles through OEMs (e.g., BMW, VW,…) or through fleet tracking (e.g., Webfleet technology).

Data Set: HERE

For each **road segment**, for each **direction**, for each **vehicle type** (cars and trucks), for each hour

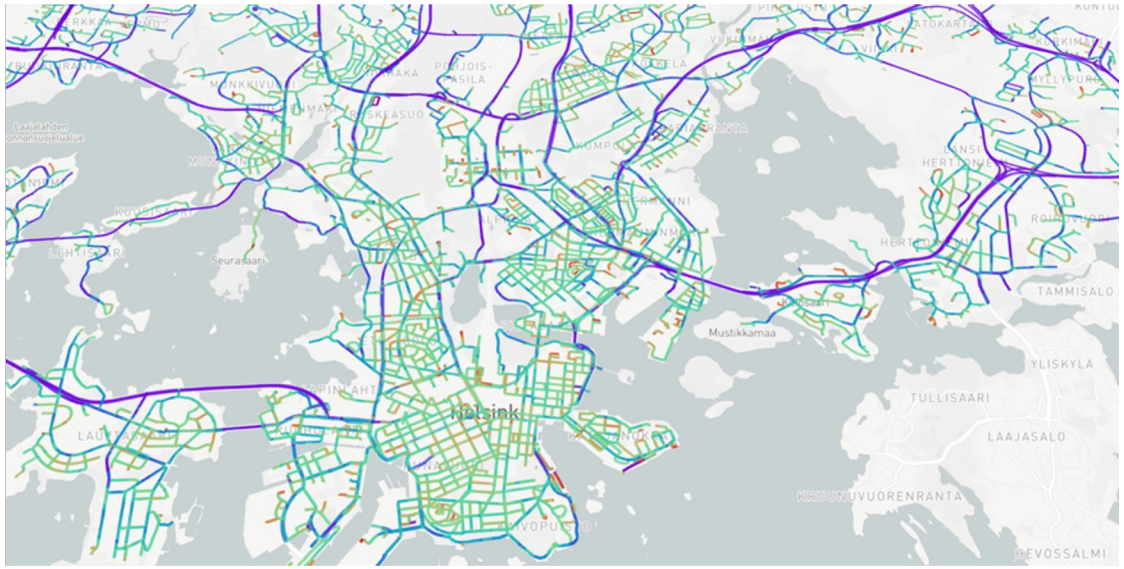
- **probe count**: the number of probes that are used to compute the statistic
- **mean speed**: the mean speed during that hour
- **stddev**: the standard deviation of the mean speed
- **min**: the minimum speed recording during that hour
- **max**: the maximum speed recording during that hour
- **confidence**
  - An indication of how good the captured data is. This takes into account the number of probes in the data.
- **pct-5** → **pct-95**: the speed percentiles.
  - The pct-85 velocity is the velocity that at least 85% of road users had during the registered hour.



Data Set: TomTom Traffic Stats

For a selected date range and time period (e.g., commute hours 7am-11am during September) provides aggregate floating vehicle data with properties:

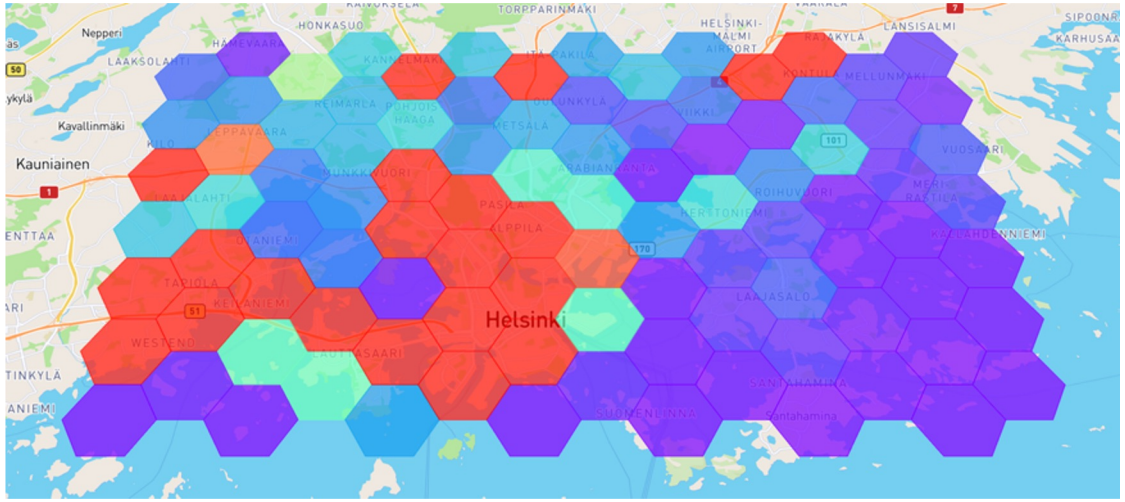
- The street segment **identifier**
- The street **name**
- The speed **limit**: the maximum driving speed allowed
- **Sample size**: the number of samples that contribute to the statistics
- **Distance**: the length of the specific segment
- **Median Speed**
- **Average Speed**
- **5-pct → 95-pct Speed**: the speed that at least 5%, 10%,… of the drivers drive. We only exposed the 50 percentile and 85 percentiles in the platform to reduce the number of attributes
- **Average Travel Time**: The average travel time to drive the segment, in seconds.
- **Median Travel Time**: The mean travel time to drive the segment, in seconds.
- **Travel Time Ratio**: The ratio of the travel time of this time window vs the travel time of the morning commute window.



Data Set: TomTom Origin Destination

Trip counts and percentages between different regions in Helsinki, in the data sets for morning/evening/entire day aggregated:

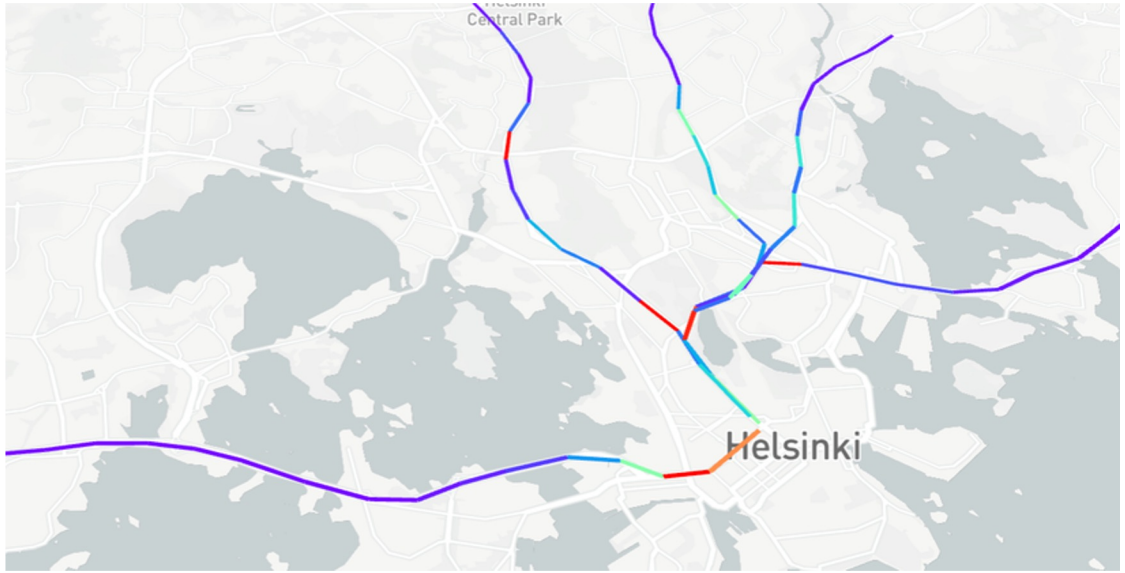
- The **main** region
- The **other** region
- **The type of flow**: incoming or outgoing. Incoming means that data is about trips incoming in the main region and originating from the other region. Outgoing means that the data is about trips leaving the main region and going to the other region.
- **Trips**: the number of trips flowing from main to other or from other to main depending on the above type.
- **Percent**: the percentage of trips
- **Trips morning**: same as above but for morning commute hours
- **Percent morning**
- **Trips evening**: for evening commute hours
- **Percent evening**



Data Set: ODIQ (Google) static link data

The ODIQ static link data contains fixed routes and has information on how traffic flows along this route during several days or weeks with an interval of between 1 and 15 minutes

- A route consisting of segments. In the xyzt.ai platform, this is ingested as a GeoJSON file.
- For each segment along the route, following **constant** properties are provided:
  - **route**: the route the segment belongs to (the route name is the filename that was used to import the data). This property can be used to filter out an individual route.
  - **distance**: length of the route segment (in meters)
  - **distanceInRoute**: how far along the route this segment is (in meters)
- For each segment along the route, following **temporal** properties are provided:
  - **delay**: the delay (in seconds per km) experienced on average at the corresponding time instance
  - **speed**: the average speed (in kmh) recorded at the corresponding time instance



Data Set: Traffic counting data

Contains (mostly) hourly statistics on traffic counts (number of vehicles) and average speeds at fixed locations throughout Helsinki.

Each records contains:

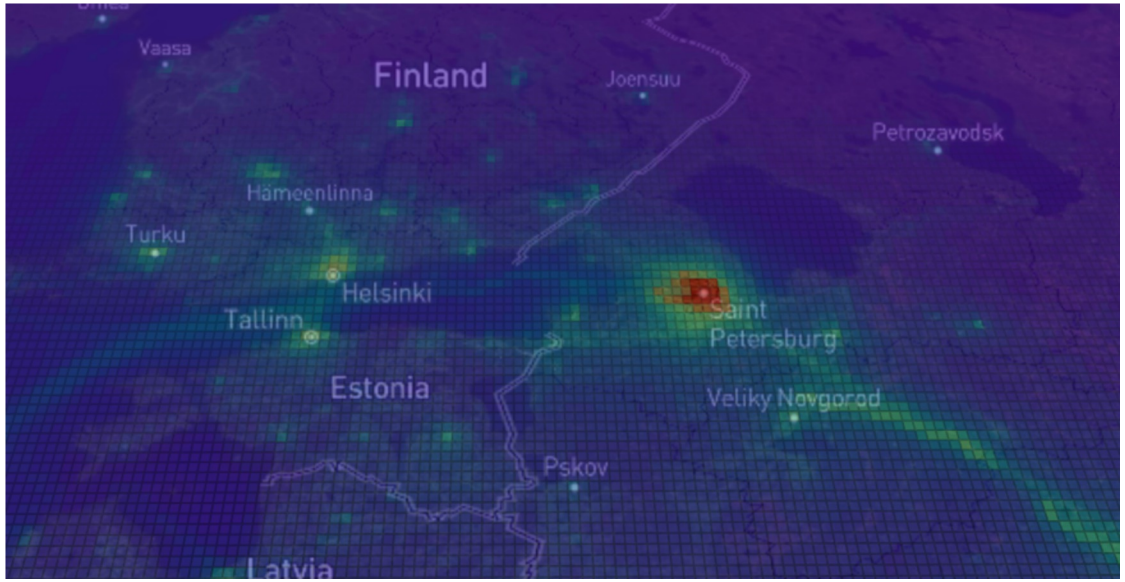
- An **identifier** linking to the counter location
- A **timestamp**
- **Traffic count** (number of vehicles per hour)
- **Average speed** (aggregate per hour)
- **Direction** (e.g., east or west, north or south)

Data Set: Copernicus Atmospheric Monitoring Data

The data consists of hourly measurements of NO2, CO, PM10 (PM = particle matter), and PM2.5 for September 2022.

- **NO2**: Mass concentration of nitrogen dioxide in the air (hourly values, g/m3)
- **CO**: Mass concentration of carbon monoxide in the air (hourly values, g/m3)
- **PM10**: Mass concentration of PM10 ambient aerosol in the air (hourly values, g/m3)
- **PM2p5**: Mass concentration of PM2.5 ambient aerosol in the air (hourly values, g/m3)

Note that CAMS provides many more variables, such as O3, NO,…



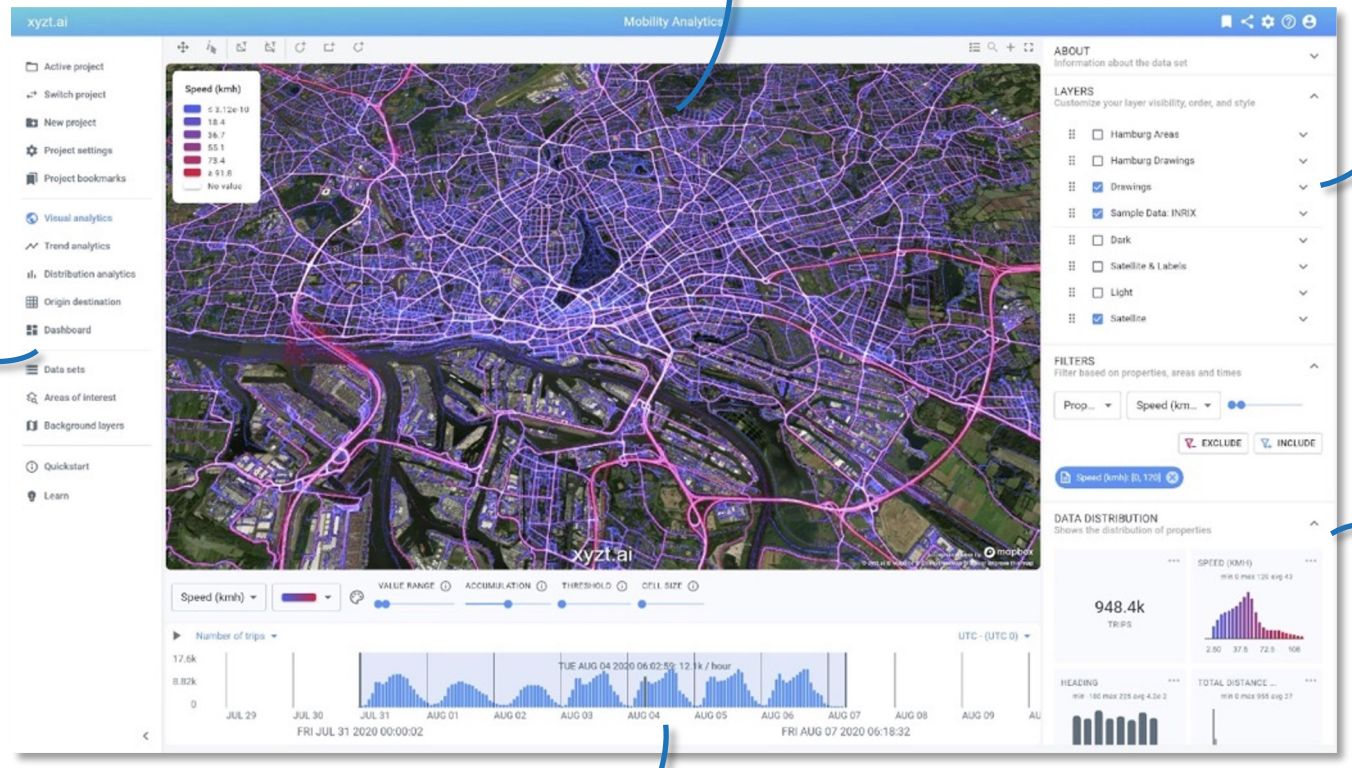
xyzt.ai platform overview

High-level overview of the xyzt.ai mobility analytics platform.

xyzt.ai

Unparalleled performance scaling to **billions of** floating vehicle data or traffic counting records

Built-in visual, trend, distribution, origin-destination **analytics**



Bring your **own data** & fuse **multiple** data sources

As **easy** as a BI tool, no coding required

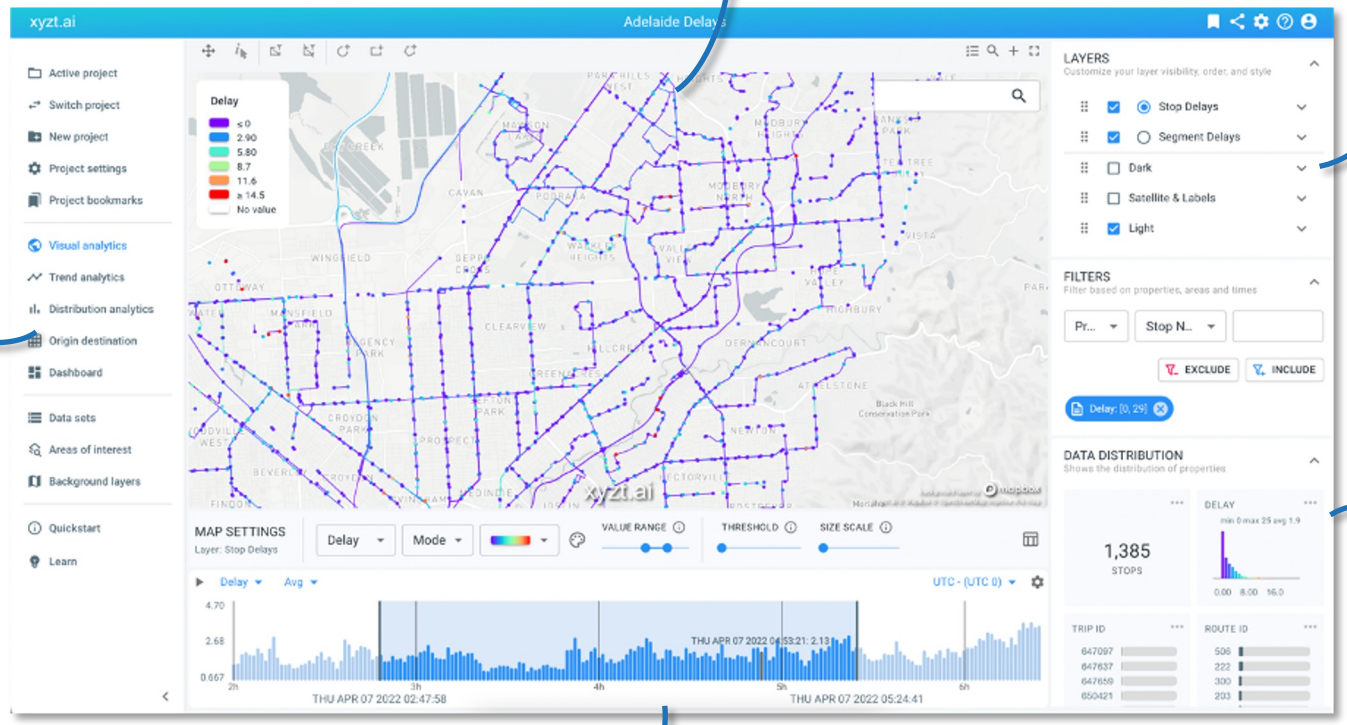
time handled as a core dimension, enabling before/after comparison analysis

xyzt.ai

Unparalleled performance scaling to **billions of** data records, both free-form and as time-series data

