WISDOM OF CROWDS – DATA SET DEFINITIONS

Footfall data

Footfall data captures the amount of people that passed through the grid during an hour, based on routing the people movement to the street network.

Figure 1. Example visualization of the footfall data (N.B. grid will be aligned with the Statistics Finland grid)

Data set

- Area: Uusimaa region
- Geographical resolution: 250 x 250 m YKR grid
- Time period: 1.1.-31.3.2018
- Time resolution: Hourly
- File size: ~32 MB per day, ~2900 MB per 3 months

Data format

- count: amount of people passing through a grid cell (int)
- grid_id: unique identifier of grid cell (int)
- time: date and hour, where time denotes the start of the hour in Finnish time (d.m.y h.m.s)

count	grid_id	time
15935	5379281	23.1.2018 16.00.00
14342	5379280	23.1.2018 16.00.00
14092	5379281	23.1.2018 15.00.00
…	…	…

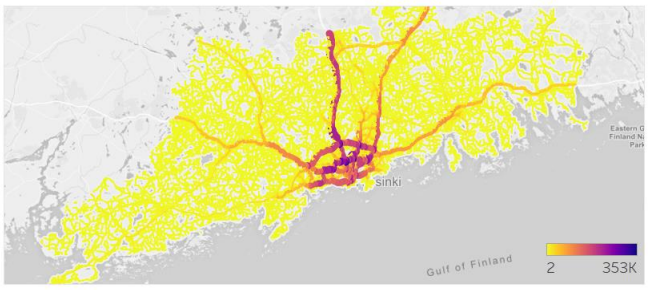
Joining with shapefile

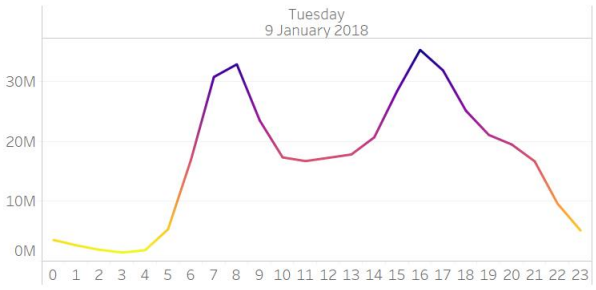
Column 'grid_id' is the id of the footfall grid which can be joined with the 'id' column in the shapefile.

Junction data package example

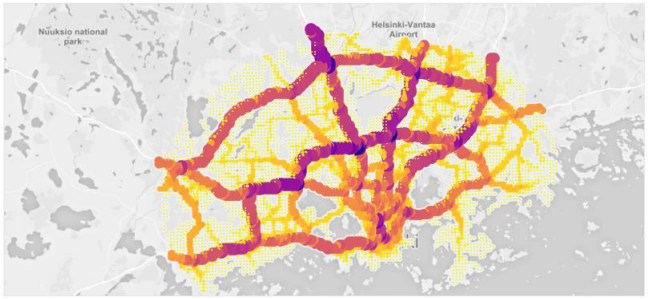
Footfall data coverage in Uusimaa and data set example from Jan 9th 2018.

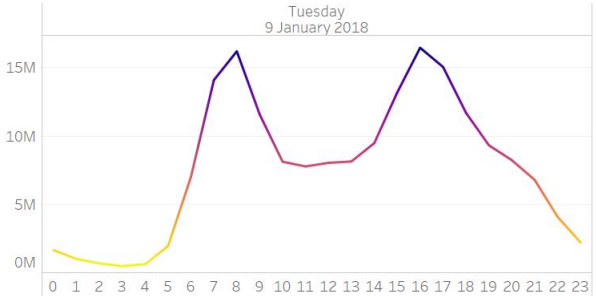
HOURLY FOOTFALL IN UUSIMAA PROVINCE ON JAN 9th 2018





HOURLY FOOTFALL IN HELSINKI CAPITAL REGION ON JAN 9th 2018





Activity data

Activity data reports the number of people dwelling in area for a certain amount of time. More precisely, activity count calculates the number of unique persons with an activity per spatial and temporal unit. Here, the spatial unit is one Telia MTC grid cell, and time unit is an hour. An activity is identified if person is dwelling at least 20 minutes in the same area. As a consequence, counts refer to distinct persons only within a time slot and MTC grid cell.

Let's clarify this with examples:

- A person having an activity from 08:00 to 08:25 in grid cell 10001, and another activity from 08:35 to 08:57 in grid cell 10002 will be calculated once per both grid cells
- A person having an activity from 09:55 to 10:20 in grid cell 10003 will be calculated once for hour 09-10, and once for hour 10-11.

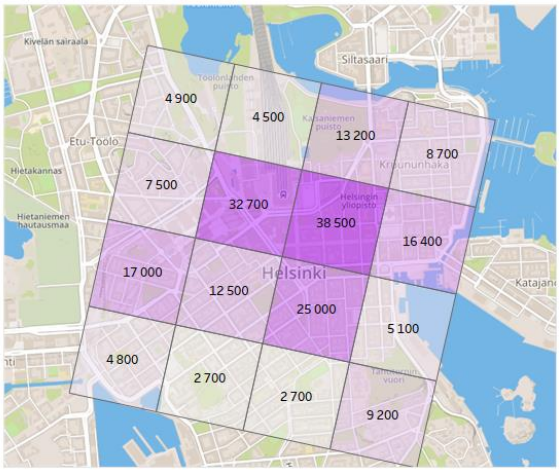


Figure 2. Example visualization of the activity data (N.B. grid will be aligned with the Statistics Finland grid)

