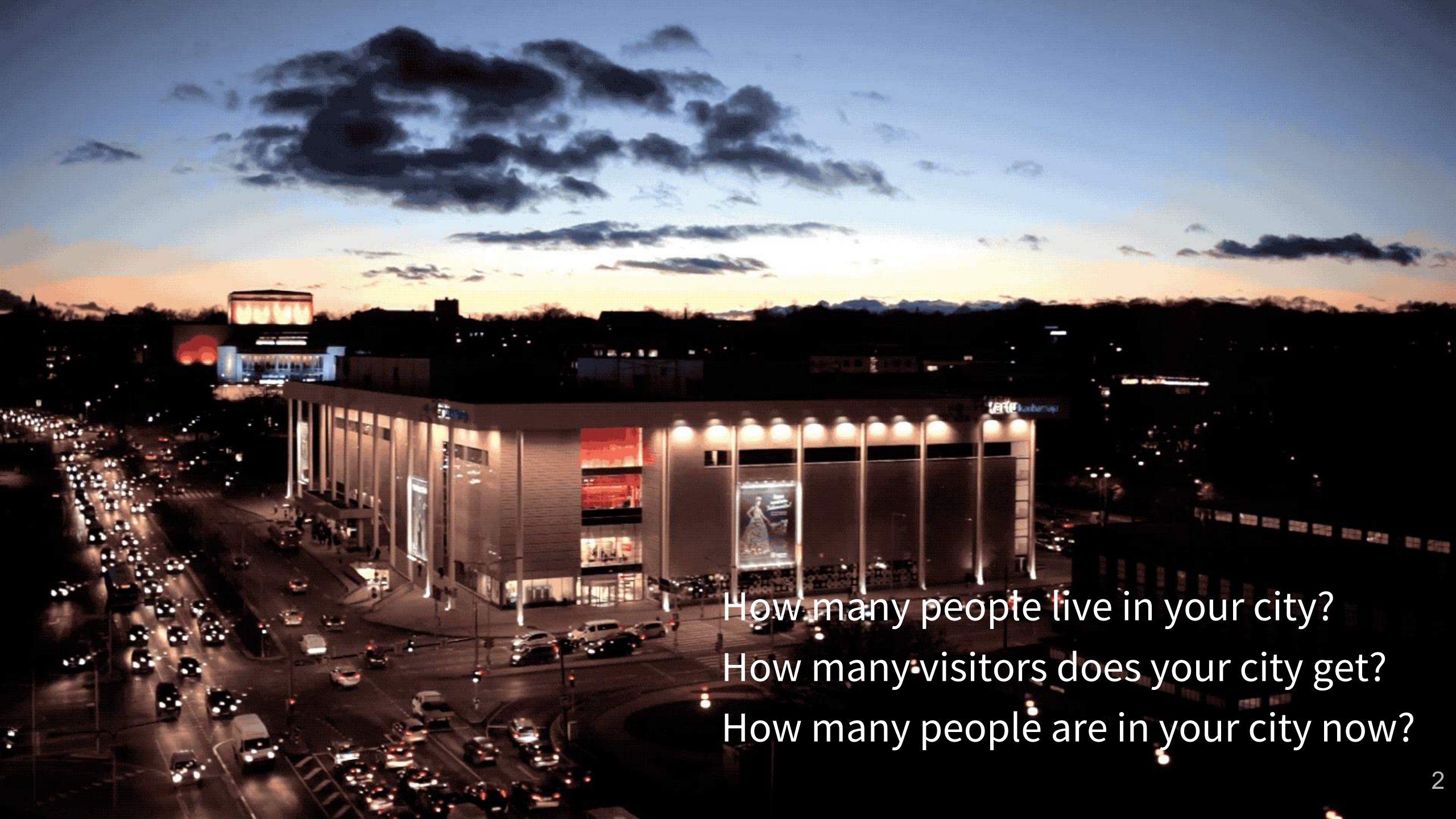


# Mobile Phone Data (MPD) for Dynamic Population Mapping

UN Committee of Experts on Big Data and Data Science for Official Statistics has prepared guiding materials for the use of mobile positioning data for dynamic population mapping for the statistical community