Bring your **own data** & fuse **multiple** data sources

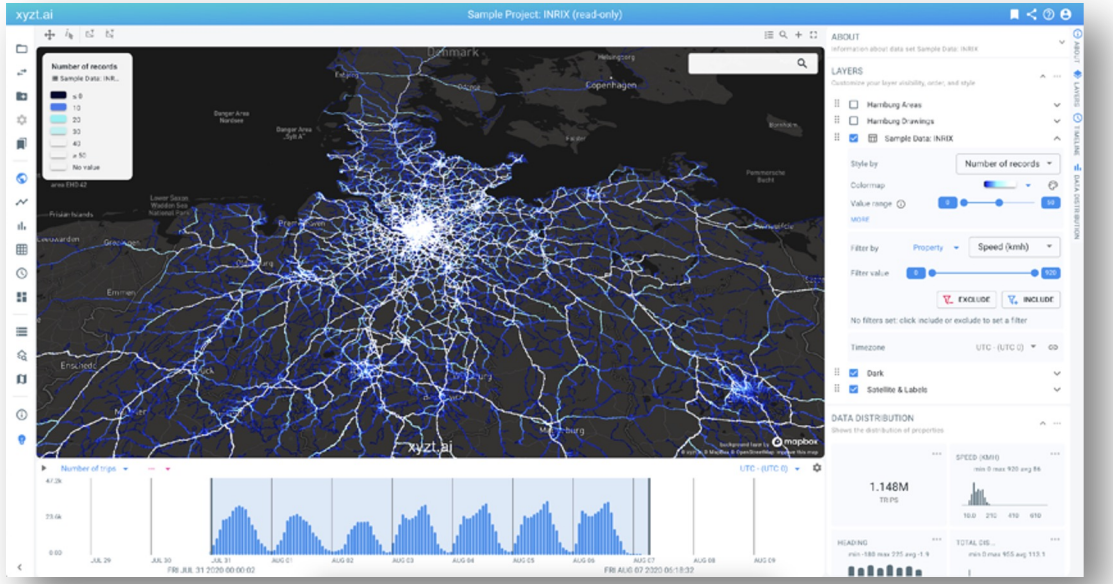
Built-in visual, trend, distribution, origin-destination, dwell **analytics**



As **easy** as a BI tool, no coding required

time handled as a core dimension, enabling before/after comparison analysis

MOVEMENT DATA



- ID=DeviceAB01
- Lon=51,244
- Lat=20,4555
- Time=15u14
- Velocity=6 km/h

record 2

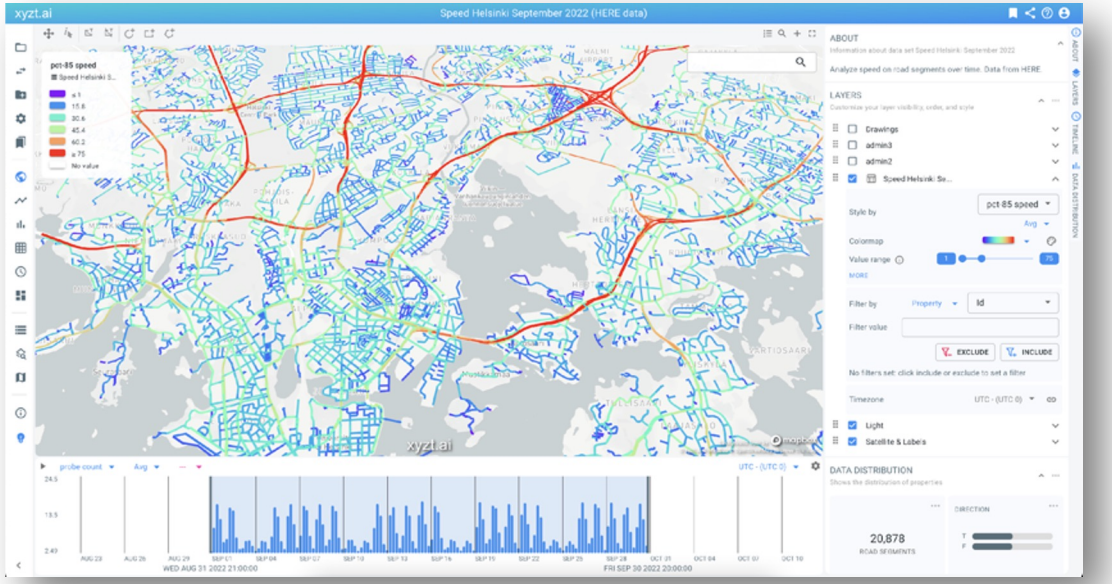
DeviceAB01

record 1
- ID=DeviceAB01
- Lon=51,234
- Lat=20,456
- Time=15u13
- Velocity=5 km/h

record 3
- ID=DeviceAB01
- Lon=51,257
- Lat=20,455
- Time=15u15
- Velocity=1 km/h

record 4
- ID=DeviceAB01
- Lon=51,258
- Lat=20,4545
- Time=15u17
- Velocity=0 km/h

TIME SERIES DATA



Average speed

Record1
ID=S123
Time=15u
speed=66

Record2
ID=S123
Time=16u
speed=60

time

Averag e speed

Record1
ID=S567
Time=15u
speed=57

Record2
ID=S567
Time=16u
speed=58

time

- ID=S123
- Lon0=50,124
- Lat0=20,152
- Lon1=50,224
- Lat1=20,176

Street 1

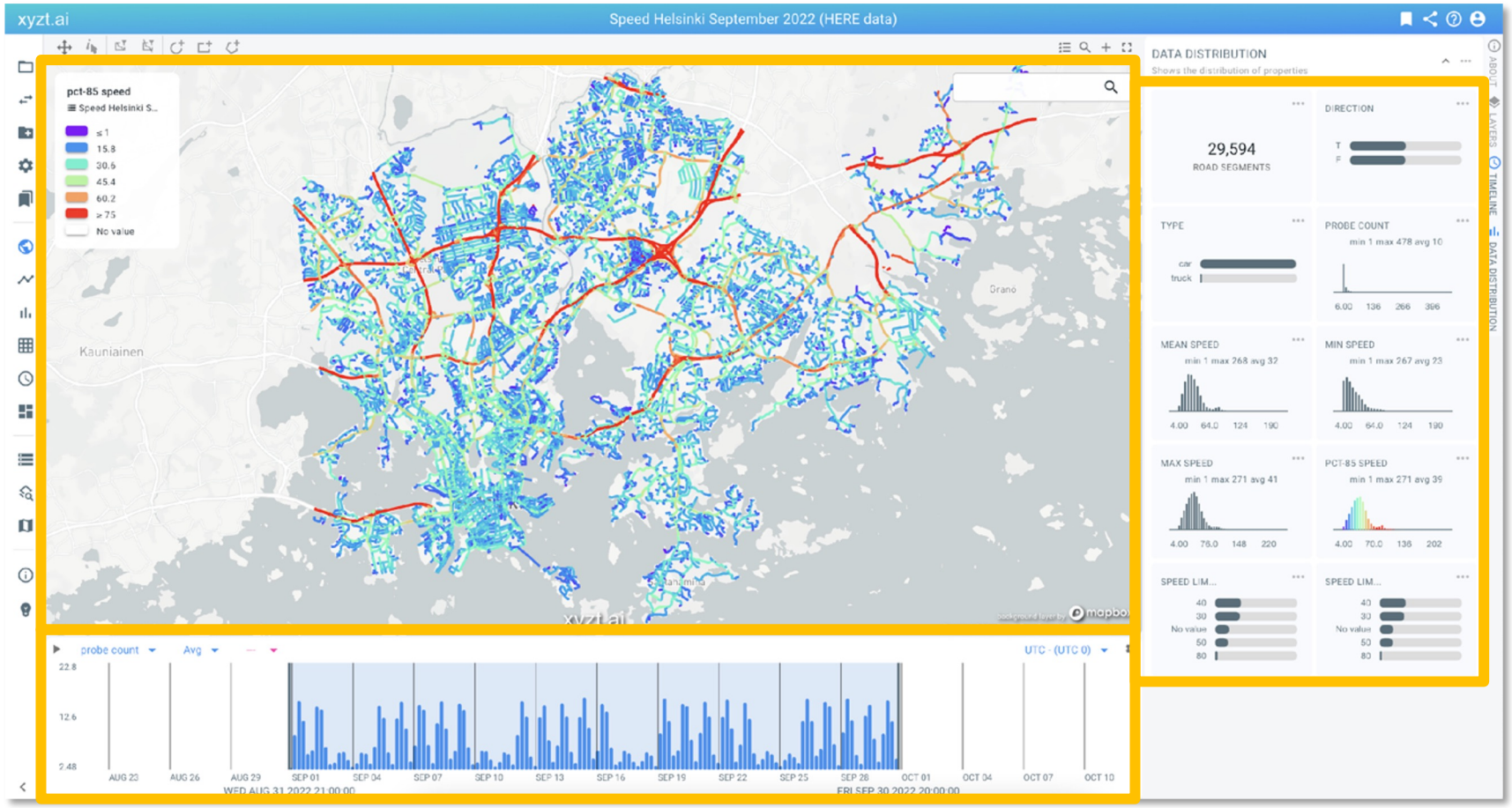
- ID=S567
- Lon0=50,224
- Lat0=20,256
- Lon1=50,124
- Lat1=20,296

Street 2

xyzt.ai: Visual Analytics

1 or 2 maps with data driven styling

- Color mapping
- Density mapping
- …



Distributions on data properties

- Min
- Max
- Average
- Distribution

Interactive timeline with data-driven trends (number of vehicles, average speed,…)

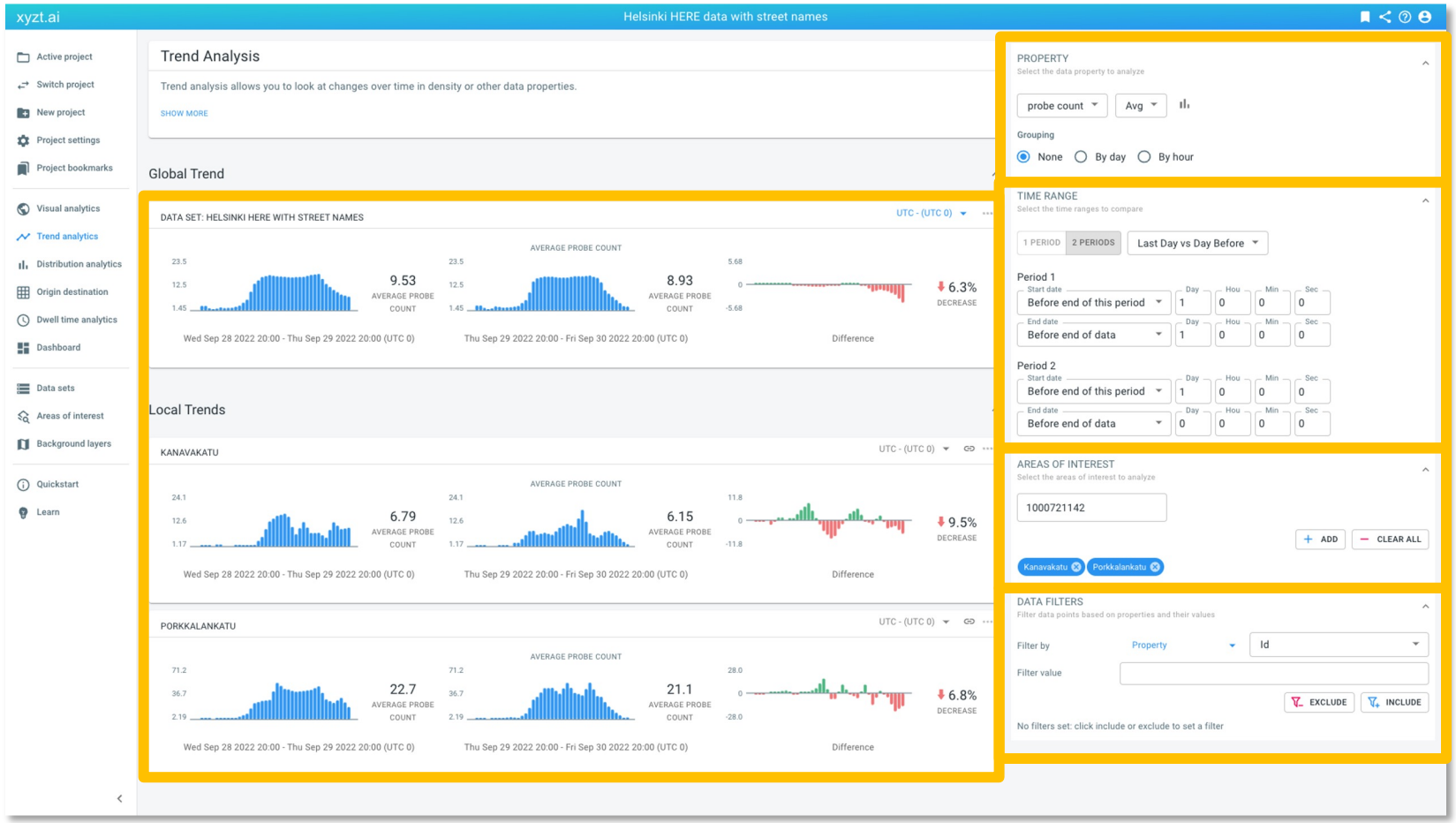
xyzt.ai: Trend Analytics

Global and local trends

1 or 2 periods

Trends of data properties such as number of vehicles, speeds,…

Local trends on regions in the data, e.g., streets, areas,…



Select property, aggregation, and/or grouping by hour/day

Select 1 or 2 time periods

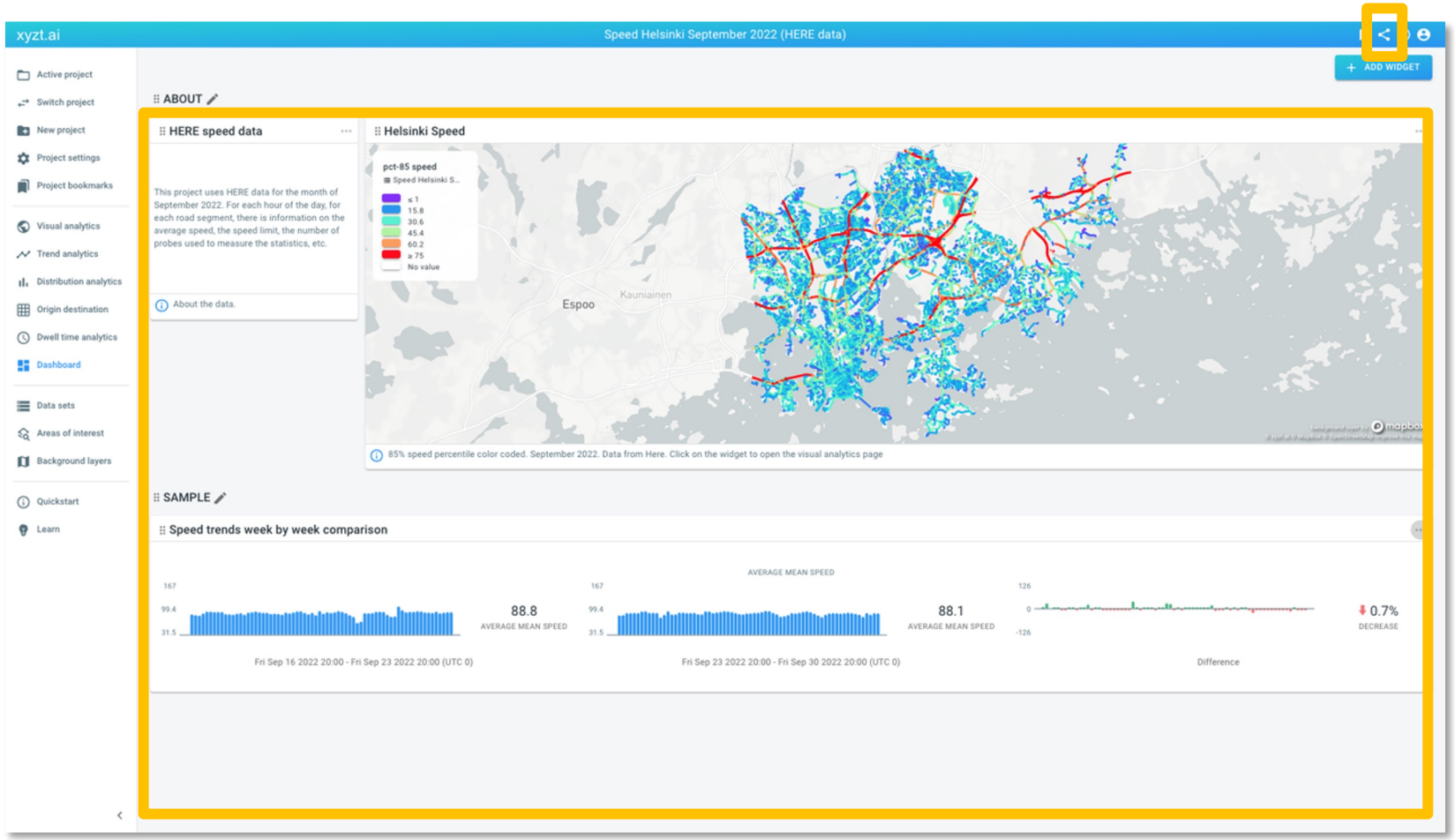
Select local areas (e.g., streets)

Further filter on the data

xyzt.ai: Dashboarding & reporting

Add any widget (map, trend line, HTML widget, distribution bar chart,…) on a dashboard for easy reporting.

Download widget data as CSV or PNG



Create shared links of the dashboard.

xyzt.ai: Dashboarding & reporting

Two main ways to use xyzt.ai for reporting

- Create a **dashboard** with widgets and create a **shared link**
  - Allow recipients to click on widgets and do further analysis, or
  - Restrict recipients to dashboard

- Extract parts of the platform to include, for example in a Word doc
  - **Save as PNG** available on every visual
  - **Save as CSV** and open for example in XLS/PowerBI/…

In addition, you can load/**draw polygons with labels** on the map

And you can **embed** any map/view/dashboard of xyzt.ai using iframes

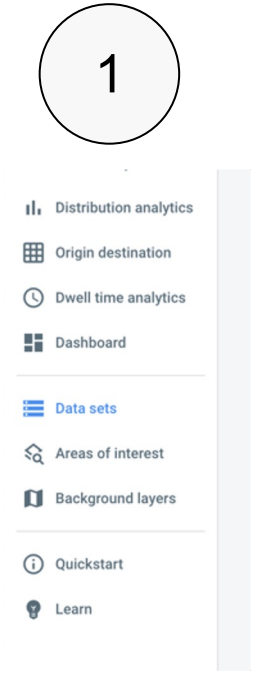
xyzt.ai: Adding data

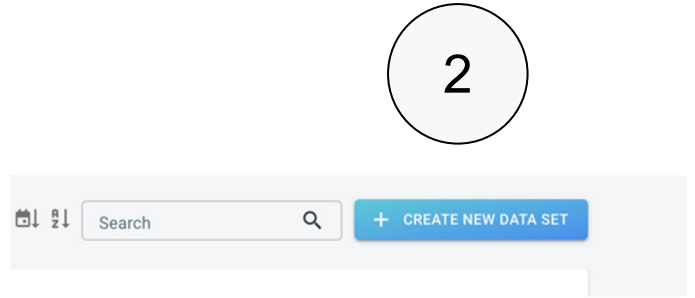
Using the **user interface**

- Data Sets (1) → Create New Data Set (2)
- Follow steps (3)
  - Upload geometry as GeoJSON
  - Upload time series as CSV
- Follow a tutorial here:
  - https://docs.platform-xyzt.ai/tutorials/upload-your-own-time-series-data/goal.html
  - https://docs.platform-xyzt.ai/tutorials/upload-your-own-data/goal.html

Using the **REST API**

- Used often for real-time data, but can also be used for historical data ingestion and automation
- https://docs.platform-xyzt.ai/tutorials/using-the-api/goal.html





Data Coverage and Quality Overview

Overview of the origins and coverage/quality of the different data sources used in the project.