Data set

- Area: Uusimaa region
- Geographical resolution: Telia MTC grid aligned with Statistics Finland grid
- Time period: 1.1.-30.6.2018
- Time resolution: Hourly
- Data volume: ~1,5 MB per day, ~270 MB per six months

Data format

- dominant_zone: unique identifier of MTC grid cell (int)
- count: unique people dwelling in an area for at least 20 minutes (int)
- time: date and hour, where time denotes the start of the hour in Finnish time (d.m.y h.m.s)

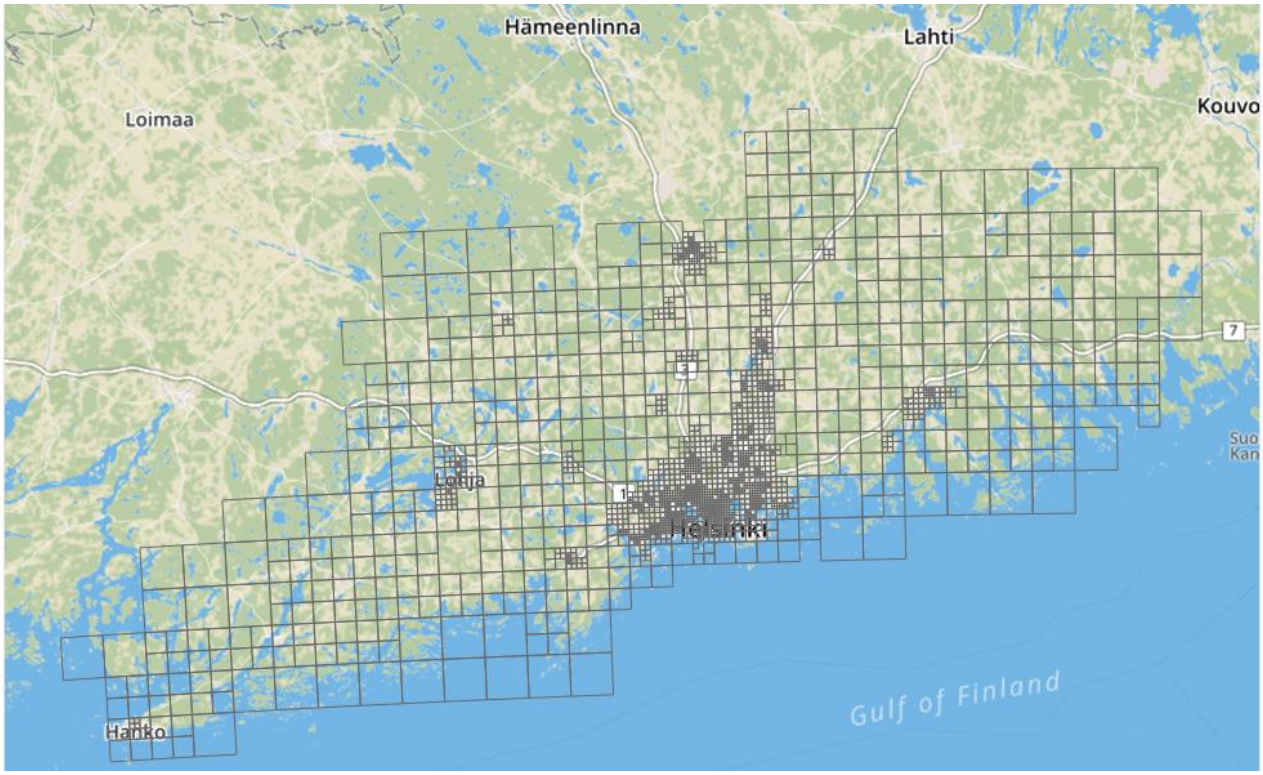
dominant_zone	count	time
2	299	23.1.2018 0.00.00
2	283	23.1.2018 1.00.00
2	229	23.1.2018 2.00.00
…	…	…

Joining with shapefile

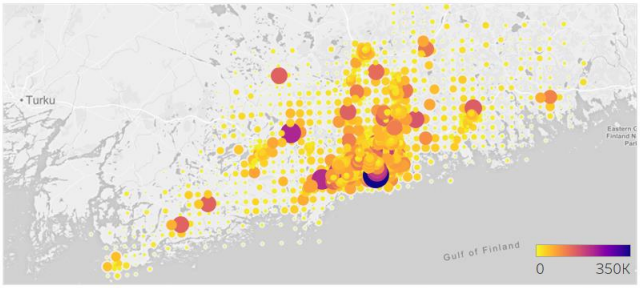
Column 'dominant_zone' is the grid cell in which the activity took place. This can be joined with 'ID' column from the shapefile.

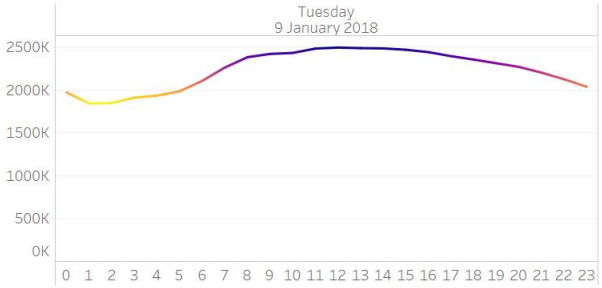
Junction data package example

Activity data coverage in Uusimaa and data set example from Jan 9th 2018.



HOURLY ACTIVITY LEVEL IN UUSIMAA PROVINCE ON JAN 9th 2018





HOURLY ACTIVITY LEVEL IN HELSINKI ICE HOCKEY ARENA ON JAN 9th 2018