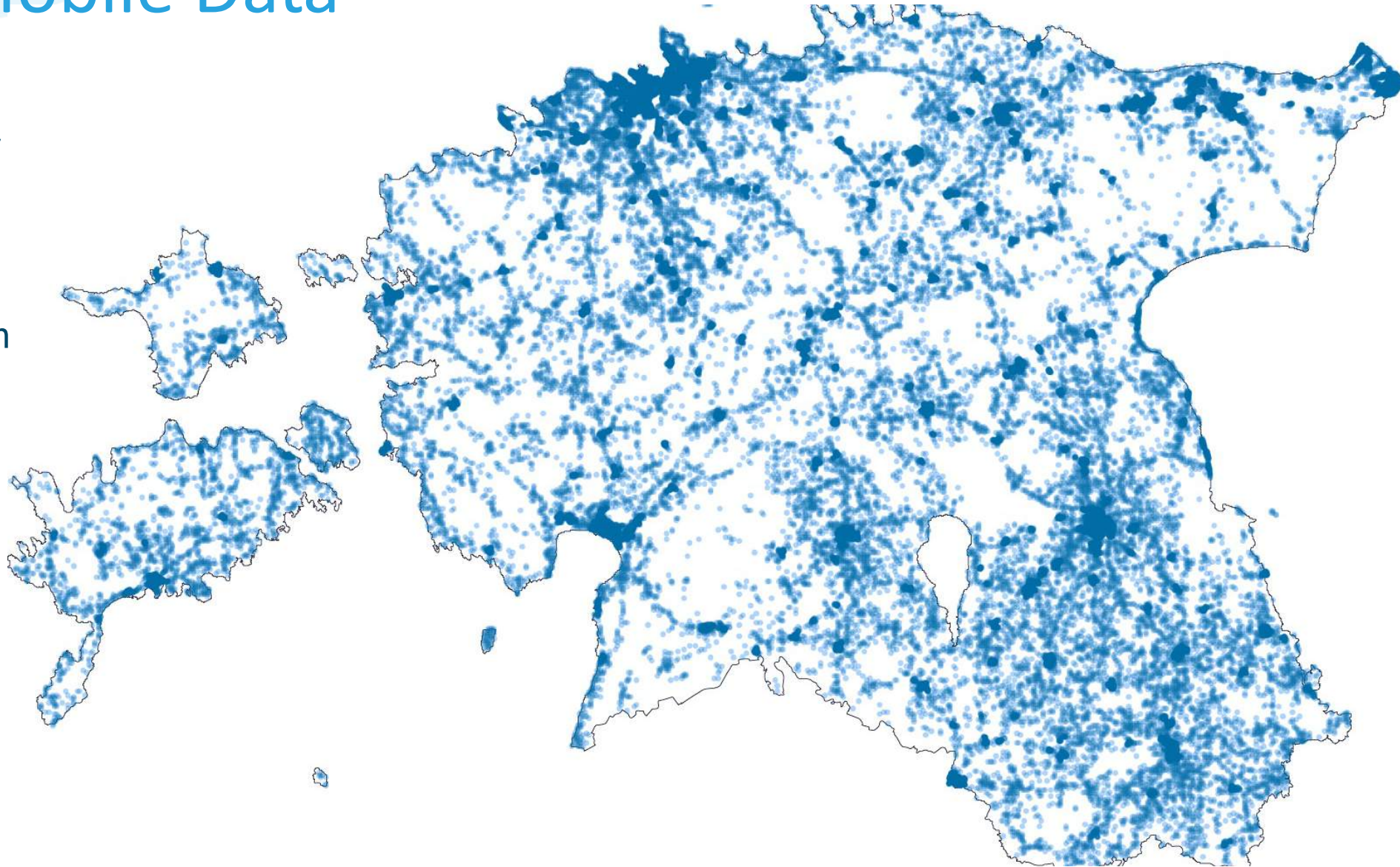
Presenter: Kaisa Vent, Positium, Estonia



How many people live in your city?  
How many visitors does your city get?  
How many people are in your city now?

# 1 Day of Mobile Data

Here is how one day of domestic and inbound data looks like in Estonia. 18M location events from call detail records



# 9 Ways Mobile Phone Data can be used for Dynamic Population Mapping

Anonymous data used, not individuals are tracked, results are presented as aggregated and extrapolated statistical indicators.

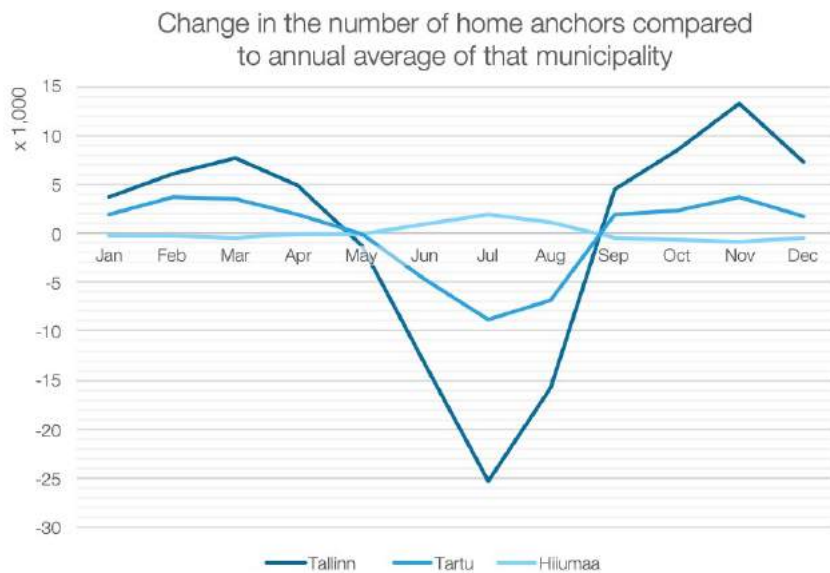
# In the Handbook

1. Resident population mapping
2. Daytime population mapping
3. De facto population mapping
4. Monitoring population redistributions caused by COVID-19 mobility restrictions
5. Infrastructure and resource planning
6. Creating dynamic sample frames for surveys
7. Census
8. Disaster preparedness planning and response
9. Dynamic population denominators to estimate disease incidence