Applicable Mobility Use Cases

Overview of mobility and traffic analytics use cases that can be solved with the data and the platform.

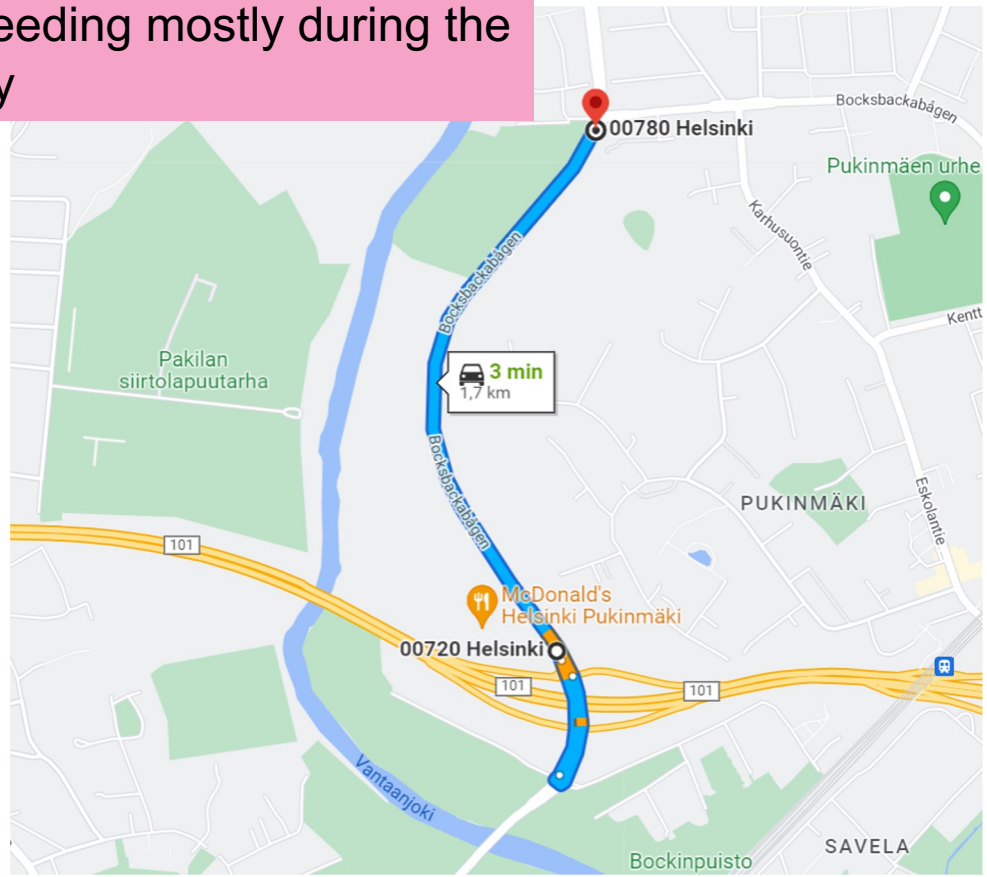
	Speeding/Congestion analysis	Origin-destination analysis	Travel time analysis	Delay analysis	Traffic density analysis	Multi-modality analysis	Dangerous events (acceleration, breaking, cornering,…)	Traffic light performance analysis	Emission analysis
	Analyze where and when traffic is speeding or where and when traffic is congested.	Analyze where traffic comes from and goes to, e.g., analyzing how many people move between different areas during different times of the day.	Analyze how long trips take between different areas and possibly analyse how travel times are impacted due to for example road works, detours, etc.	Analyze where most delay is seen along different trips.	Analyze the amount of traffic and change in amount of traffic at certain locations and times.	Analyze the modal split, such as what percentage of people drive by car, bicycle,…	Where and when do most dangerous events happen, such as people cornering on a highway ramp.	Where do traffic lights perform poorly, i.e., where do people have to wait a long time in front of a red light.	What is the impact of a LEZ, what is the impact of a change in traffic plan on the emissions?
HERE road stats	✅				✅ (only relative)				
TomTom MOVE	✅	✅	✅	✅	✅ (only relative)			✅ (Junction analytics)	
ODIQ	✅		✅	✅					
Traffic counts	✅	✅			✅	✅			
Telia origin destination		✅				Telia is working on this			
Copernicus atmospheric monitoring data									✅ (very coarse though)
Otonomo, Wejo, Bridgestone, INRIX,…	✅	✅	✅	✅ (less straightforward)	✅ (only relative)	Some data providers are working on this	✅ (e.g., Bridgestone abc data)	✅	✅ (model based on vehicle type, speed,…)

Helsinki Specific Mobility Use Cases

Use cases brought forward by Helsinki to investigate during the project with the data and platform.

Speeding Analysis

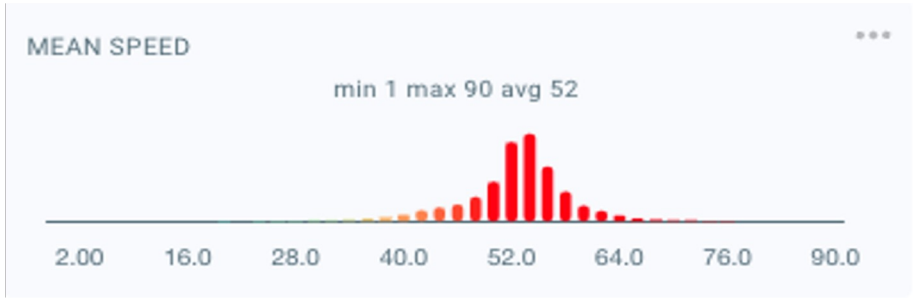
Conclusion: lots of speeding mostly during the day

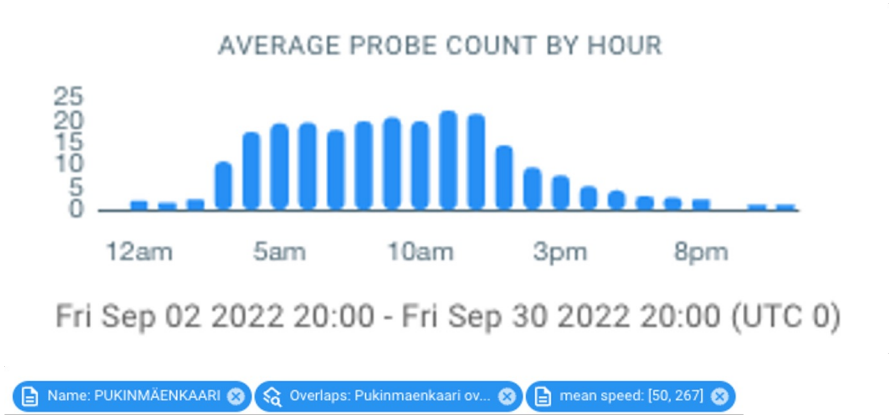


Pukinmäenkaari
https://www.google.com/maps/dir/60.2424945,24.9806478/60.2512567,24.9819422/@60.2466703,24.971509,14.73z/data=!4m2!4m1!3e0

Speed limit = 50 kmh

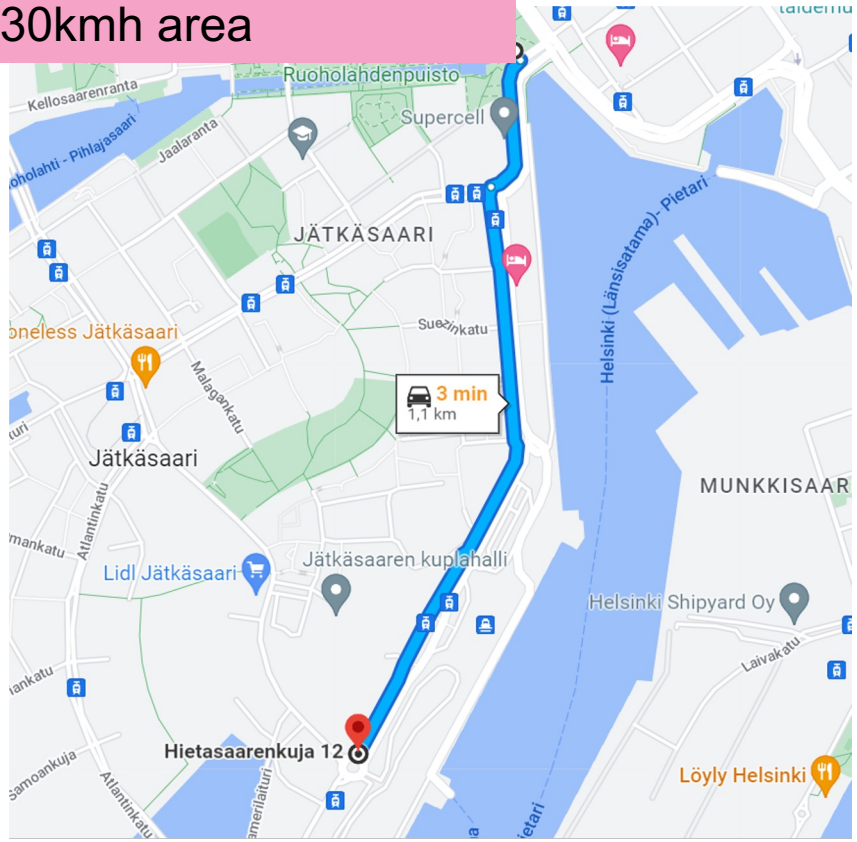
Mean speed all directions





Data used: HERE

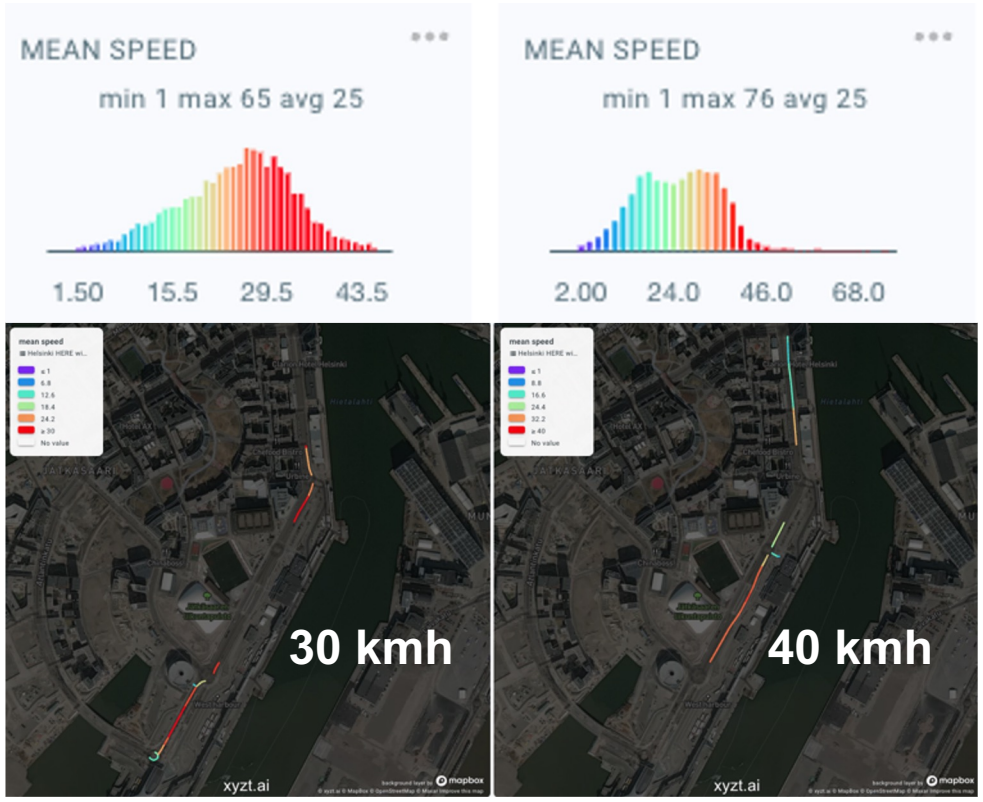
Conclusion: speeding in the 30kmh area



Tyynenmerenkatu:
https://www.google.com/maps/dir/60.2424945,24.9806478/60.2512567,24.9819422/@60.2466703,24.971509,14.73z/data=!4m2!4m1!3e0

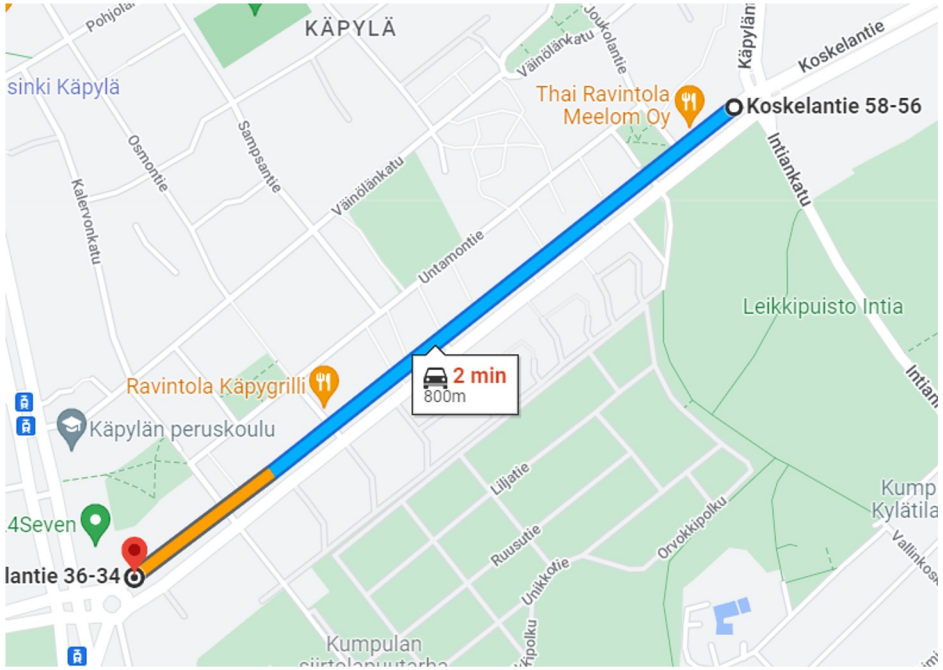
Speed limit = 30 kmh and 40 kmh

HERE mean speed all directions



Data used: HERE

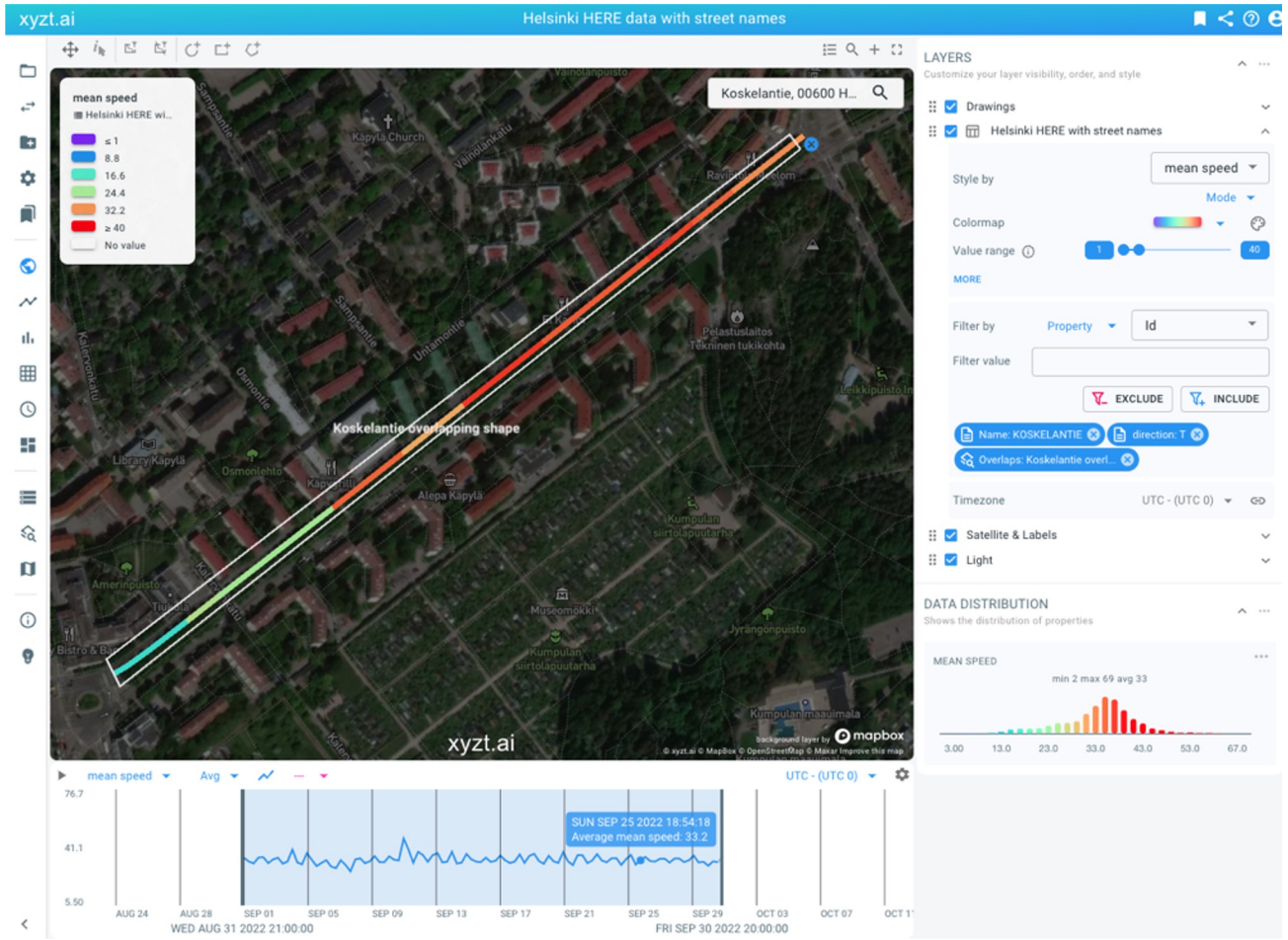
Conclusion: some speeding



Koskelantie (a school close, lots of feedback from residents): https://www.google.com/maps/dir/60.2134243,24.9581828/60.2089233,24.9465951/@60.2111784,24.951203,15.58z/data=!4m2!4m1!3e0

