MSc Thesis big data & food

Florine Kremer — MSc Food Security and Sustainable Consumption

Supervisors: Lector Dr. Sigrid Wertheim-Heck & Anke Brons

Master thesis in Environmental Policy, Wageningen University and Research.
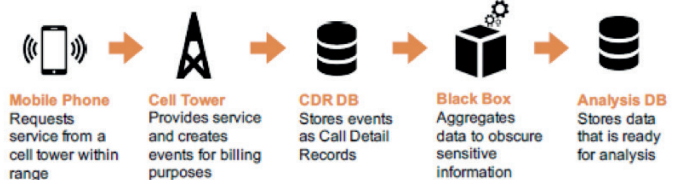Executed within research group Food and Healthy Living - Aeres University of applied Sciences Almere

AERES
UNIVERSITY OF APPLIED SCIENCES
ALMERE

in collaboration with **WAGENINGEN** UNIVERSITY & RESEARCH
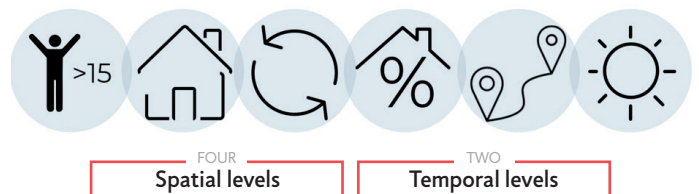
Background

Understanding people's everyday food consumption practices is essential in achieving the transition towards more healthy and sustainable diets. Because of the complex and context-specific nature of food practices, there is a prevalence of qualitative research. As these results are often based on small-scale datasets, the impact of social practice research on public policy is still limited. Larger datasets on food purchasing are often unavailable for scientific research, as they are often held by commercial companies. However, the collaboration with data analyst Mezuro allowed me to analyze aggregated and anonymised mobile phone location data of Dutch residents, which enabled me to explore the potential contribution of this big data source for consumer research. The aim of this study is to determine if and how mobile phone location data can be of value for researching food purchasing when applying a social practice perspective.

Methods

First, the characteristics of mobile phone location data from data analyst company Mezuro is described in detail. Hereafter, the data is applied to research regarding food purchasing practices in the Dutch city of Almere. The mobile phone location data is integrated with information about the food retail environment and sociodemographic profiles of the residents of Almere from the Dutch Central Bureau of Statistics (CBS) and the Flows of Food research project. Two perspectives are considered: the food retail outlets and the residents of Almere.

Data is collected and anonymized by Mezuro:

Mobile Phone Requests service from a cell tower within range

Cell Tower Provides service and creates events for billing purposes

CDR DB Stores events as Call Detail Records

Black Box Aggregates data to obscure sensitive information

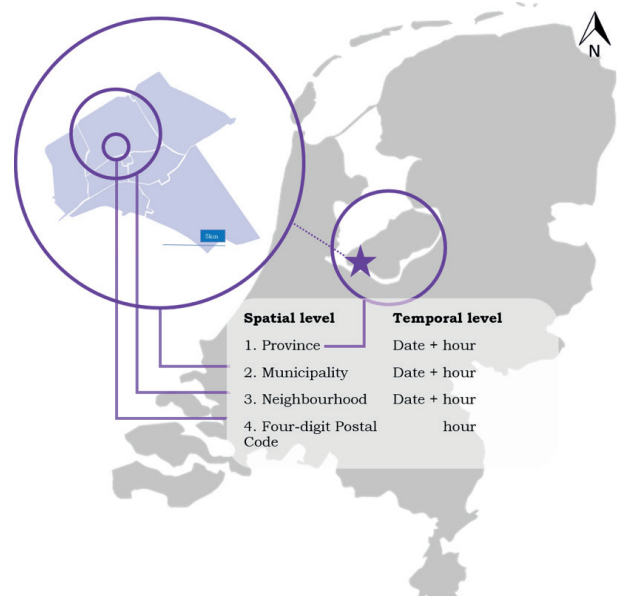**Analysis DB** Stores data that is ready for analysis

Results

The output is generated for groups of more than 15 individuals. Datasets are origin, destination, and population density. Available variables are the number of people present, origin, travel frequency, share of residents at home and away, distance travelled and weather data. Data can be collected on different spatial- and temporal levels.