# 1. Resident Population

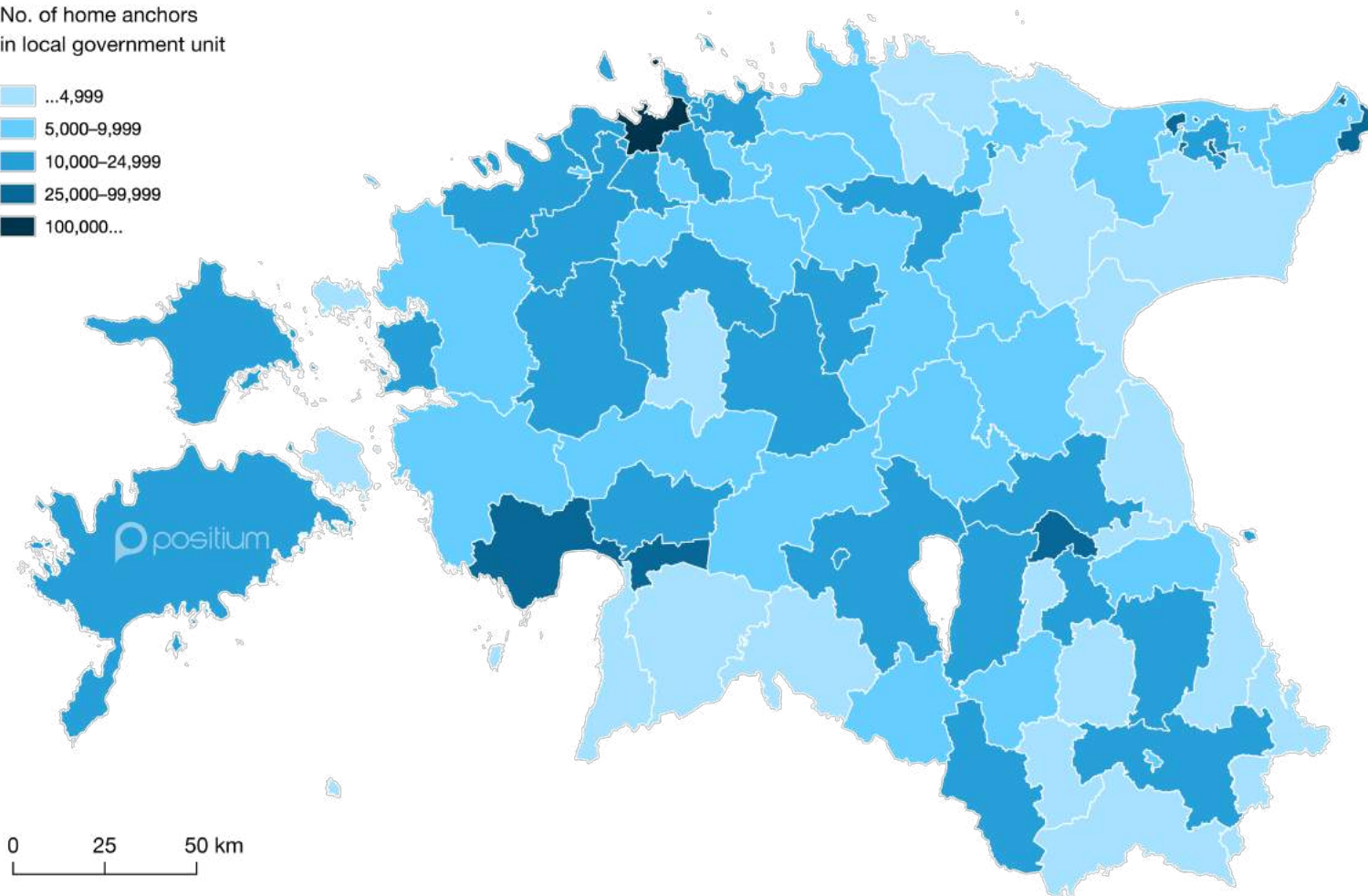
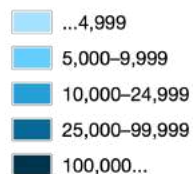
## Number of residents in local government units

When number of residents falls in Tallinn and Tartu over summer, it increases on the islands, such as Hiiumaa. In Tartu, there are many university students, who leave during summer.



Annual average number of home anchors: Tallinn 423,204; Tartu 96,133; Hiiumaa 11,320