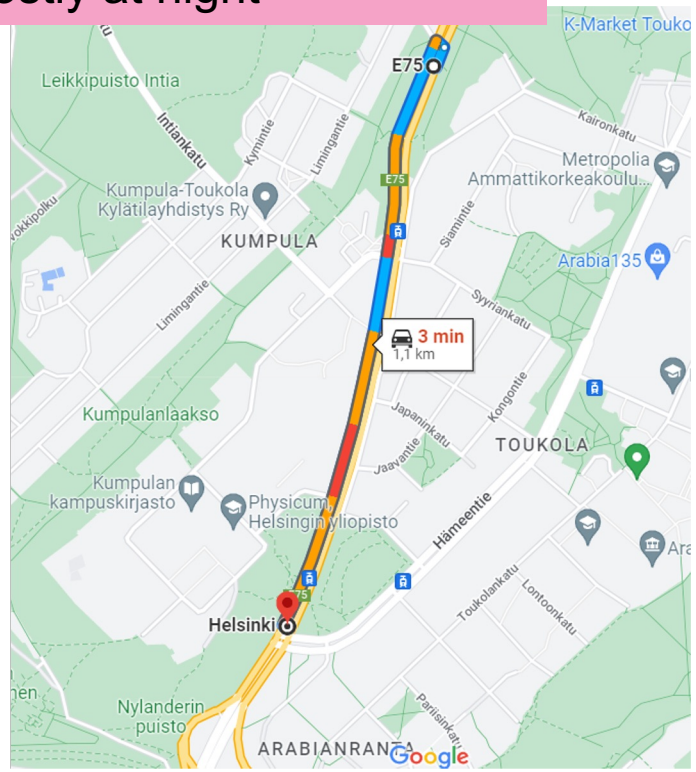
Speed limit = 40 kmh

HERE mean speed averaging at 33 kmh



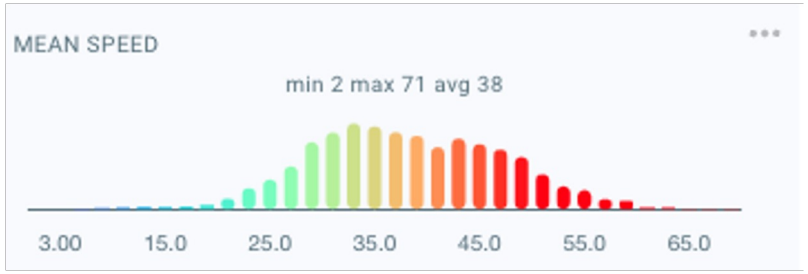
Data used: HERE

Conclusion: speeding mostly at night

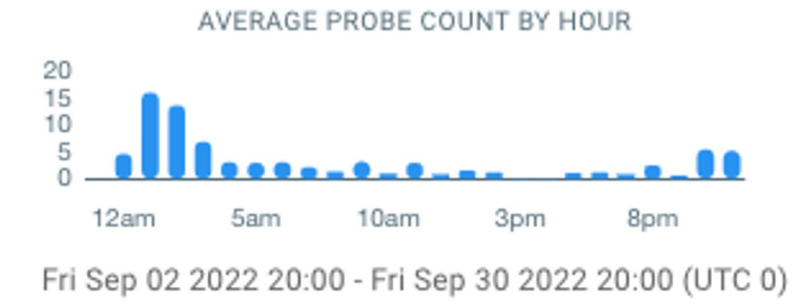


Speed limit =50 kmh

HERE mean speed averaging at 38 kmh



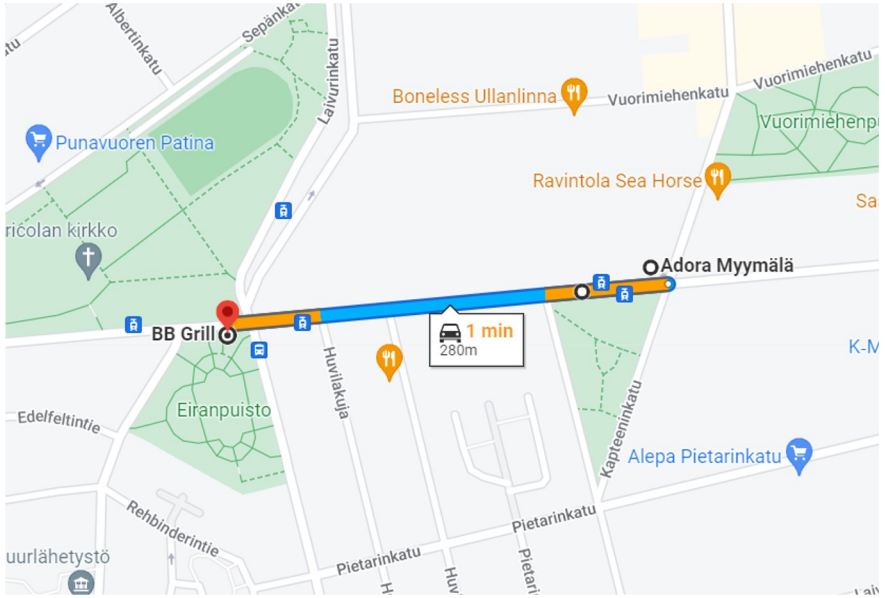
Speeding mostly at night (filtered out only mean speed > 50kmh)



Kustaa Vaasan tie:
https://www.google.com/maps/dir/60.2117602,24.9696781/60.2033044,24.9652527/@60.2071738,24.9660937,15.8z/data=!4m2!4m1!3e0

Data used: HERE

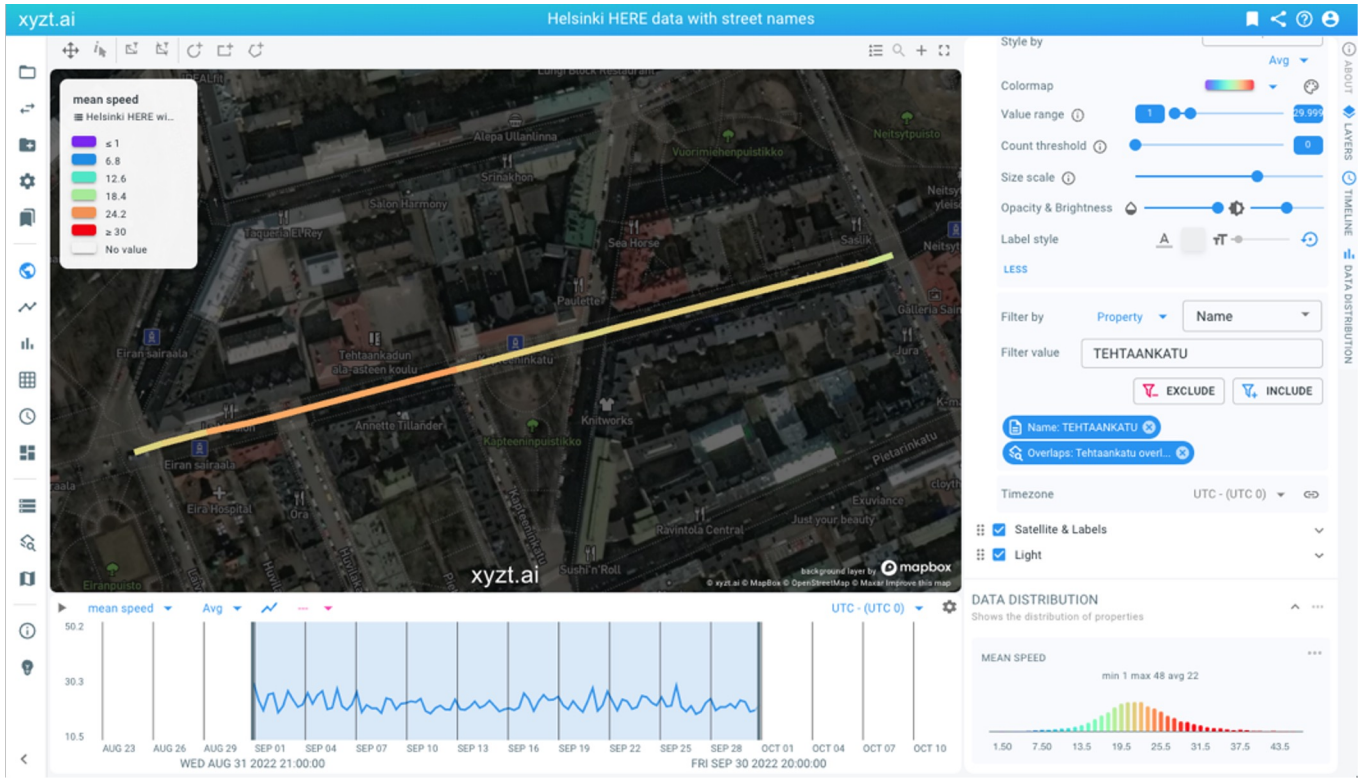
Conclusion: little evidence of excessive speeding

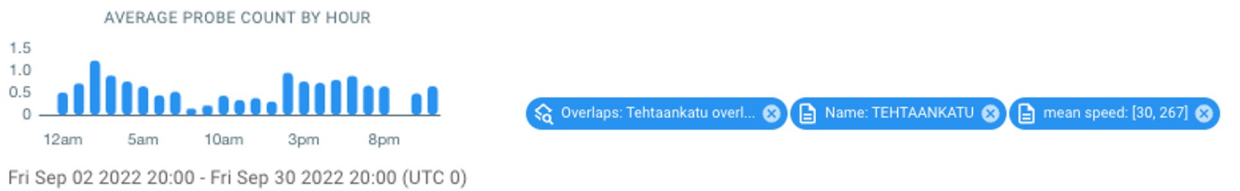


Tehtaankatu (a school close, lots of feedback from residents): https://www.google.com/maps/dir/60.1582689,24.9459715/60.1580286,24.9409591/@60.1578706,24.94141,16.51z/data=!4m9!4m8!1m5!3m4!1m2!1d24.9449569!2d60.1582107!3s0x46920bb7867cd98b:0xd46ad0f914a884f0!1m0!3e0

Speed limit = 30 kmh

HERE mean speed averaging at 22 kmh



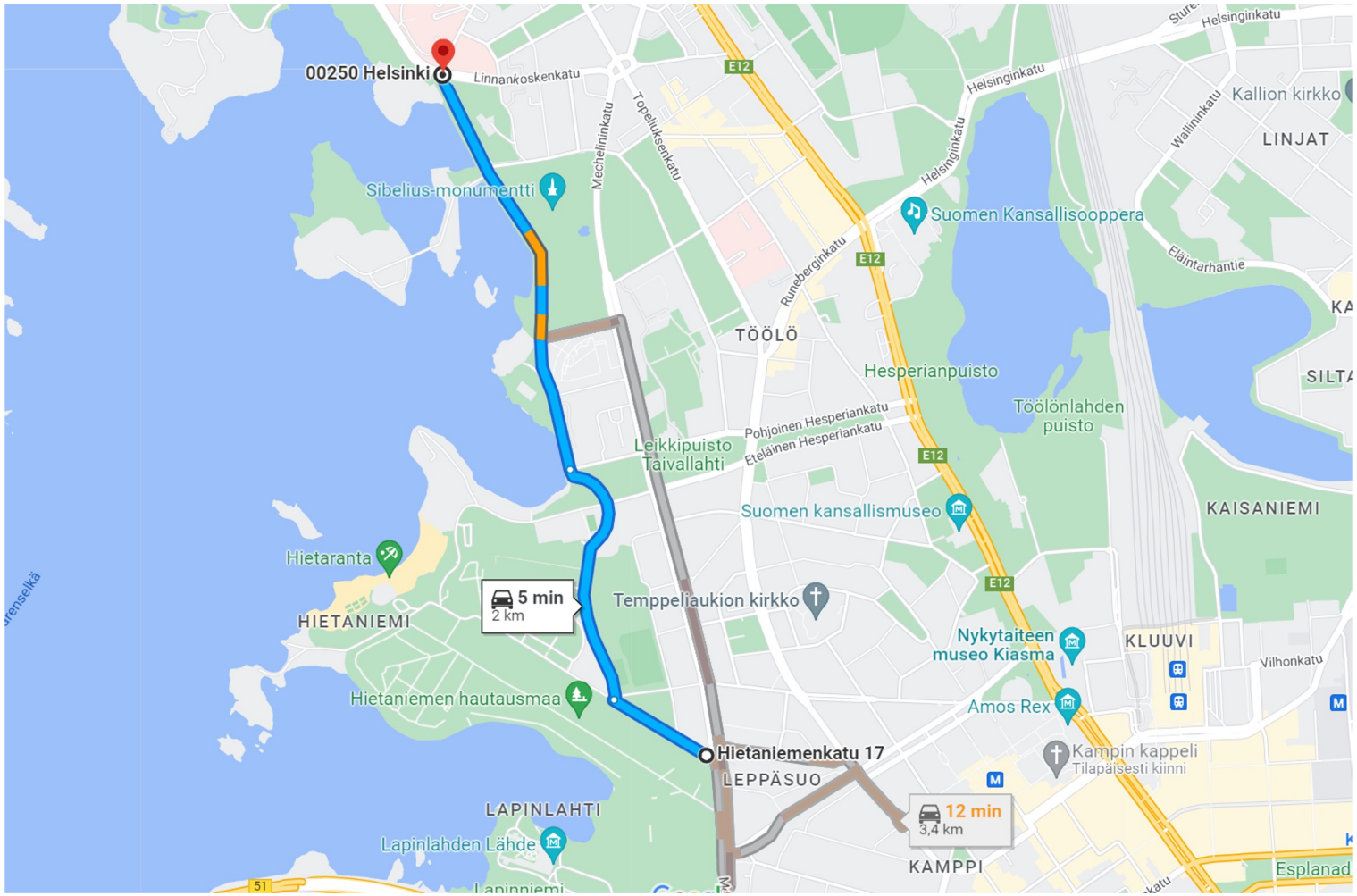


Data used: HERE

Traffic counting vs floating vehicle data analysis

Routes and alternatives analysis

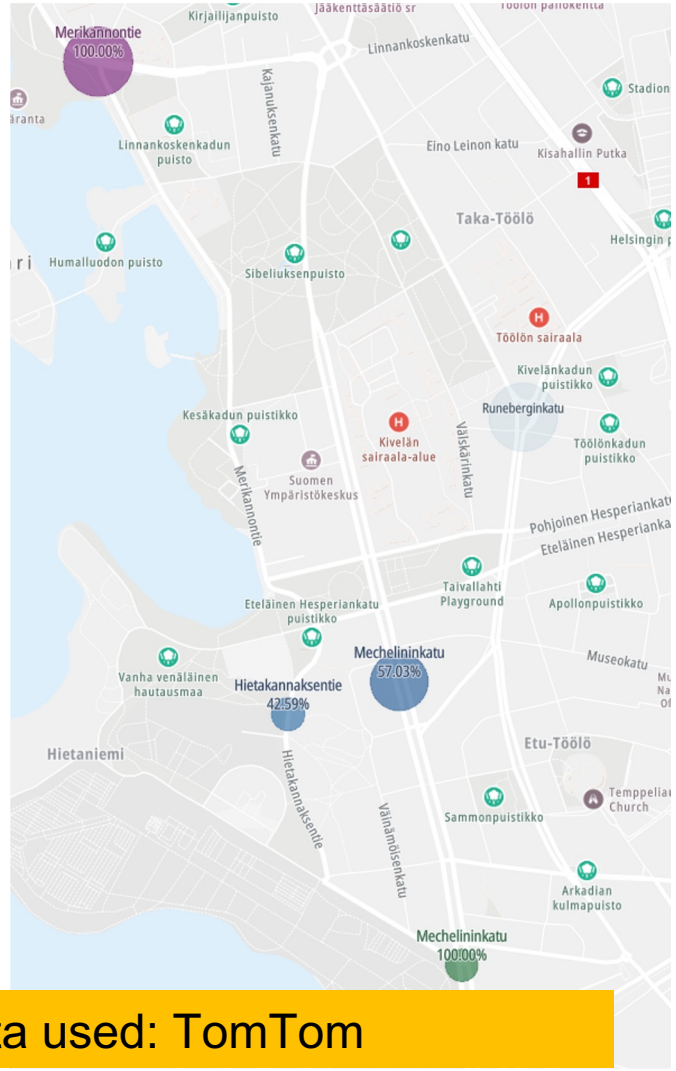
Route analysis: Mechelininkatu-Linnankoskenkatu

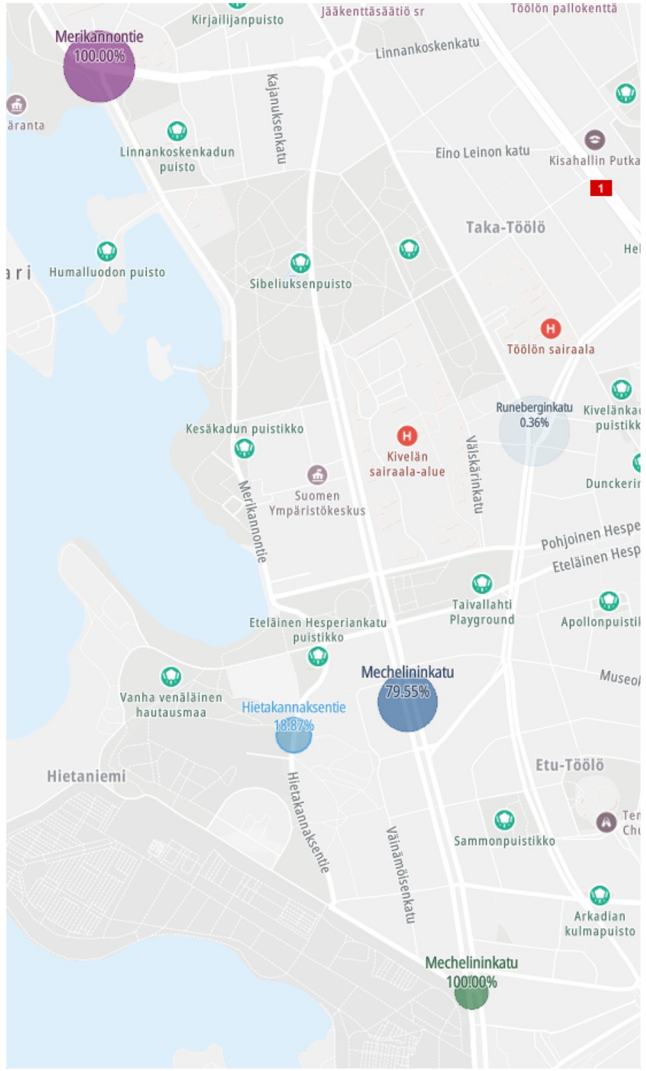


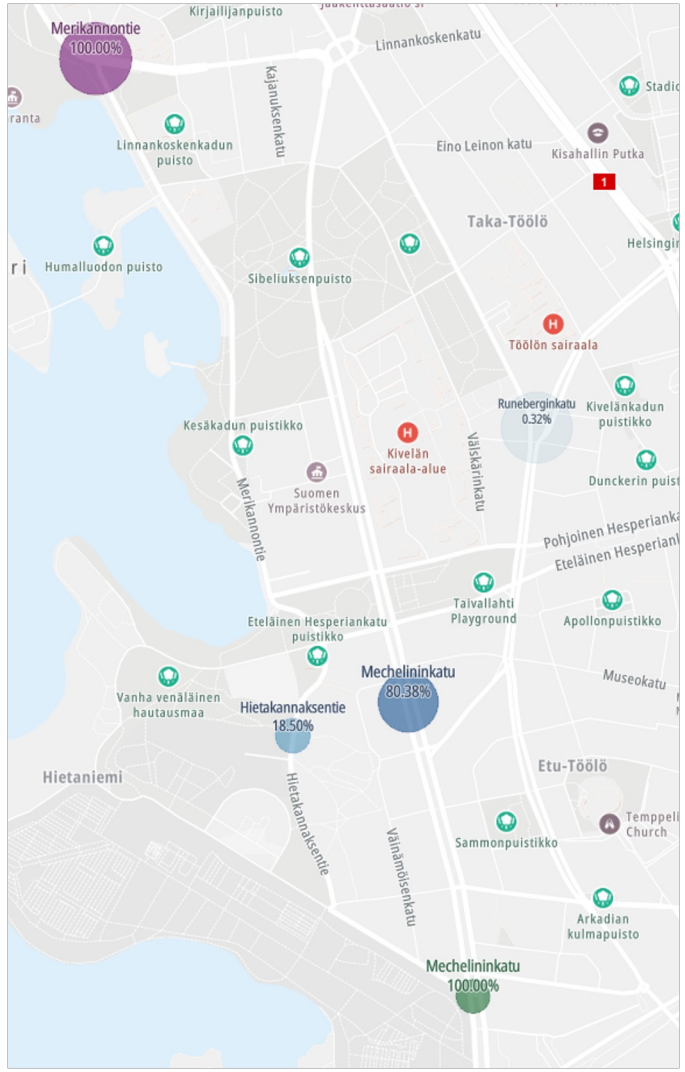
How popular is this route as a rat-running route?