FOUR	TWO
Spatial levels	Temporal levels

| Spatial level	Temporal level
1. Province	Date + hour
2. Municipality	Date + hour
3. Neighbourhood	Date + hour
4. Four-digit Postal Code	hour

In Almere, most food retail outlets are visited on Saturdays by visitors from the near vicinity (46-63% locals). Purple outlines indicate days with open-air markets. Specific food retail outlets are too detailed for analysis: data is most reliable for trips of >5 km.
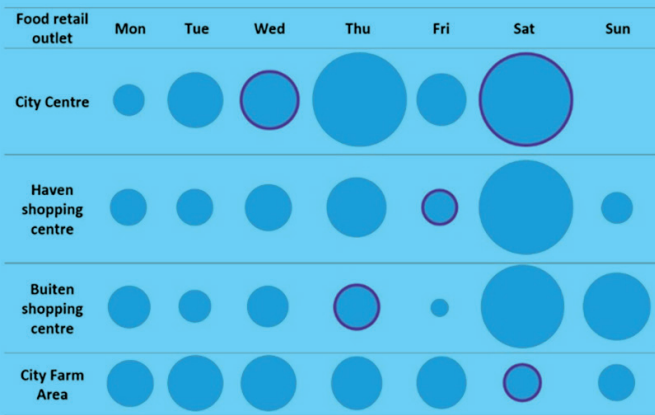
Residents of neighbourhoods with a larger distance to the nearest food retail outlet, more cars and a higher income per household tend to travel further than average.
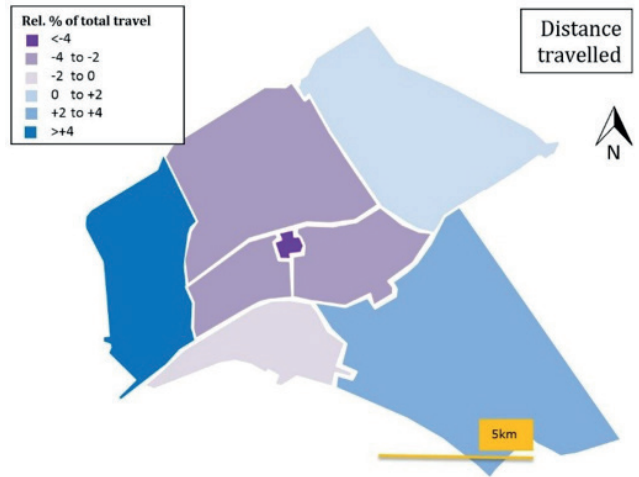
The residents of Almere travel mostly within their home neighbourhood or other areas within Almere. Another popular destination is Amsterdam. Although the reasons for travel cannot be known, it is likely people travel there for work purposes. Relatively more residents stay close to home during weekends than on weekdays. Most residents are frequent visitors, especially to Almere and Amsterdam.
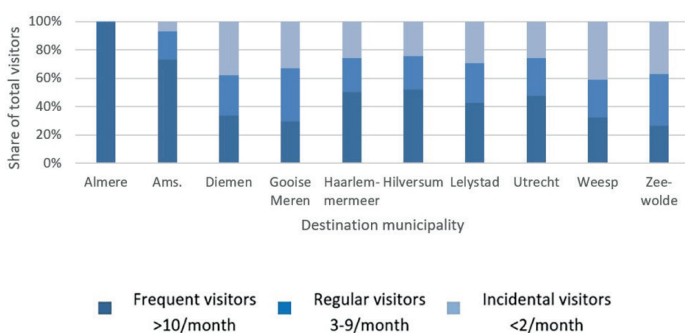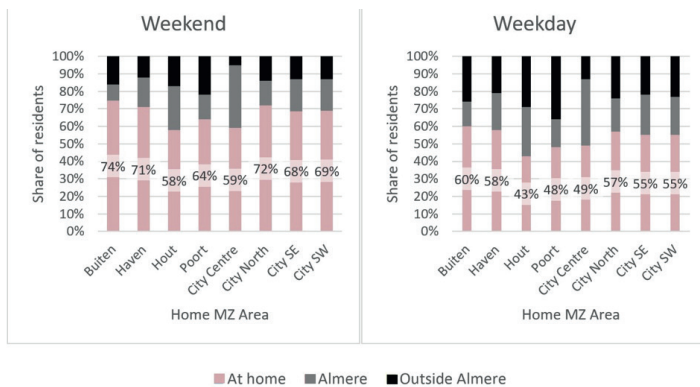
Conclusion

Mobile phone location data can serve as an addition to the toolbox of research on food consumption practices. It generates large-scale insights into the mobility and movements of the masses. With this, real-time outputs on multiple spatial scales can be used to confirm findings from qualitative research or to discover hidden behaviour patterns. The more qualitative research methods can, in turn, complement mobile phone location data by adding insights into the complexity and context of daily life. Mobile phone location data is most reliable on a larger geographical scale. Because of this, and strict privacy regulations, countries other than NL might provide more detailed insights into specific food retail outlets. Besides food consumption, the data can be used for a wide range of applications: urban planning, infrastructure, and mobility.

When combined with different sources and research methods, mobile phone location data can provide reliable and large-scale data, applicable in many different contexts. This research provides a mere glimpse of how big data can be used to generate valuable information about real-world events.