No. of home anchors in local government unit



## 2. Daytime Population

Mobile positioning data



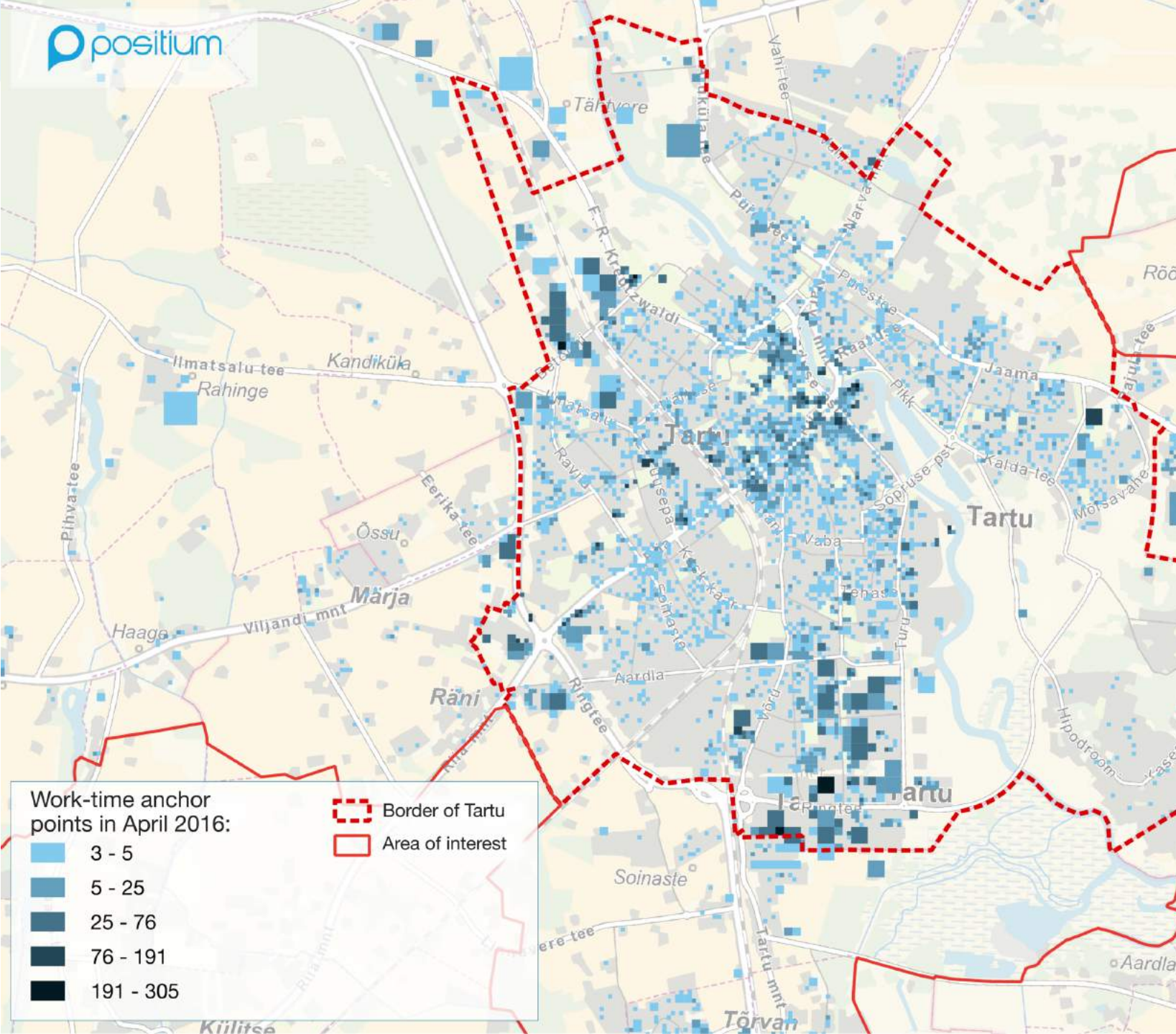
Cleaned and processed



Work-time anchors detected



Aggregated to adaptive grid



Work-time anchor points in April 2016:

- 3 - 5
- 5 - 25
- 25 - 76
- 76 - 191
- 191 - 305

- Border of Tartu
- Area of interest

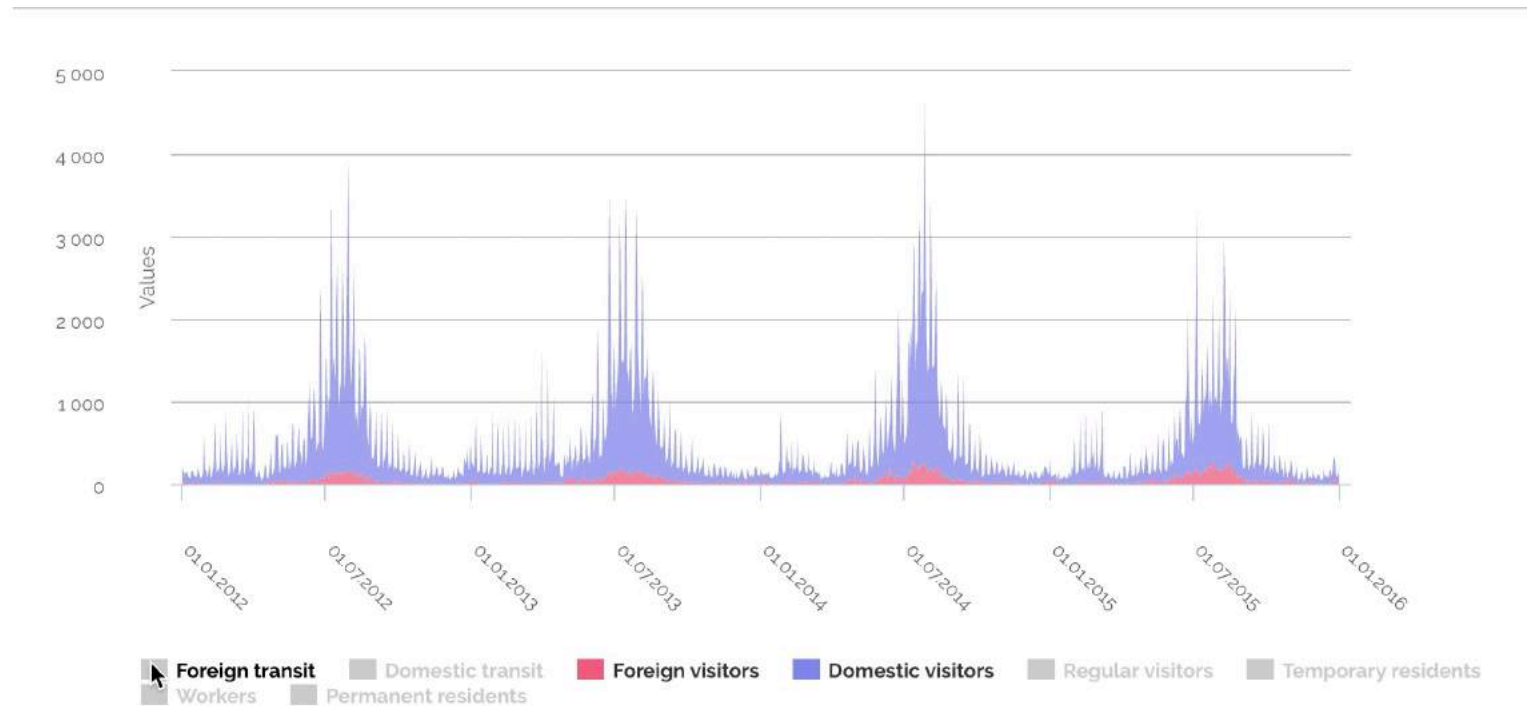
# 3. De facto Population

The number of people present in specific municipality during a period of time (example: day)

- Shows population dynamics
- Shows who are residents, temporary residents, visitors and others who stay in an area.