Shortcut is taken mostly in the morning (42%)

Route analysis: Mechelininkatu-Linnankoskenkatu







Data used: TomTom

Helsinki

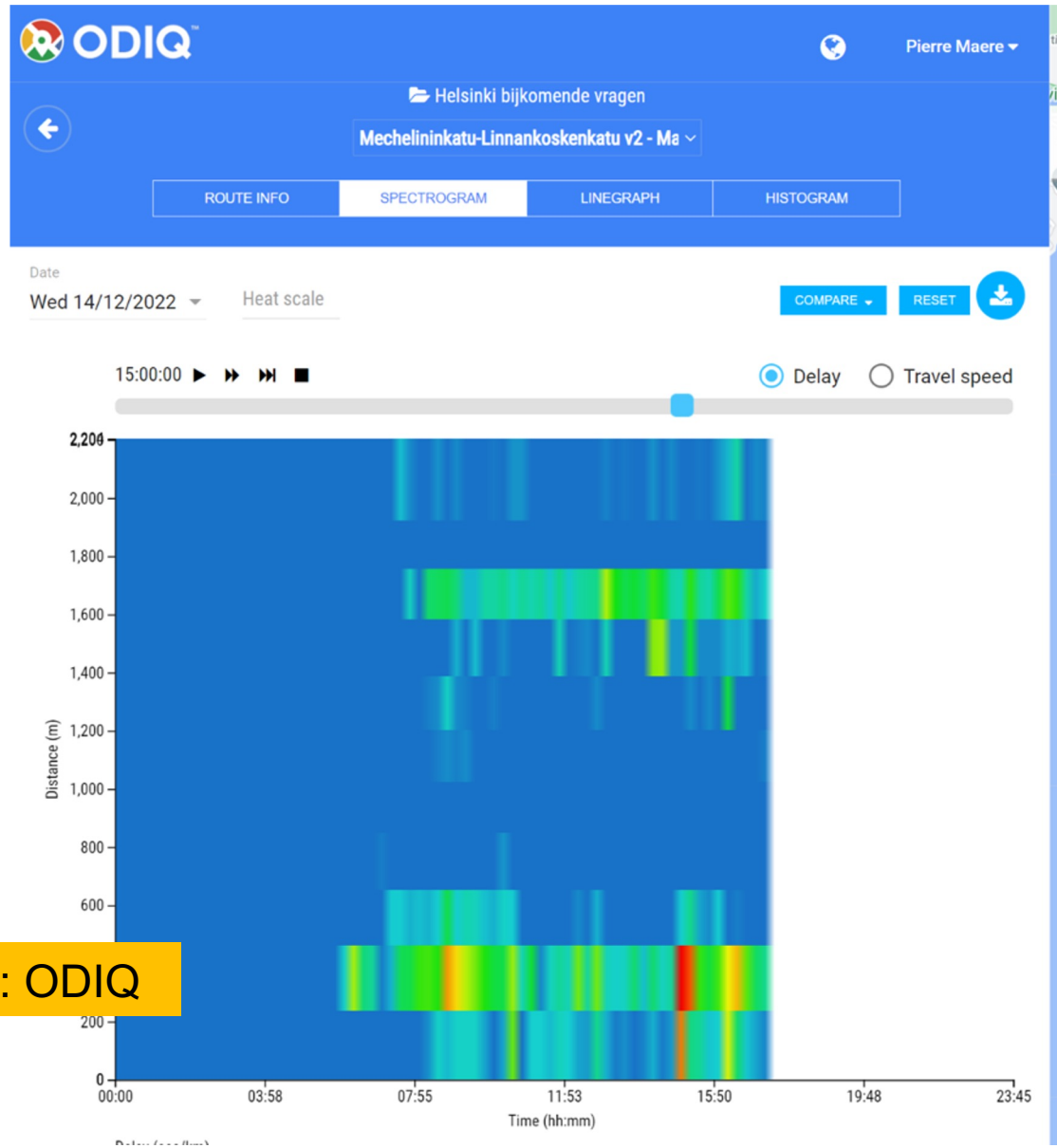
Morning (6-9 am)

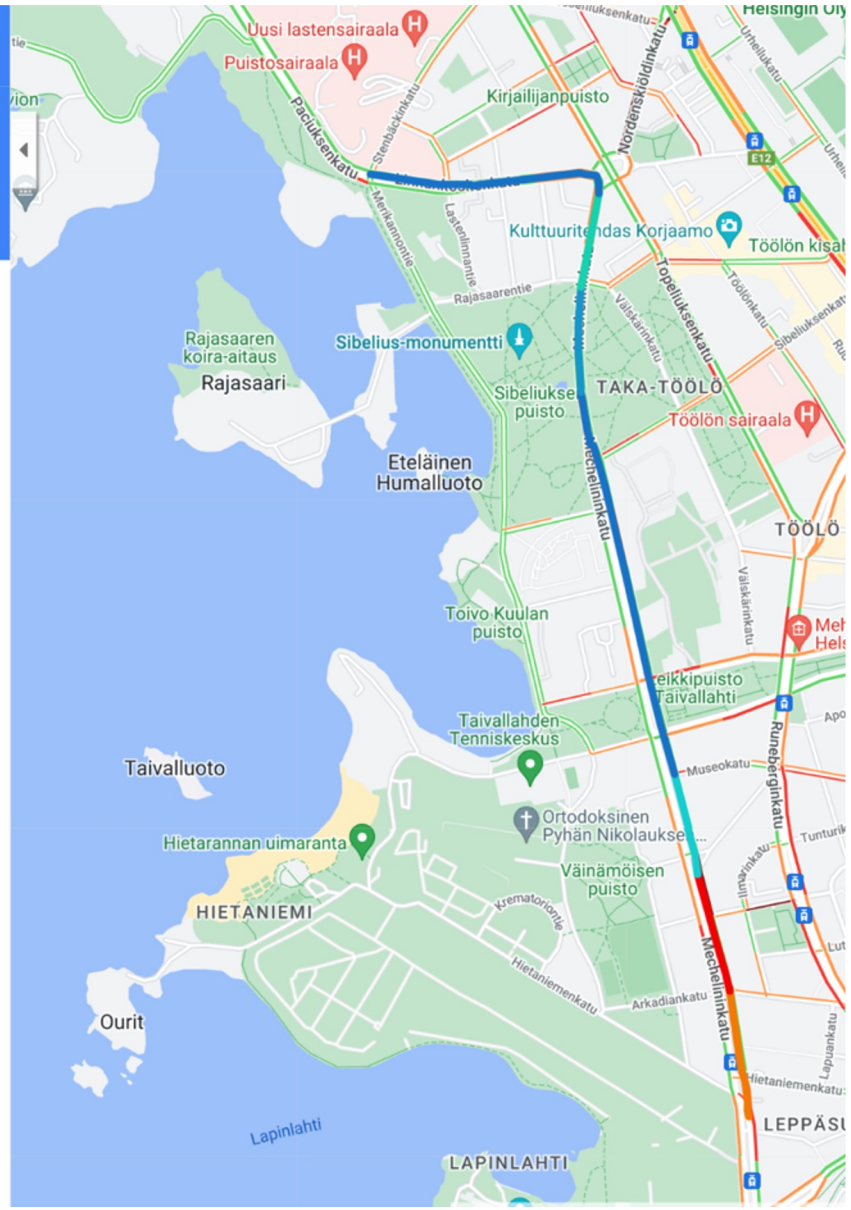
All day

Evening 16-19h

Slowing down mainly at the start of trajectory

Route analysis: Mechelininkatu-Linnankoskenkatu



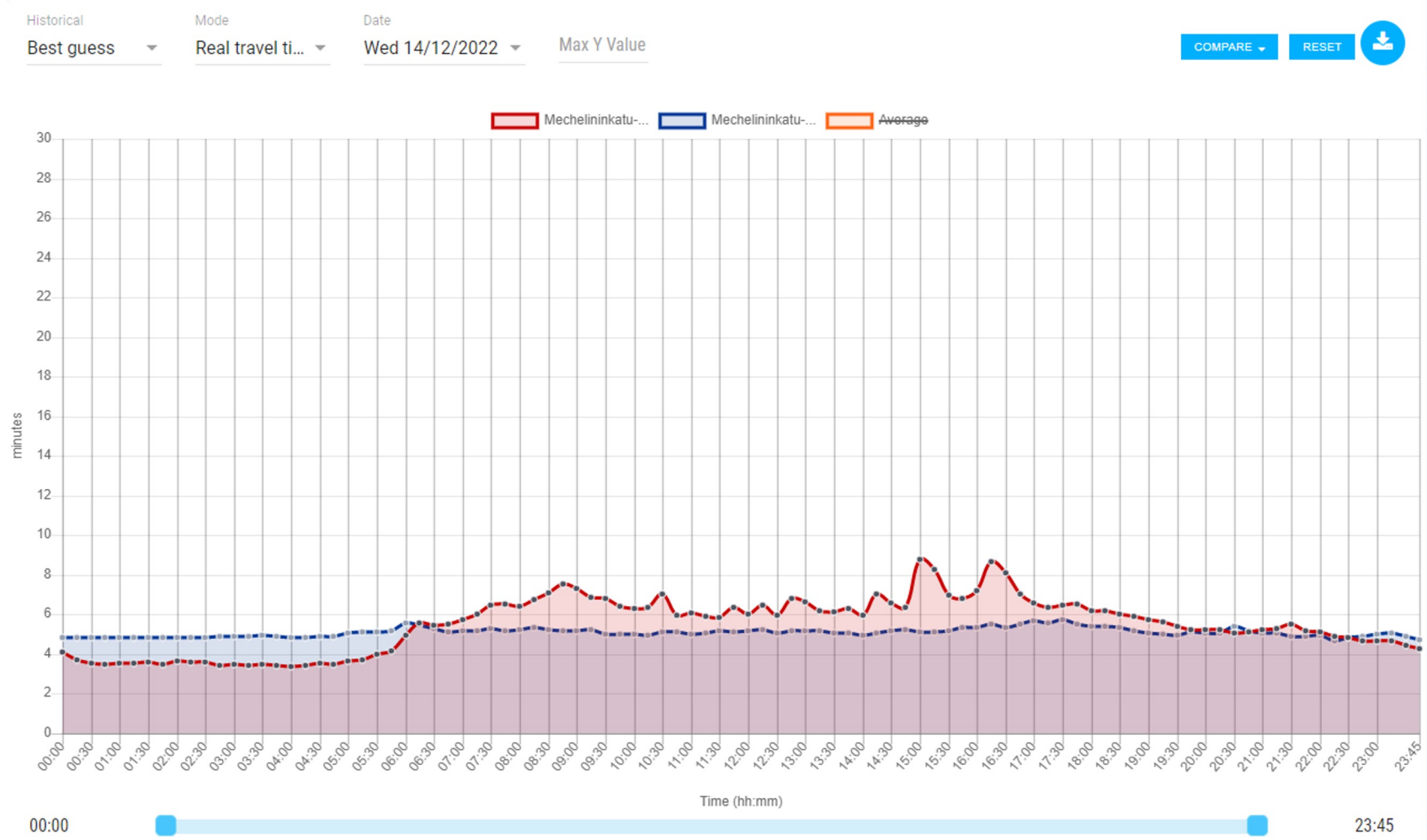


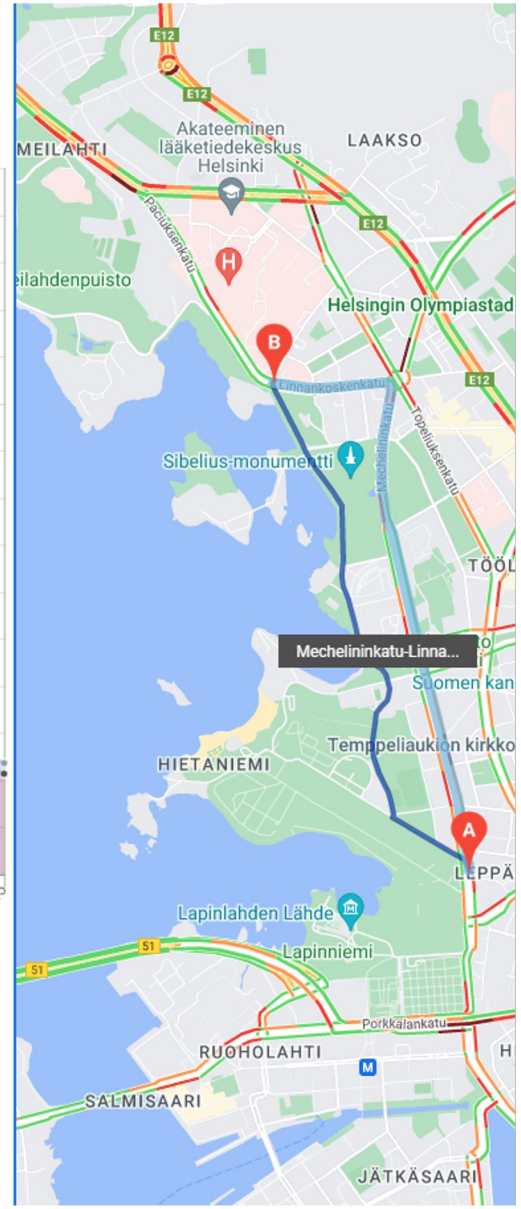
Data used: ODIQ

Helsinki

Drive-through on the blue route (Mechelininkatu-Linnankoskenkatu)

The shortcut is always faster during the day

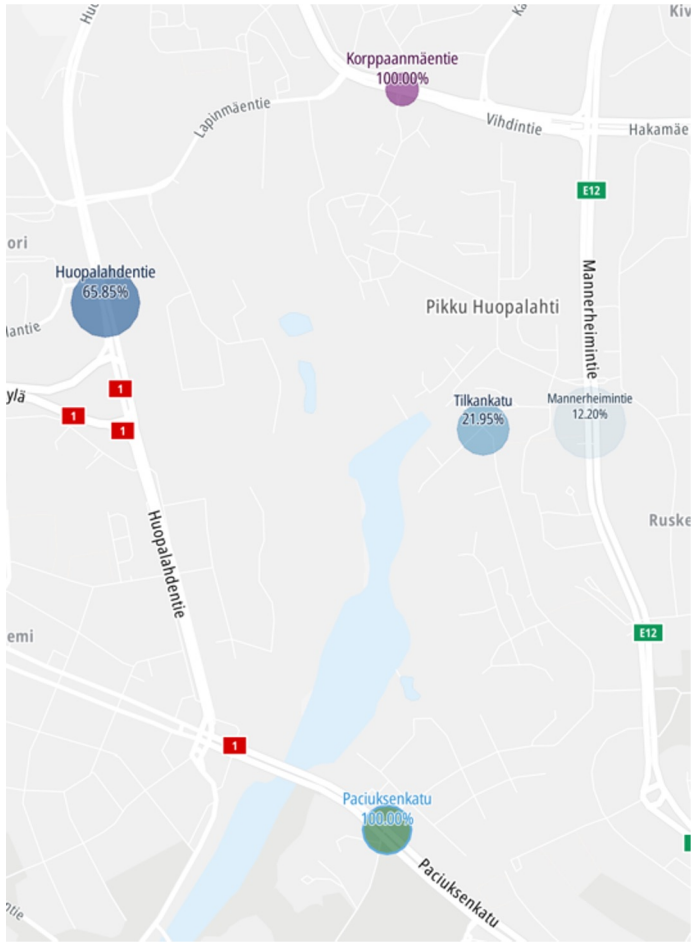


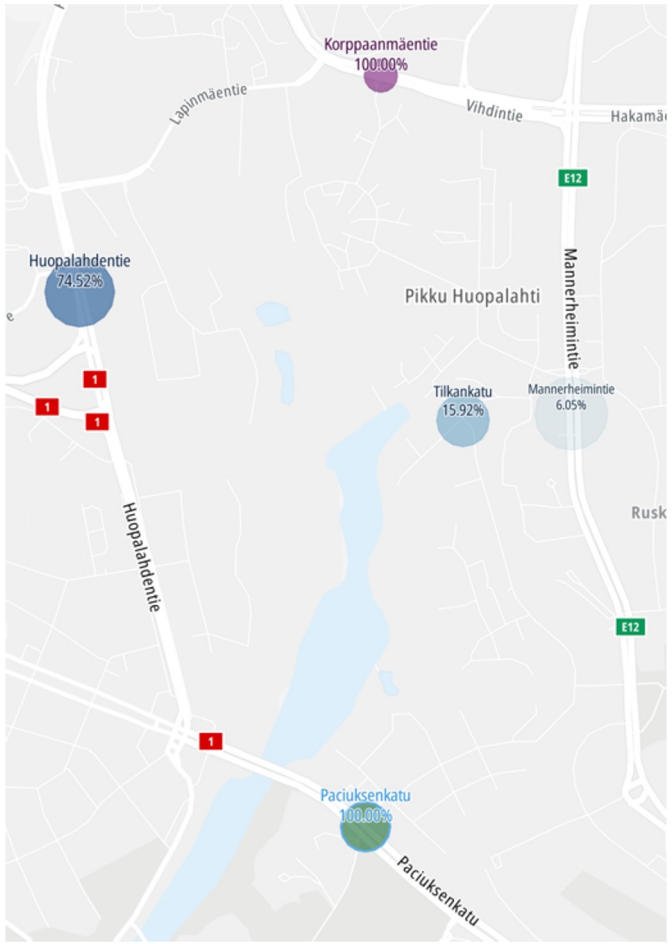


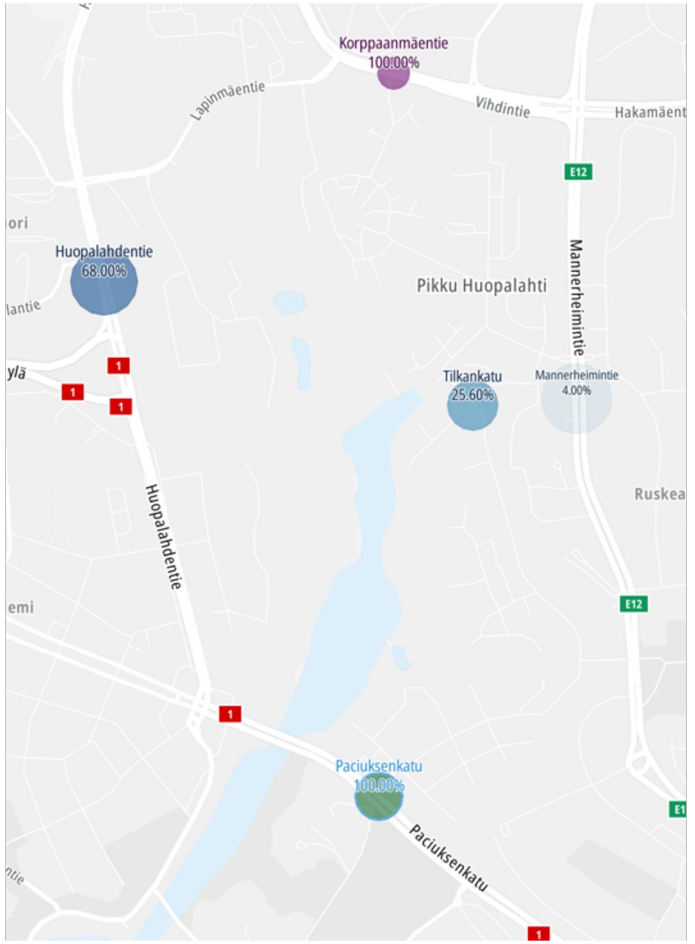
Data used: ODIQ

Drive-through on Paciuksenkatu-Vihdintie

Shortcut is taken mostly in morning and evening (20-25%)







Data used: TomTom

Helsinki

Morning (6-9 am)

All day

Evening 16-19h

Drive-through on Paciuksenkatu-Vihdintie

From
Tue 13/12/2022

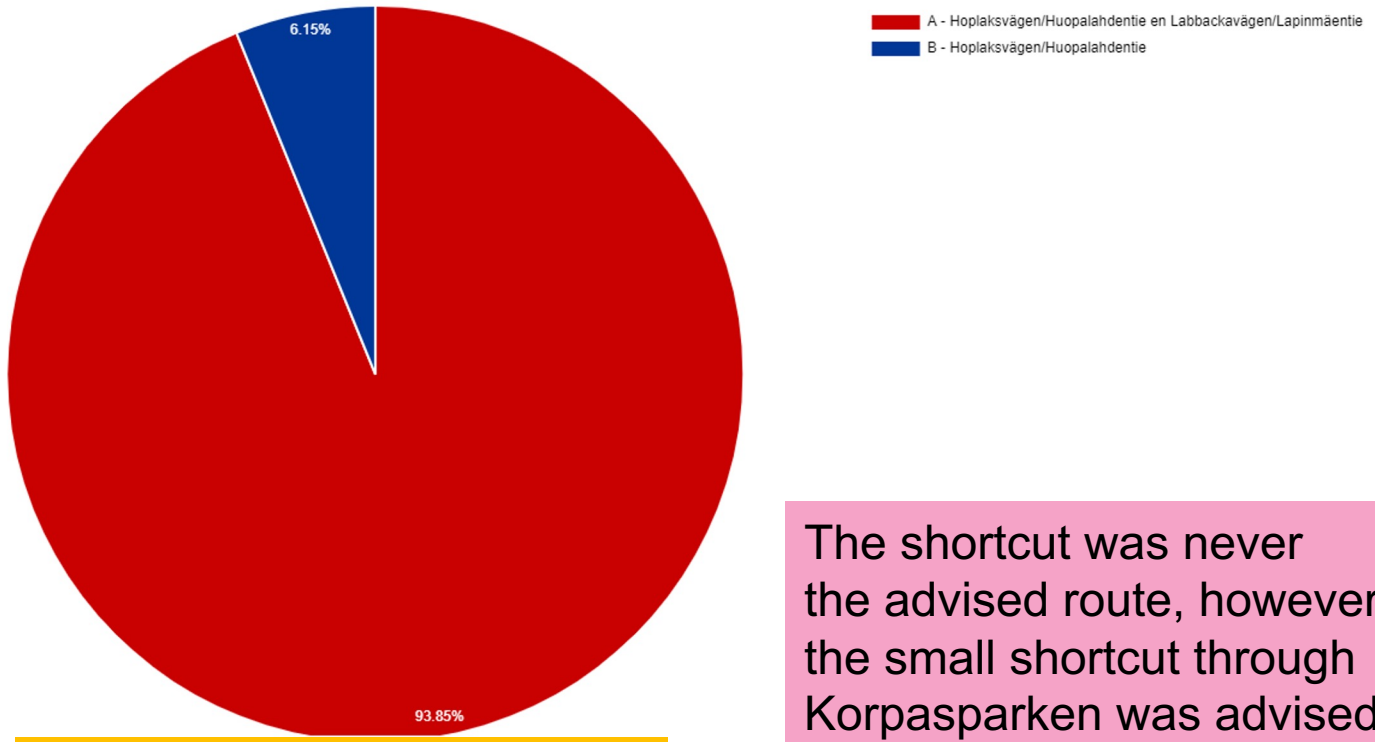
To
Thu 15/12/2022

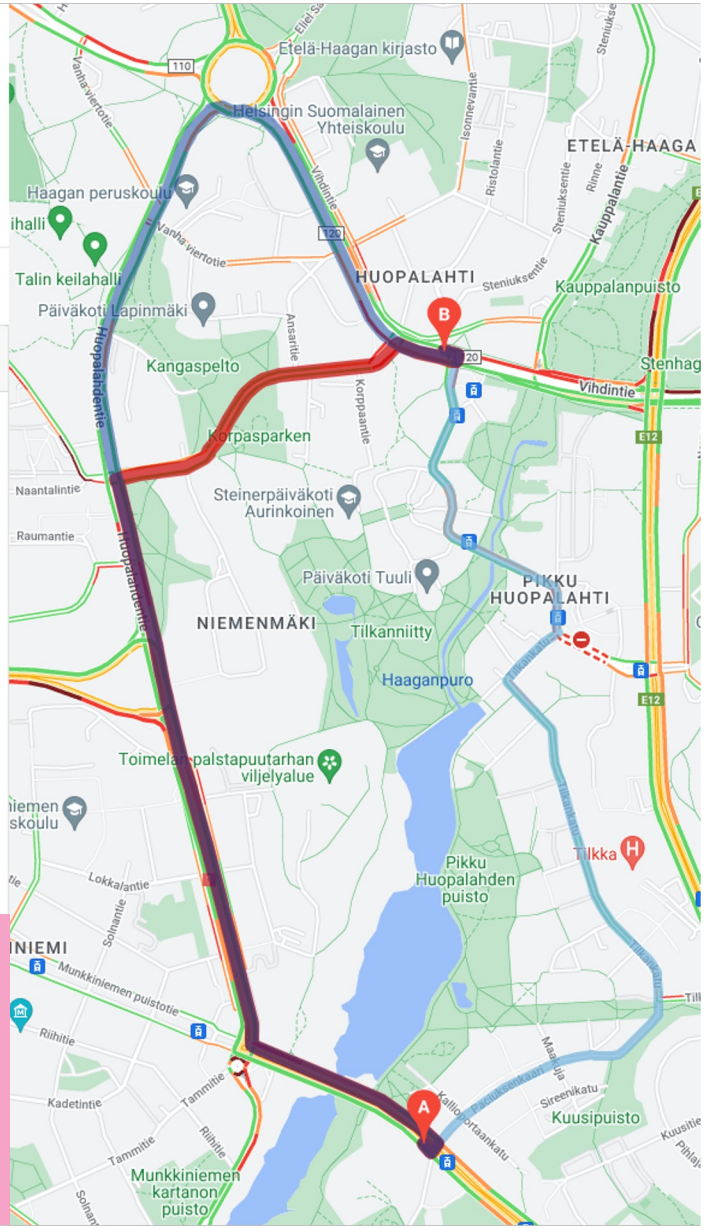
Days
MO TU WE THU FR SAT SUN

Hours
ALL(0-24)

Route information

Summary
Pie chart

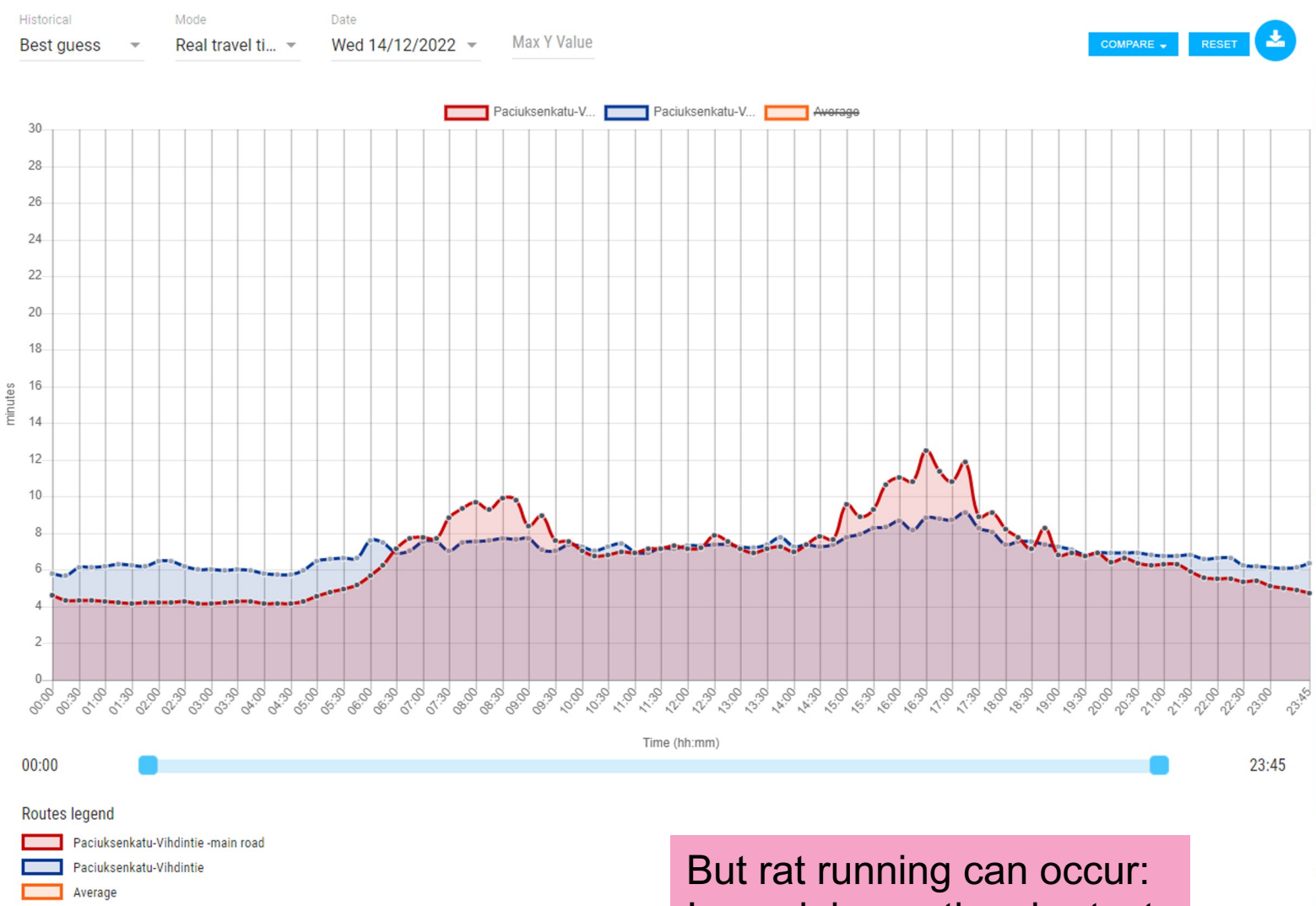


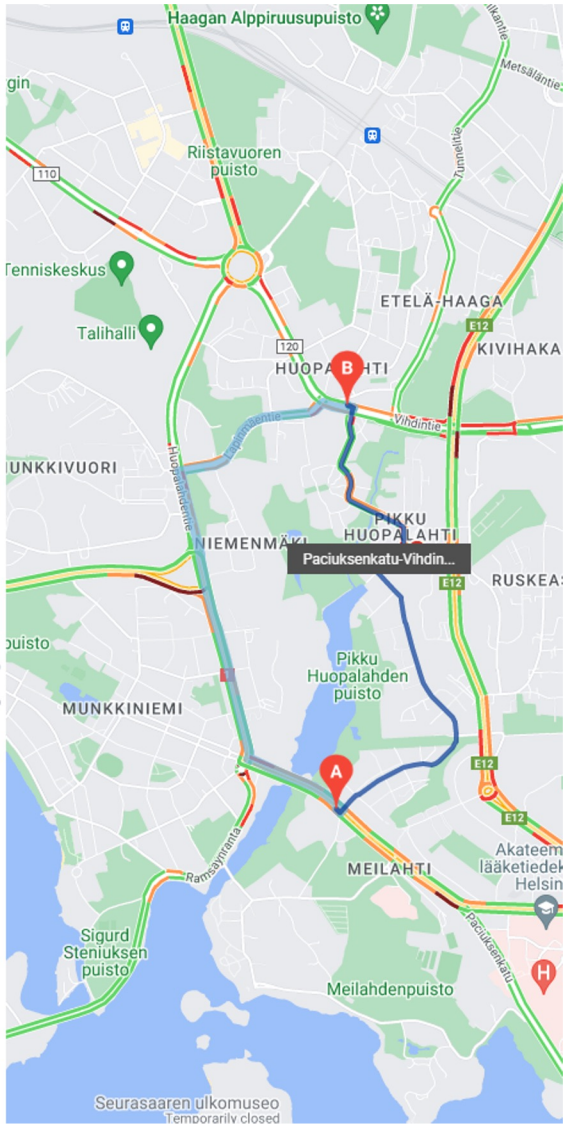


The shortcut was never the advised route, however the small shortcut through Korpasparken was advised almost 94% of time.

Data used: ODIQ

Drive-through on Paciuksenkatu-Vihdintie



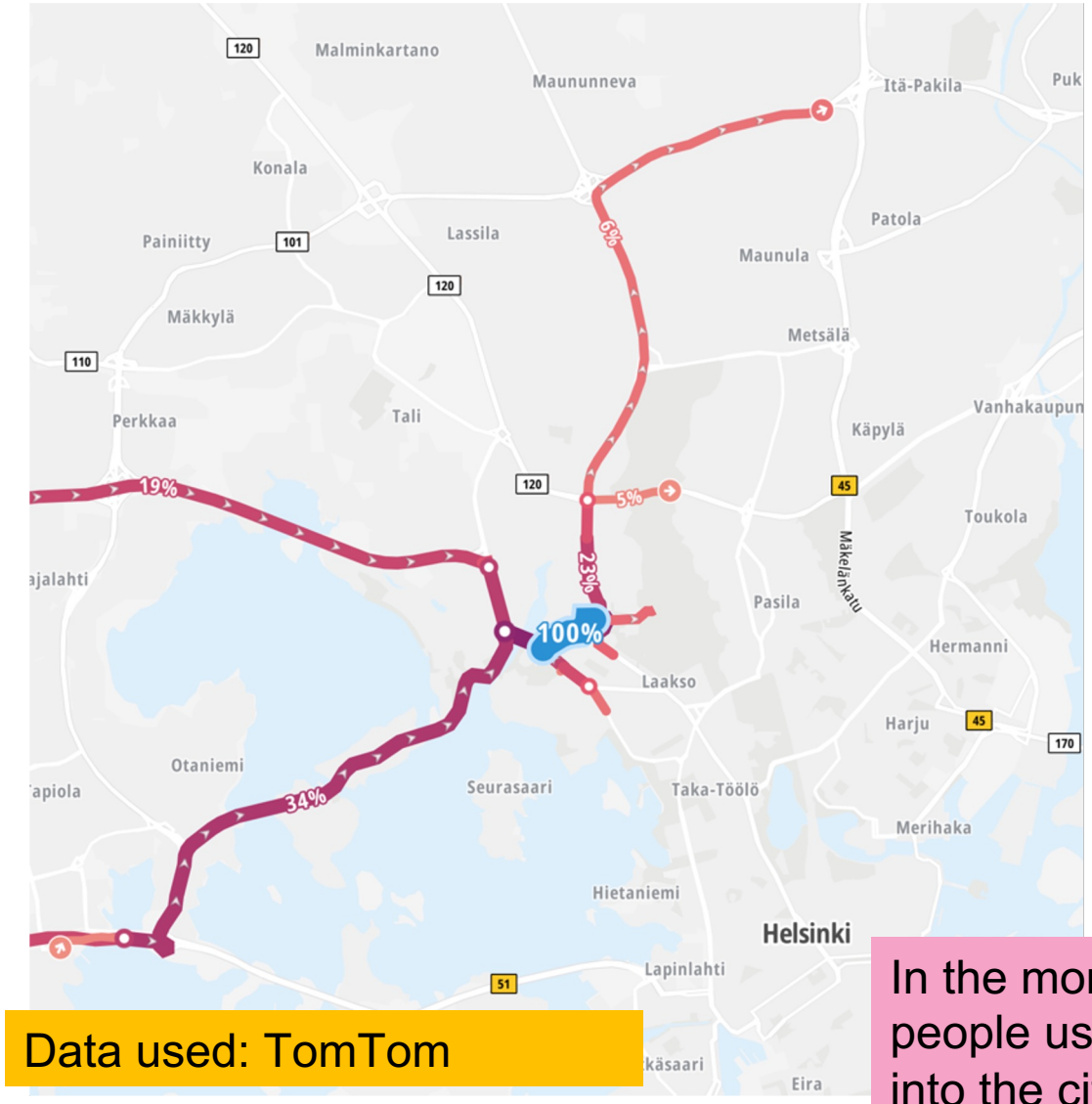


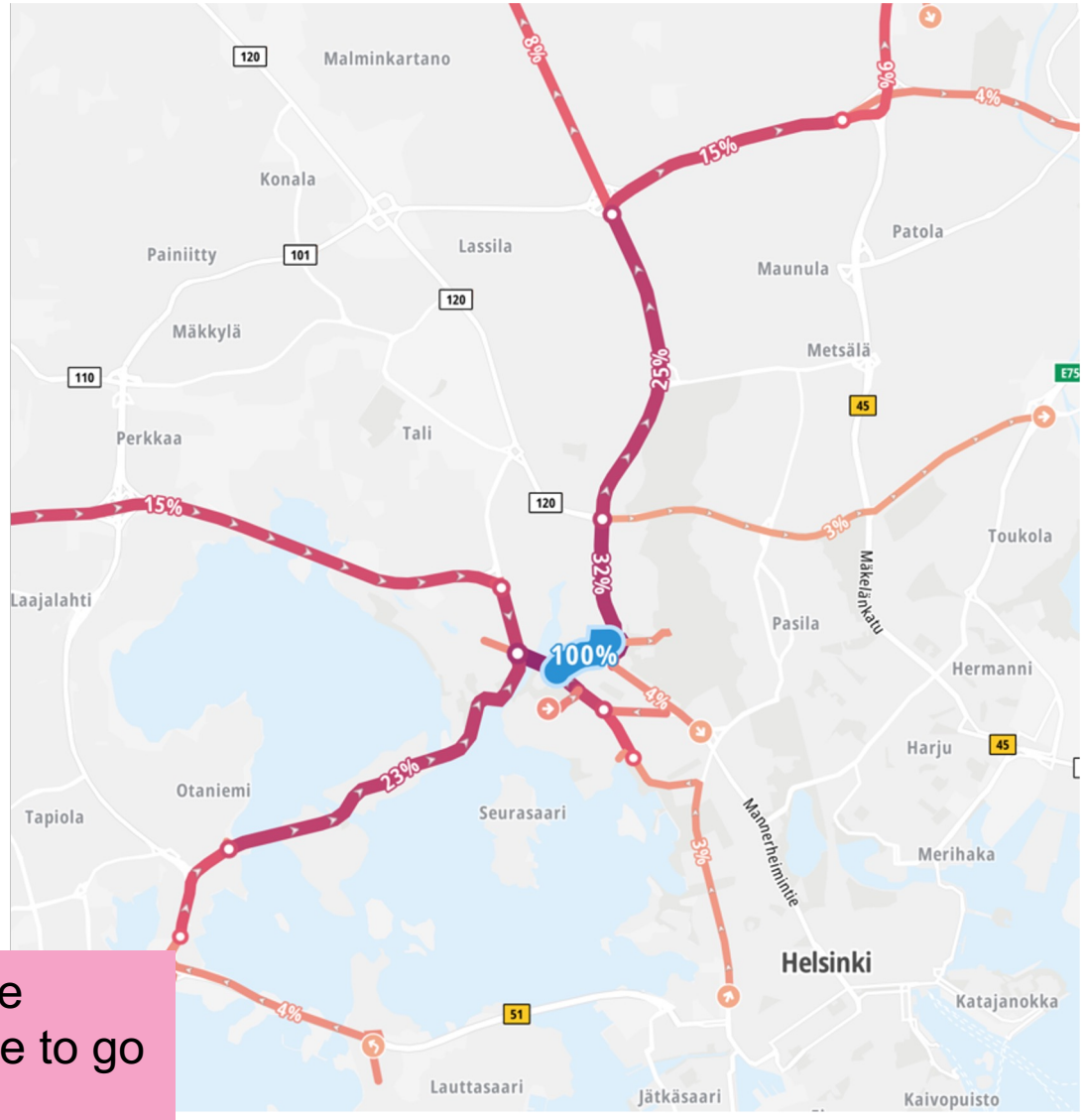
But rat running can occur: In peak hours the shortcut can be up to 4 minutes faster!