Already used by:

- national and local governments for public services planning
- law enforcement to allocate the resources based on forecast models
- disaster management preparedness in table top exercises

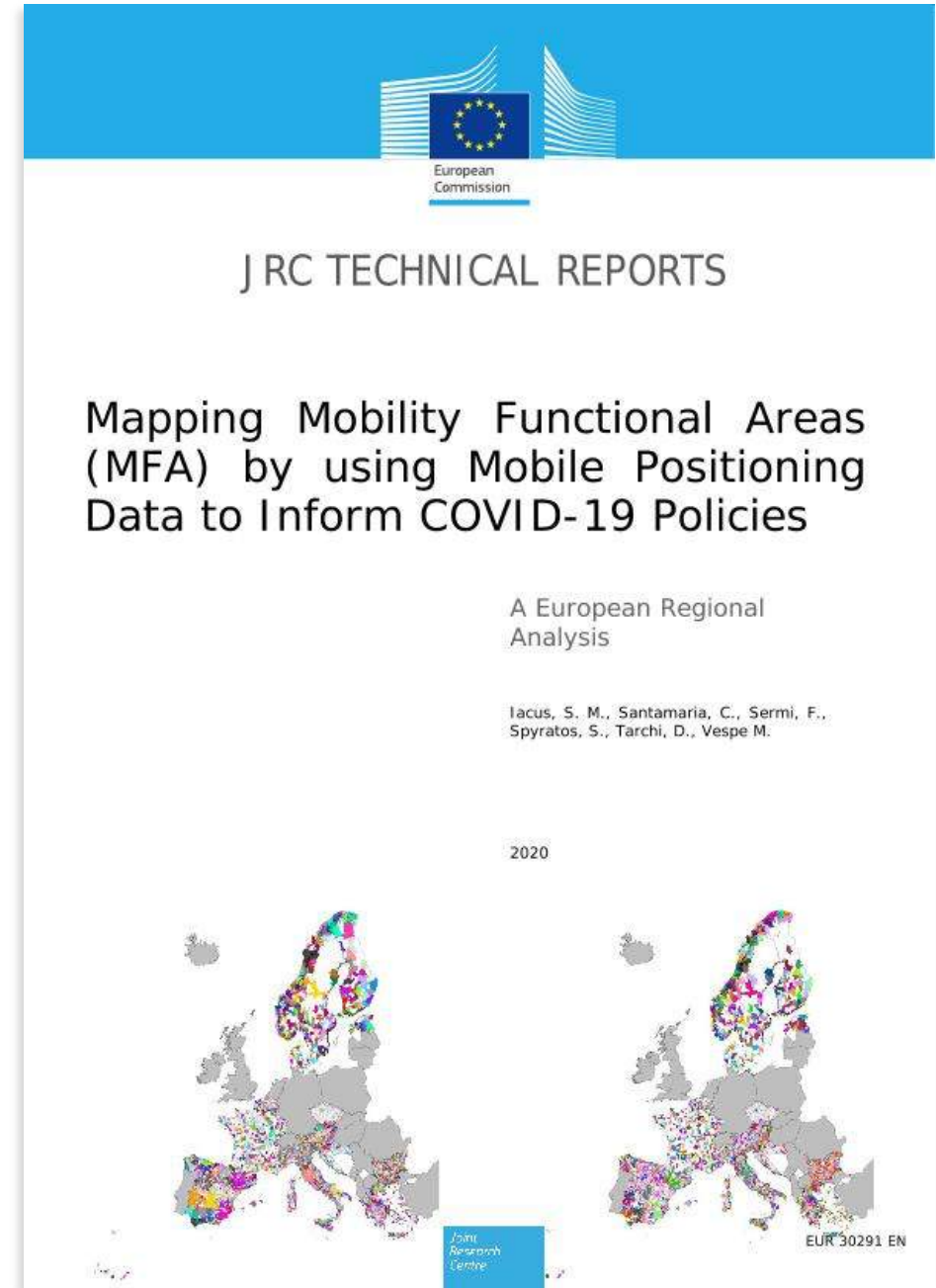




# 4. COVID-19 population mobility

During the COVID-19 pandemic, mobile phone data analysis was conducted in several countries around the world to provide insight into adherence to mobility restrictions and change in the residents (urban-rural movement)

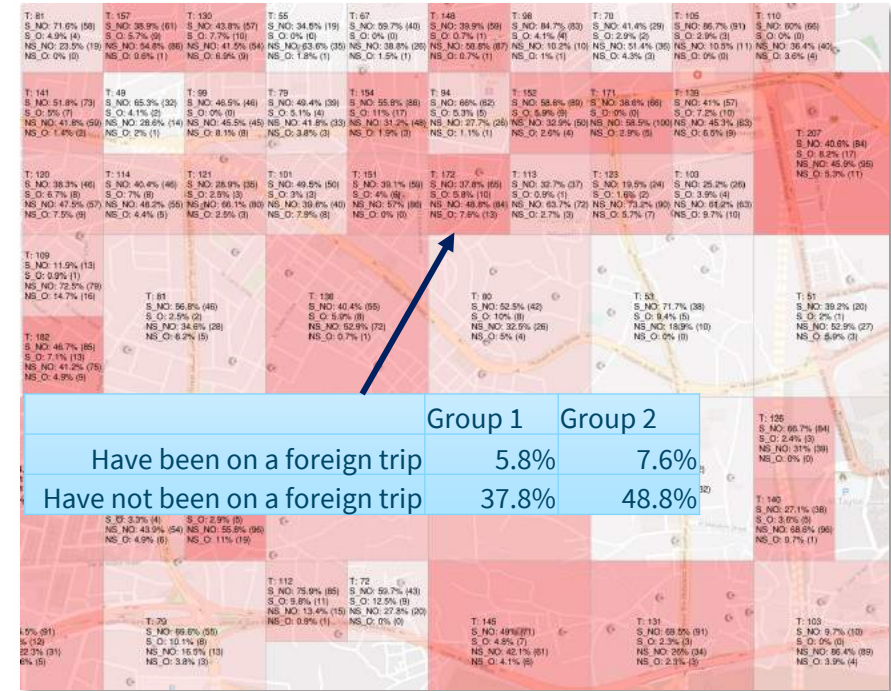
Example work from European Commission based on aggregated mobile phone data from 16 EU countries  
<https://ec.europa.eu/jrc/en/publication/human-mobility-and-covid-19-initial-dynamics>



# 5. Creating dynamic sample frames for surveys

Mobile positioning data can be used to create sampling frames in the following ways to complement surveys where:

- Reliable information does not exist
- Population data needs validation ([UNESCAP, 2019](#))
- Physical canvassing is difficult (conflict, pandemic)
- Extension of sample frame formation to specific circumstances – targeting subgroups of certain mobility pattern, e.g. nomadic, seasonal, tourism, migration.



Example use of gridded population for sampling frame generation for tourism trips

## 6. Census

Opportunities exist to complement the census during:

- **Census preparation and execution** (establishing survey frames)
- **Intercensal period** (population projections)
- **Measuring possible undercounted populations** (nomadic populations)
- **Adding features** (daytime population, migration or commuting)



# Methodological challenges

And some methods to overcome them



# Main Methodological Questions



How to build the right data model?



How to detect place of residence and select the right home detection algorithms?



How to ensure data coverage and representativity?



How to model population density?



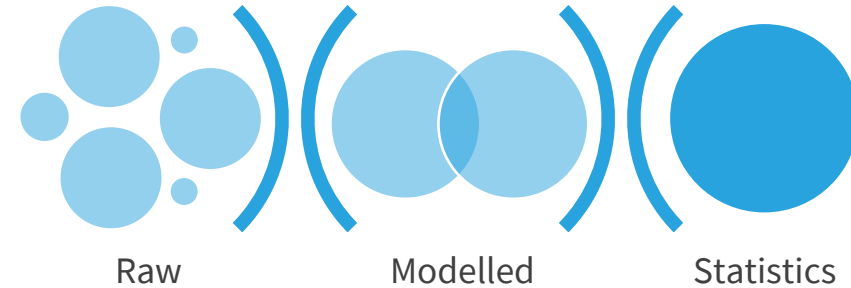
How to validate the results?

# Choice of model



## Simplified model

- Statistical calculation done directly on raw data
  - E.g. simple SIM card counts or tower density measurements in a grid
- Quick indicators
- Modelling done post-processing
- Not suitable for official statistics