Helsinki

Data used: ODIQ

Drive-through on Paciuksenkatu-Mannerheimintie





Data used: TomTom

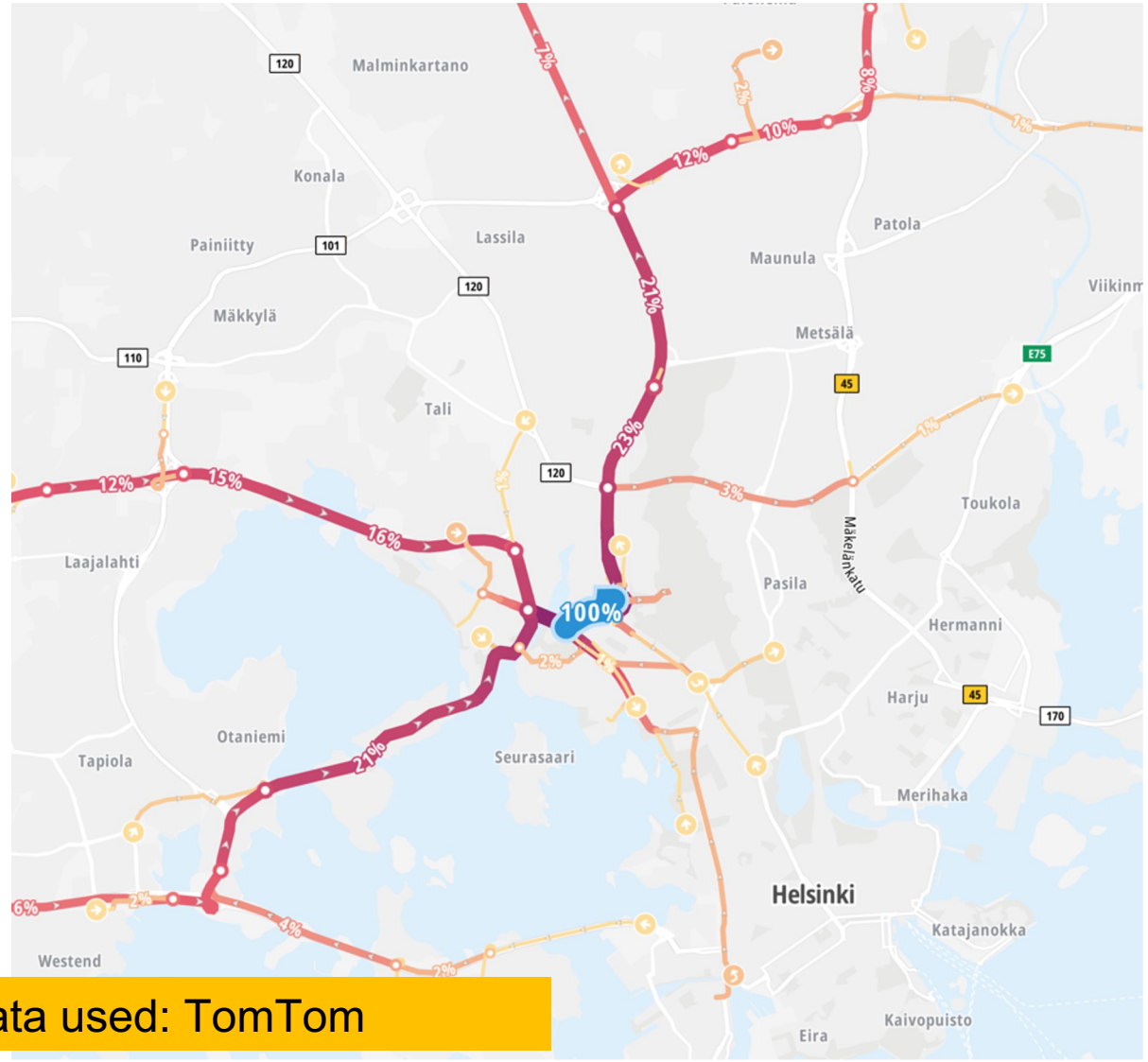
Helsinki

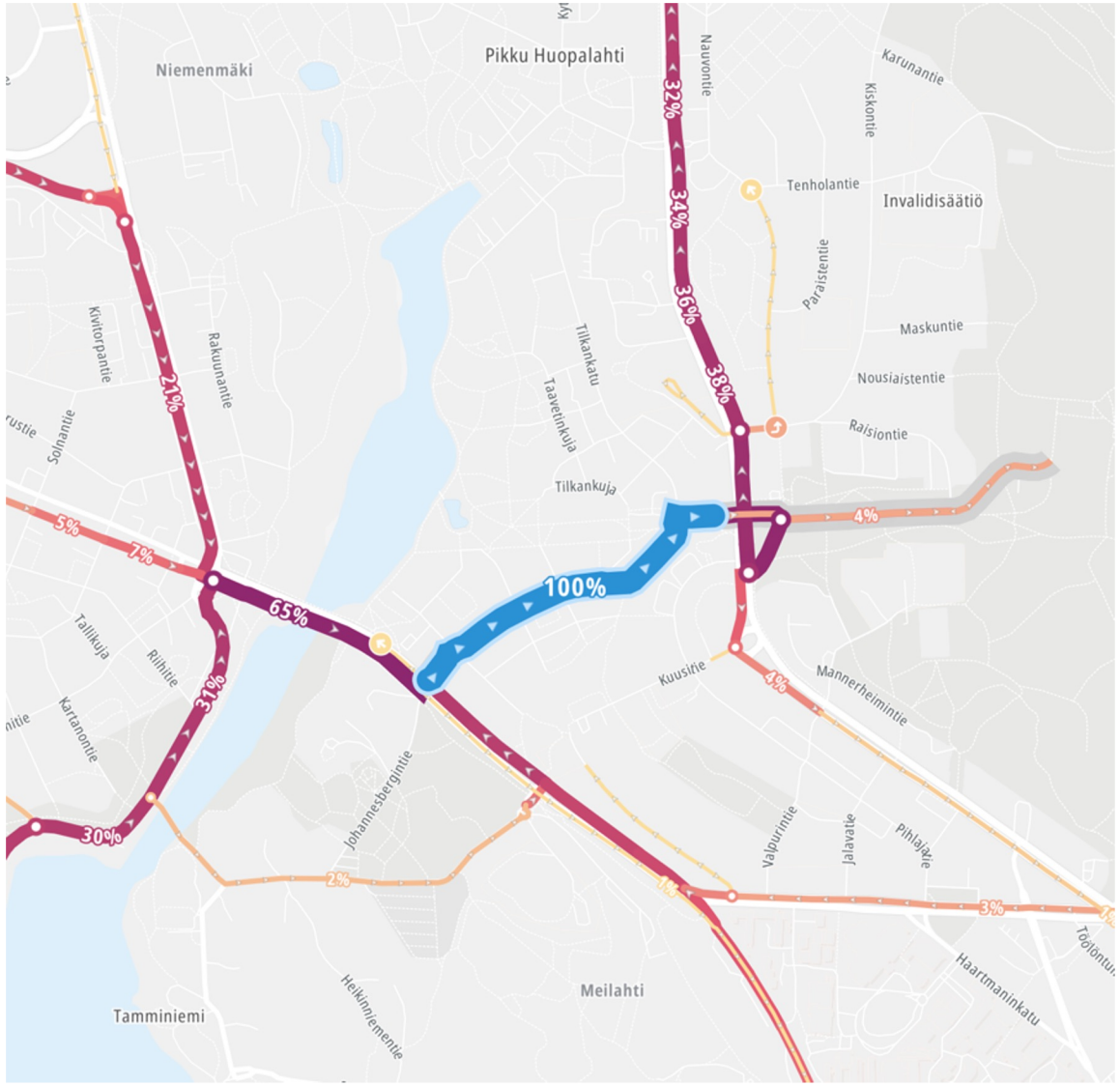
Morning (6-9 am)

In the morning, more people use this route to go into the city.
In the evening, more to go out of the city.

Evening 16-19h

Drive-through on Paciuksenkatu-Mannerheimintie



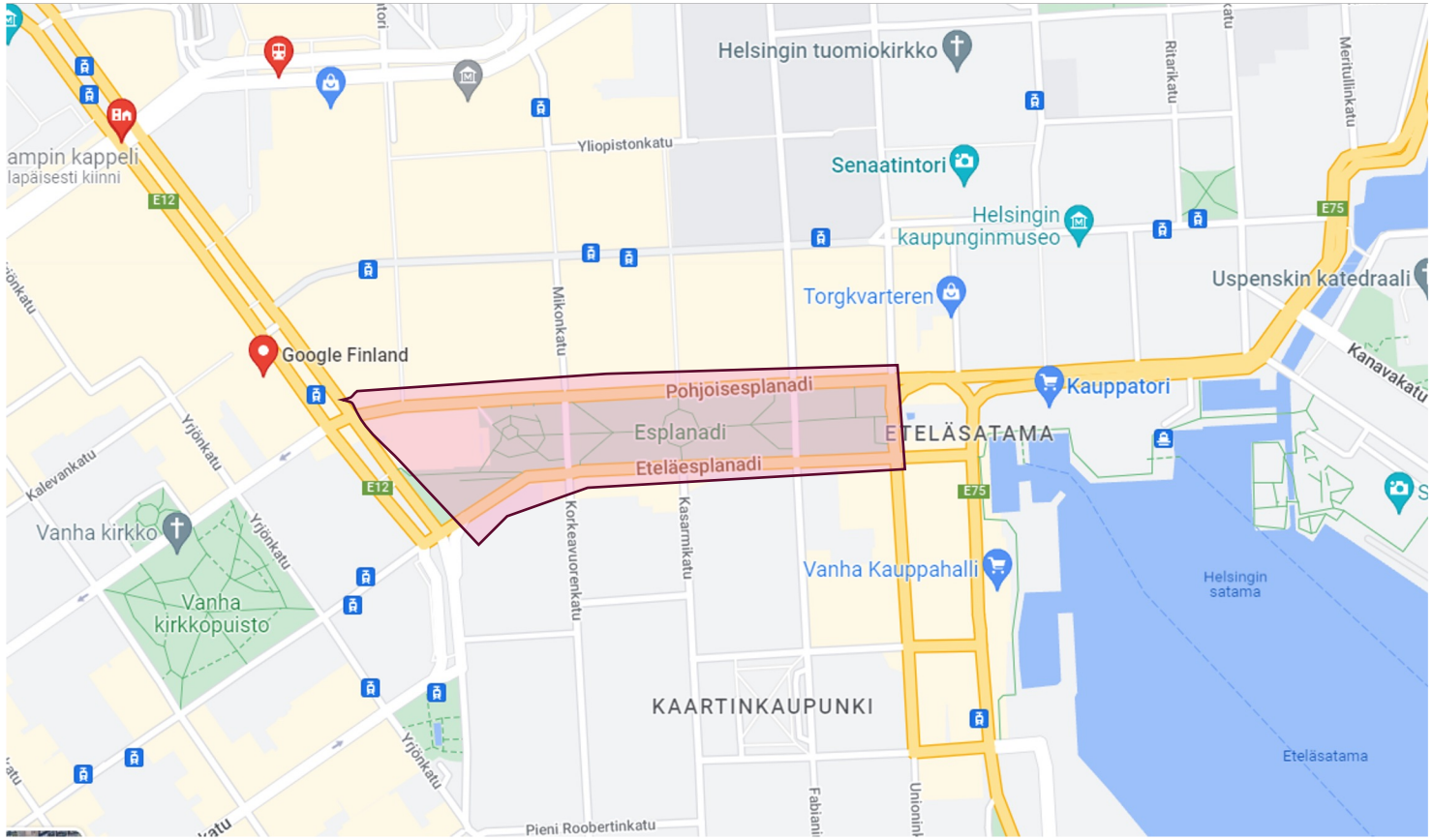


Data used: TomTom

Helsinki

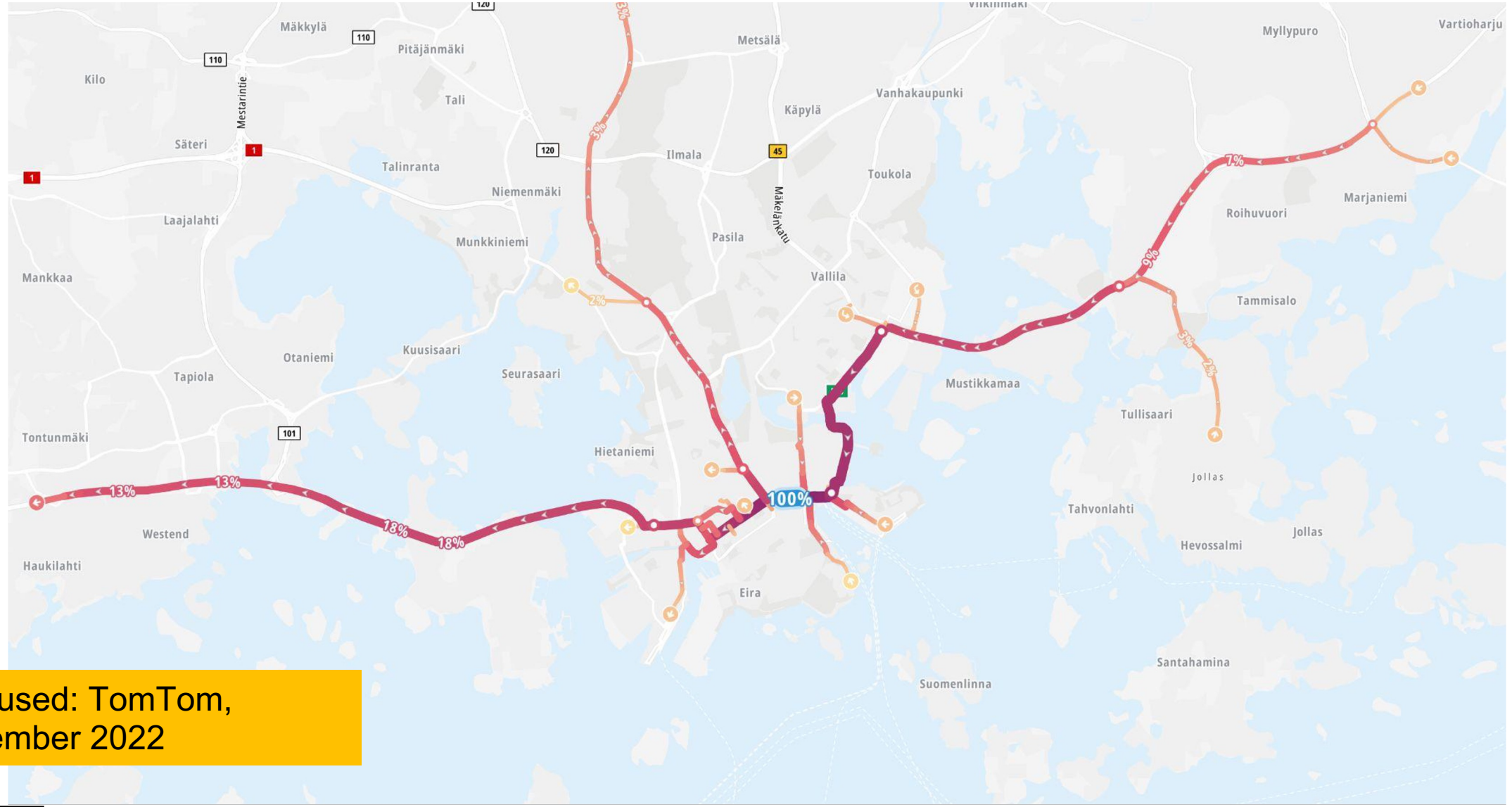
All day

Route analysis: Pohjoisesplanadi and Eteläesplanadi, OD-analysis



Where does the traffic come from and go, that use these two roads?