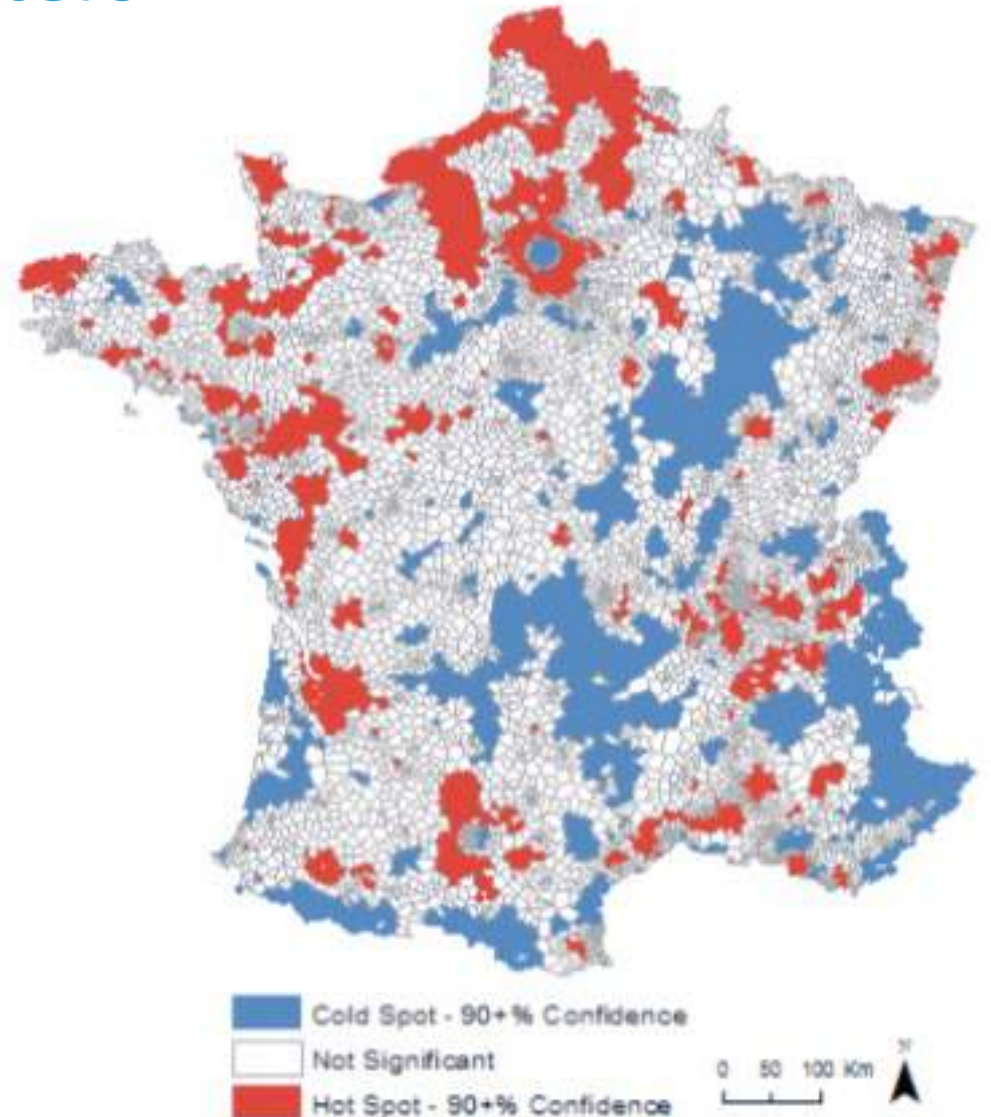
## General model

- A model of reality is built for each subscriber
- Data model matches official definitions
- Statistical concepts are applied in late data processing
- Allows combining and comparing results for many domains
- Most useful for official statistics

# Place of Residence Algorithm Matters

**Choice of criteria in home detection algorithms influences the results significantly**  
- up to 40% in France.

[Vanhoof, M et al \(2018\) Assessing the quality of home detection from mobile phone data for official statistics](#)  
Based on validation study where the results of different home detection algorithms were compared to census counts at cell tower level



# Validation of home detection

Validation through:

## 1. Aggregate comparison

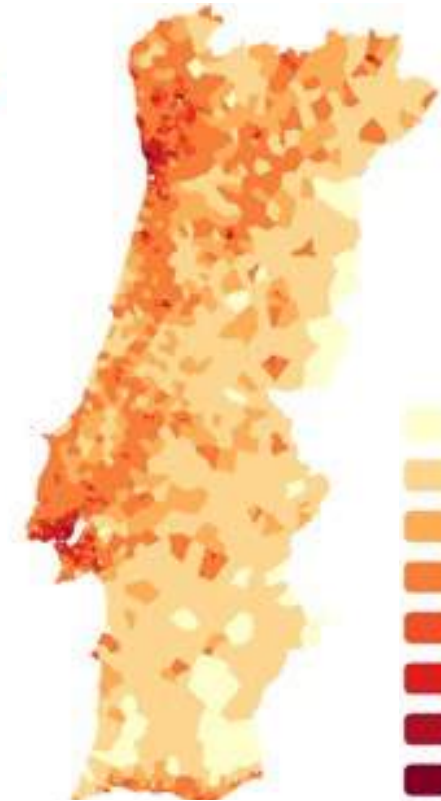
[Dynamic population mapping using mobile phone data](#). Deville et al (2014)

R = 0.89 for municipality level

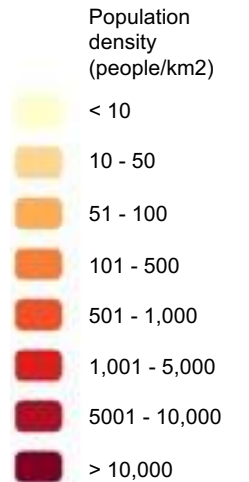
Portugal



Census



MPD



Lisbon





# Validation of home detection

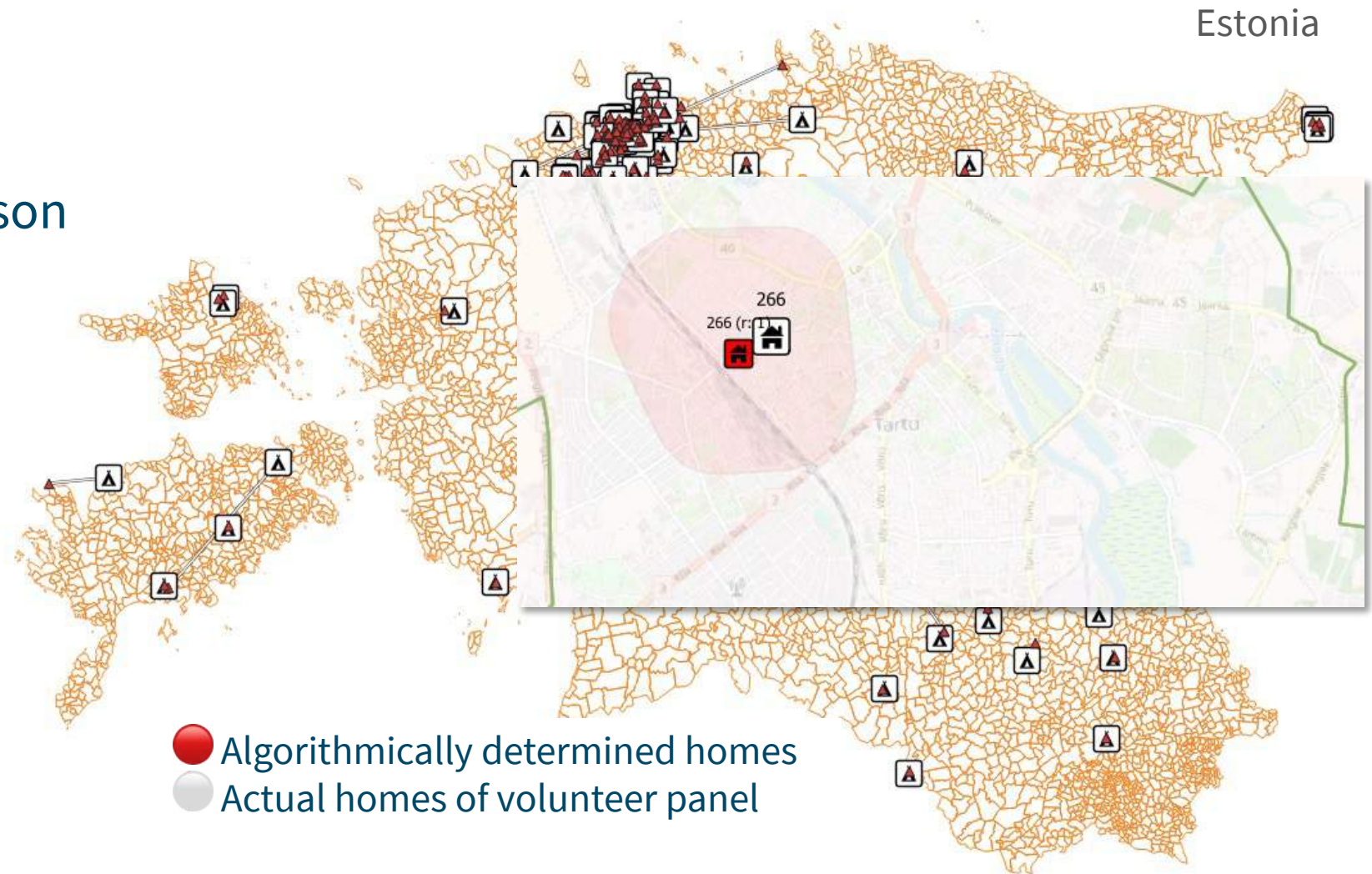
Validation through:

1. Aggregate comparison
- 2. Validation panels**

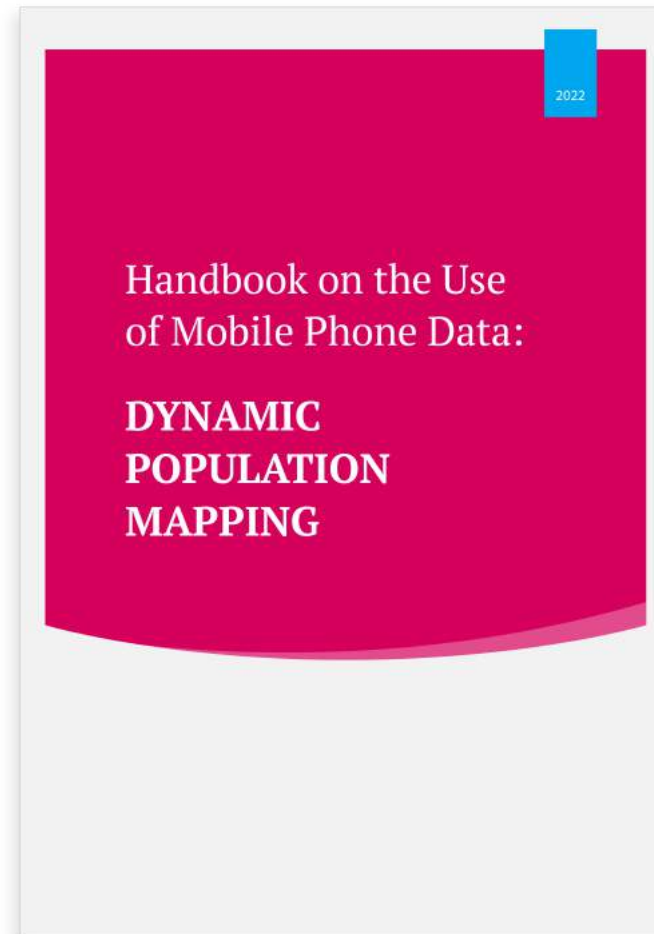
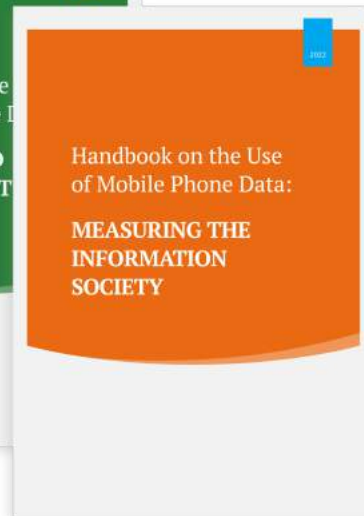
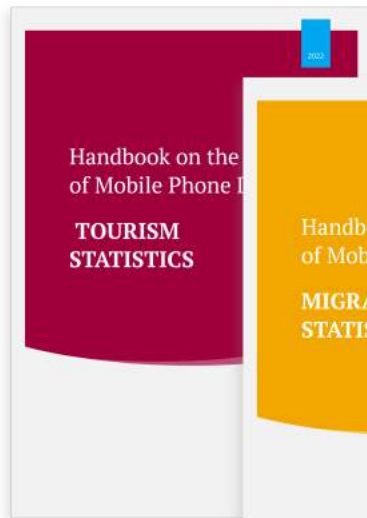
\* [Ahas et al \(2010\) Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones](#)

\*\* Based on internal validation study done with Statistics Estonia, where comparison was done between the MPD anchors calculated by Positium, population registry data and volunteers' true home address

R=0.99 for county level, R=0.94 for municipality level



# Handbook on Dynamic Population Mapping



# Thank You!

Questions?

[kaisa.vent@positium.com](mailto:kaisa.vent@positium.com)