Pohjoisesplanadi

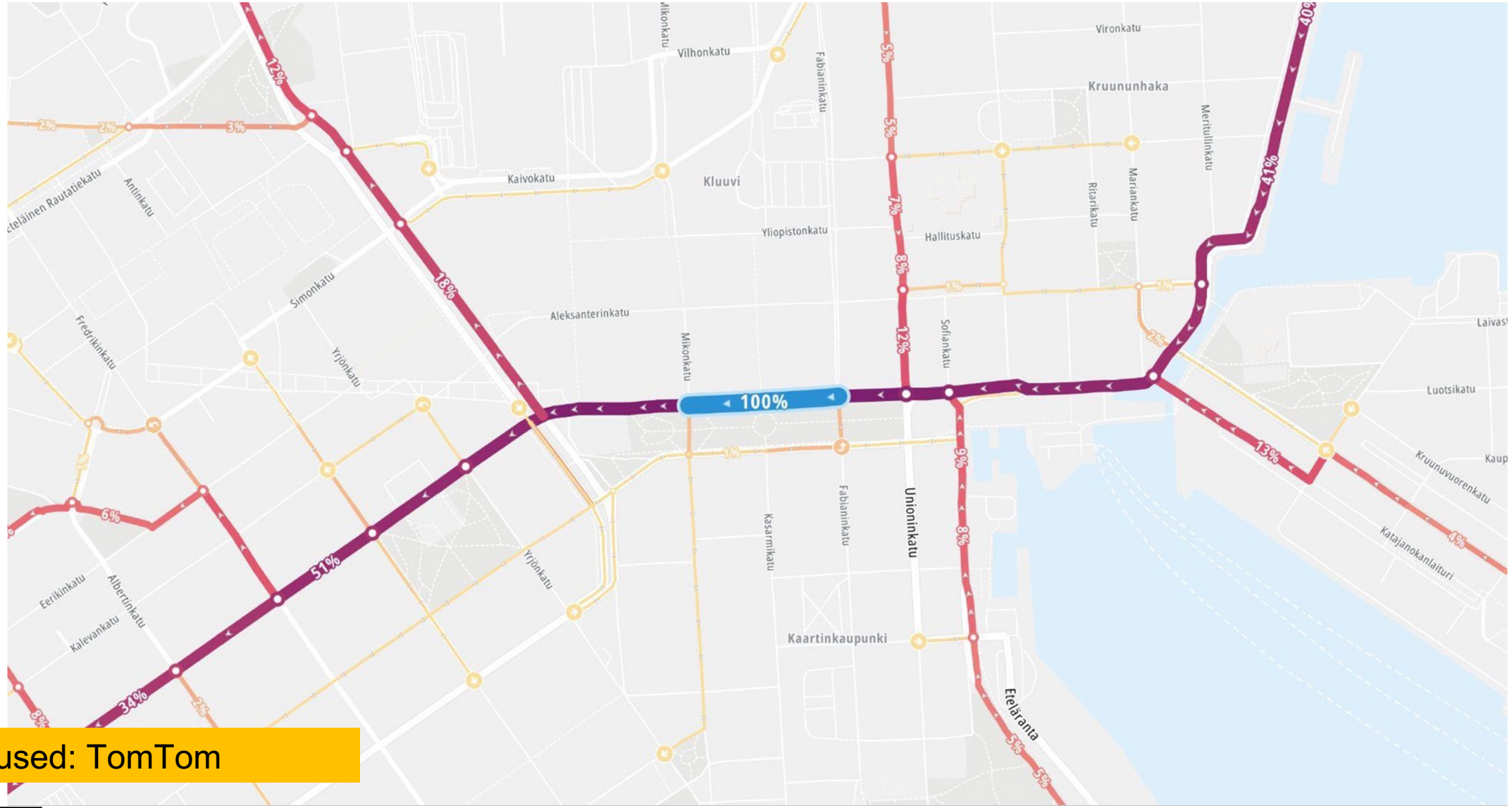


Data used: TomTom, September 2022

Helsinki

All day

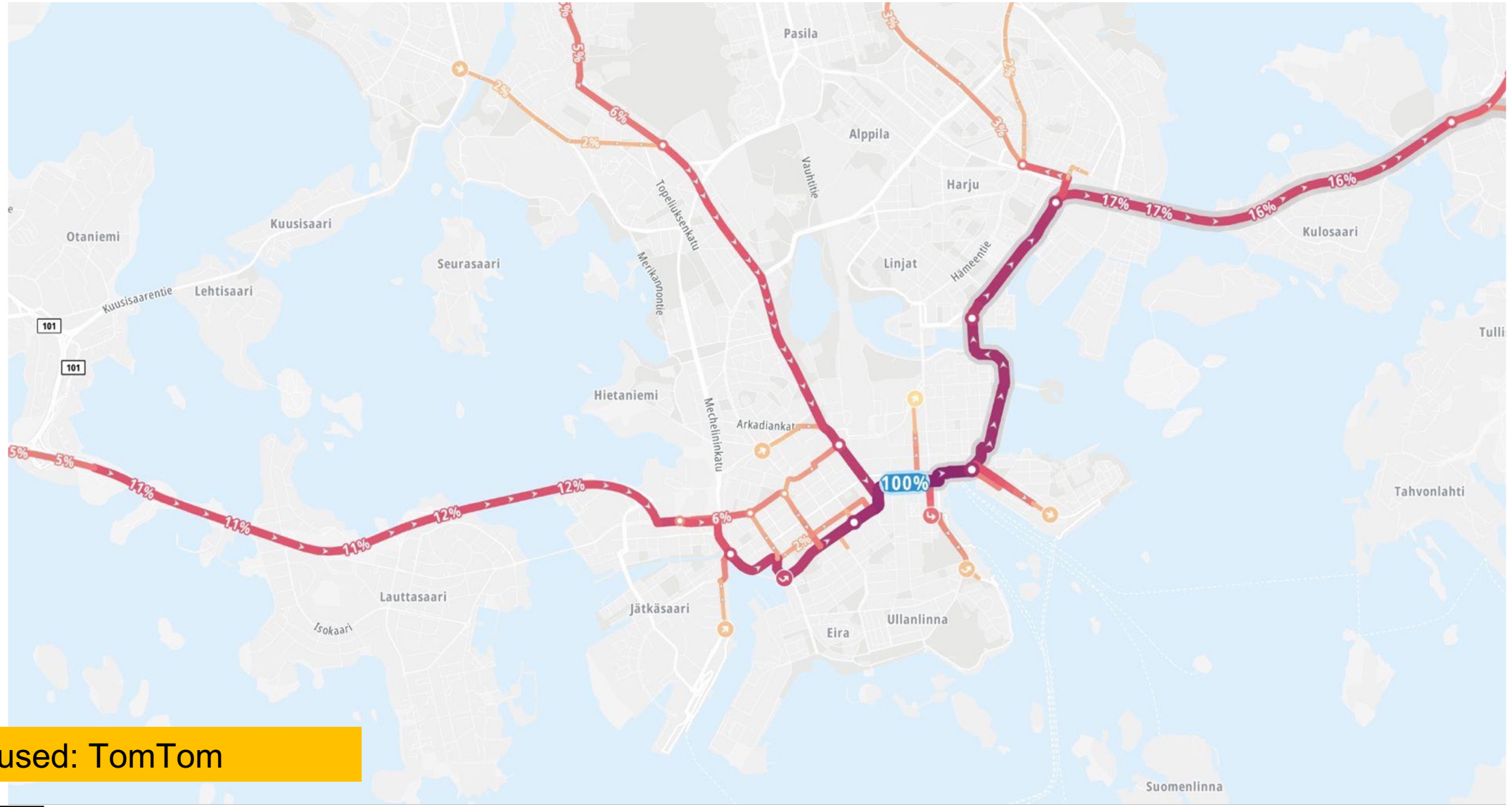
Pohjoisesplanadi



Data used: TomTom

All day

Eteläesplanadi

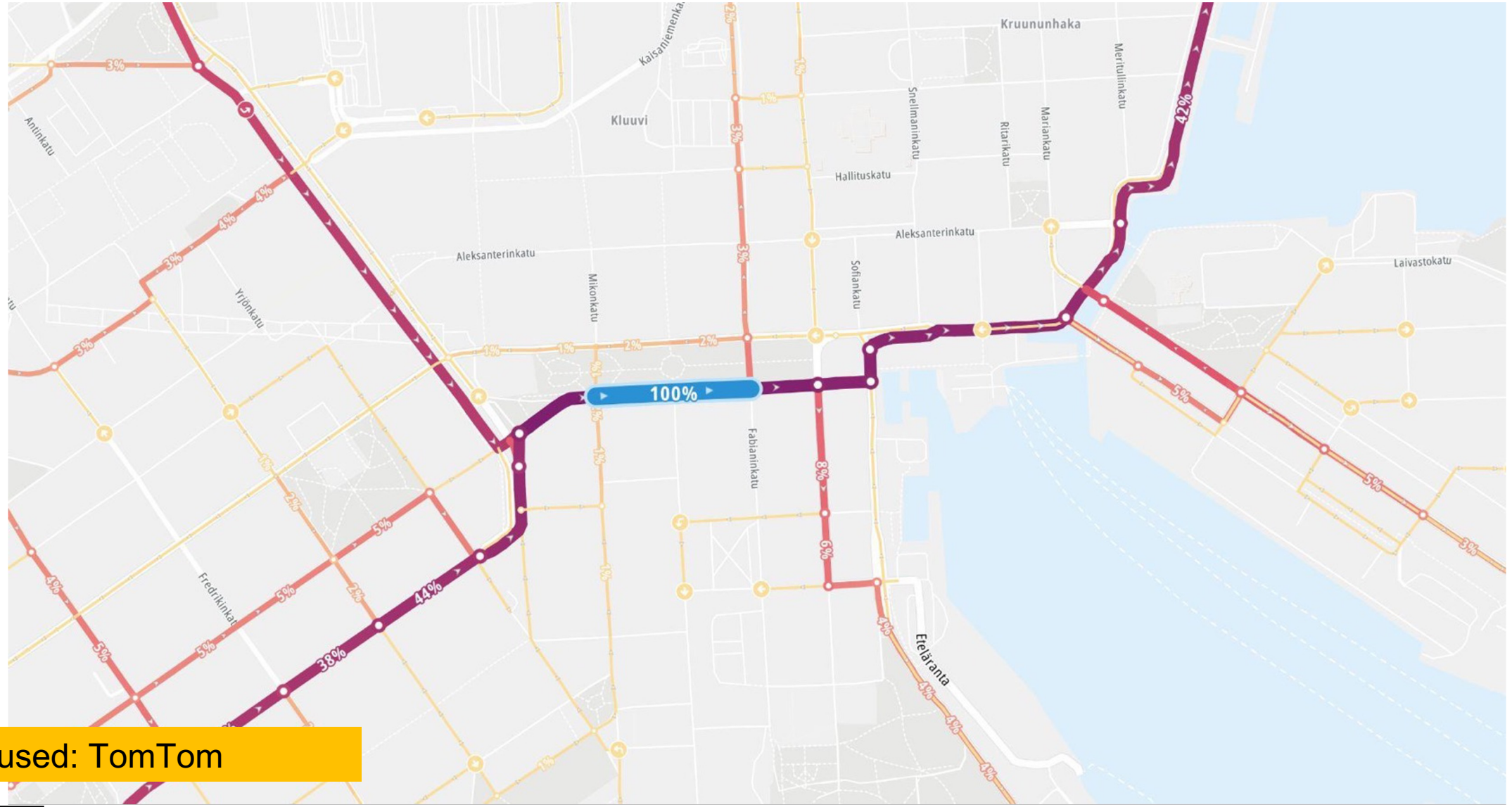


Data used: TomTom

Helsinki

All day

Eteläesplanadi



Data used: TomTom

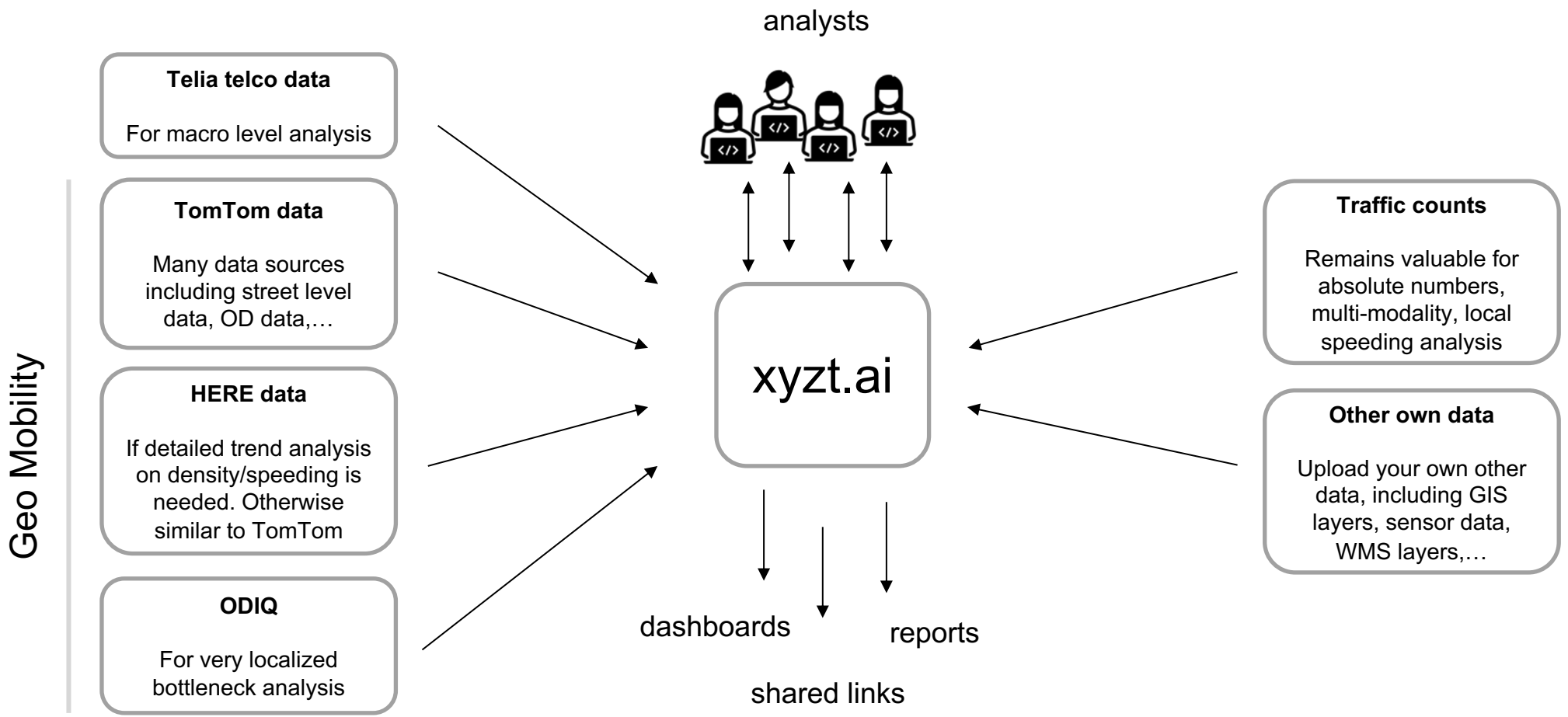
Helsinki

All day

5. Key findings

- Floating vehicle data available at different spatial and temporal aggregation levels
  - Hourly per-road statistics: HERE, ODIQ (Google)
  - Weekly/monthly per-road statistics: TomTom
  - Raw data: Otonomo, Wejo, INRIX, Bridgestone,… (not considered in the project)
- Good coverage in Helsinki area for all used data sources
- Telco data good for macro level analysis, ODIQ for micro level analysis
- Possible specific use-cases for data depend on aggregation, e.g.,
  - Origin-destination
  - Corridor analysis
  - Speeding analysis
  - Traffic density analysis
- Combination with bring-your-own-data such as traffic counting, GIS data,… powerful
- Multi-modality (pedestrians, bicycles, cars, public transport,…) still fragmented and difficult to analyze

Multi-source mobility analytics solution



Contacts

xyzt.ai

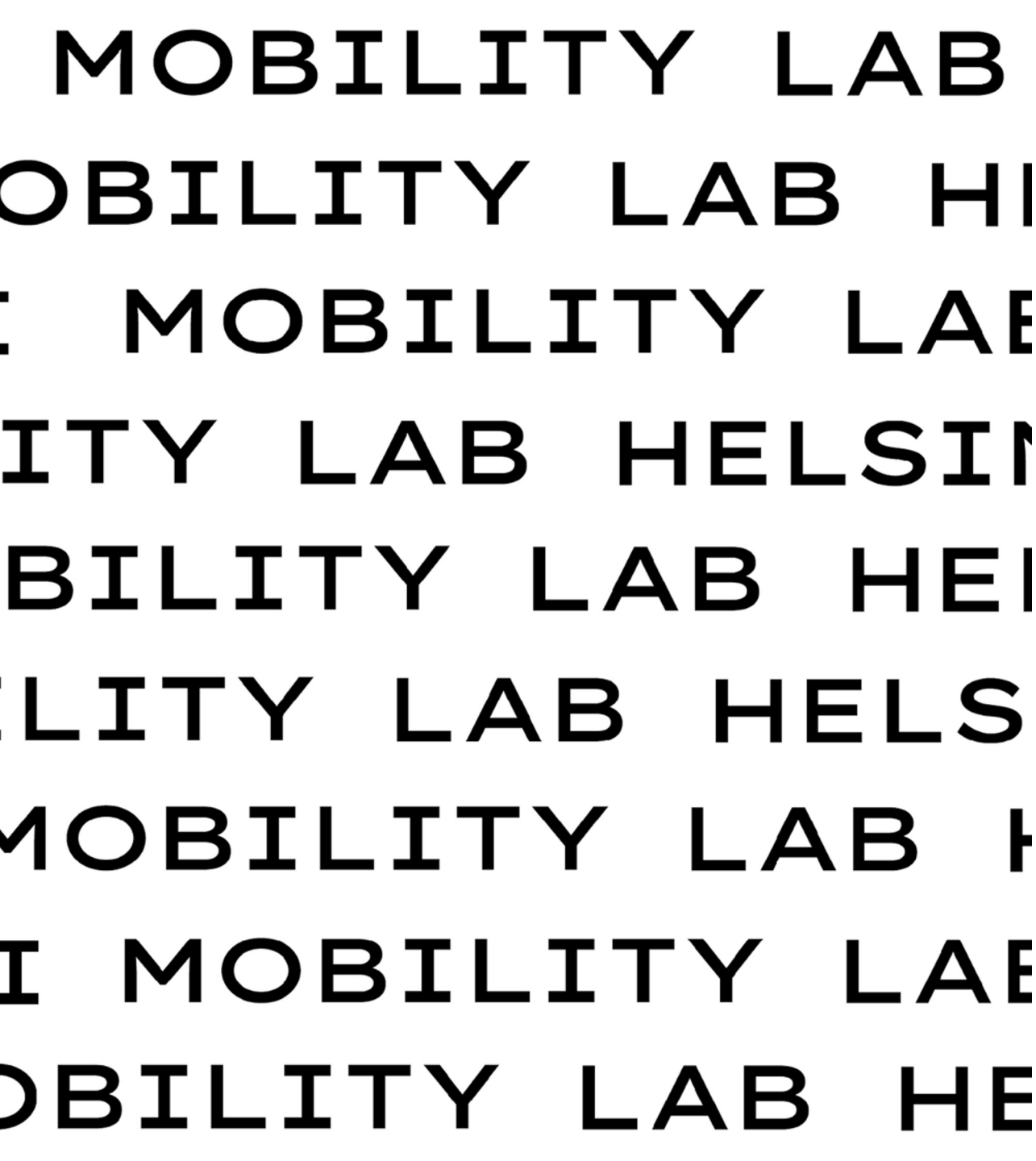
- Nick De Beer
- nick.debeer@xyzt.ai
- www.xyzt.ai

Geo Mobility

- Michiel Van Hove
- michiel@geomobility.eu
- https://www.geomobility.eu/en/home

Mobility Lab Helsinki

- Juho Kostiainen
- juho.kostiainen@hel.fi
- https://mobilitylab.hel.fi/



Juho Kostiainen
Project Manager
City of Helsinki
Juho.Kostiainen@hel.fi
+358 9 310 365 35

Janne Rinne
Project Manager
Forum Virium Helsinki
janne.rinne@forumvirium.fi
+358 41 431 0822