DATA SCIENCE MEETUP
PORTO

7 MARCH 2019
EVGHENI POLISCIUC

DSPT
DATA SCIENCE PORTUGAL

COMPUTATIONAL
DESIGN &
VISUALIZATION
LAB.

CISUC
COGNITIVE AND MEDIA
SYSTEMS

Hexagonal Gridded Map and Information Layers Applied in Retail Business Intelligence

Introduction

Retail business intelligence

- Analysis of customer-supermarket relationship
- Factors:
  - geographic location of customers;
  - demographic distribution;
  - customers' preferences;
  - accessibility to the store.

Our approach

The data

729

Hypermarkets

2

Years

2.86B

Transactions

6.6M

Client cards

278GB

Data

Aggregation by months

1st – total amounts spent, and the number of visits to the supermarkets for each customer;
2nd – customer–supermarket pairs with the frequency of visits to the nearest supermarket;
3rd – customer–supermarket pairs with the supermarket preference score.

Translates to

1st – the relationship between consumption and frequency of visits;
2nd – if the customers consume in the nearest supermarkets and how frequently;
3rd – if the customers prefer to visit the nearest supermarkets or not.

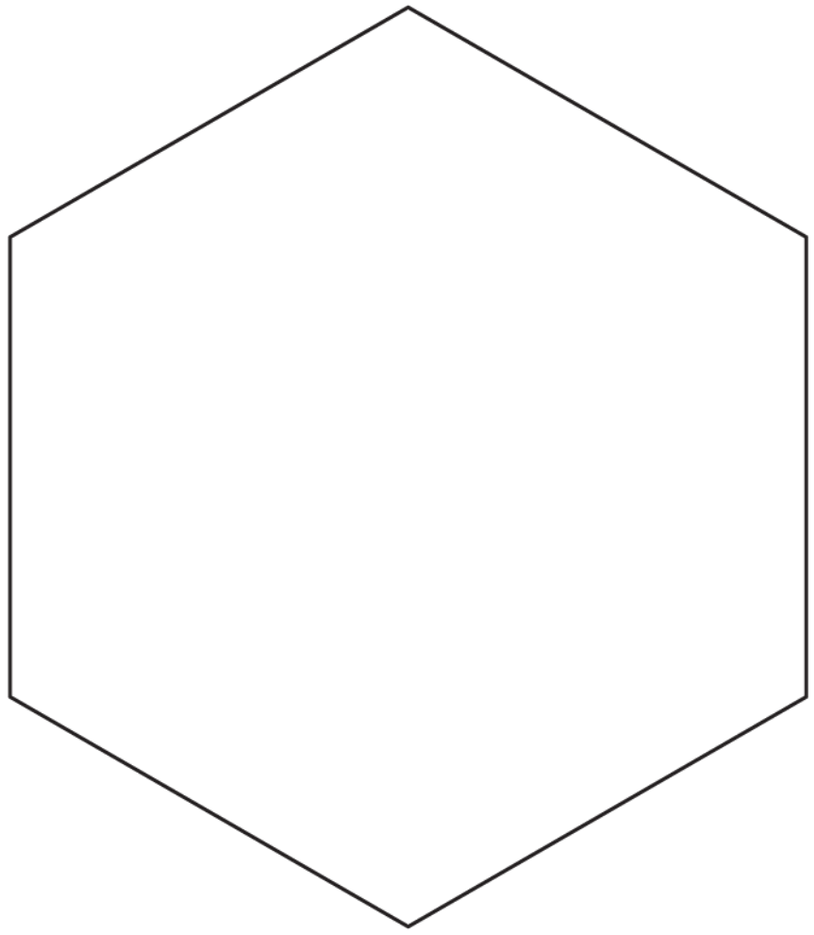
Direct representation
of frequencies of visits

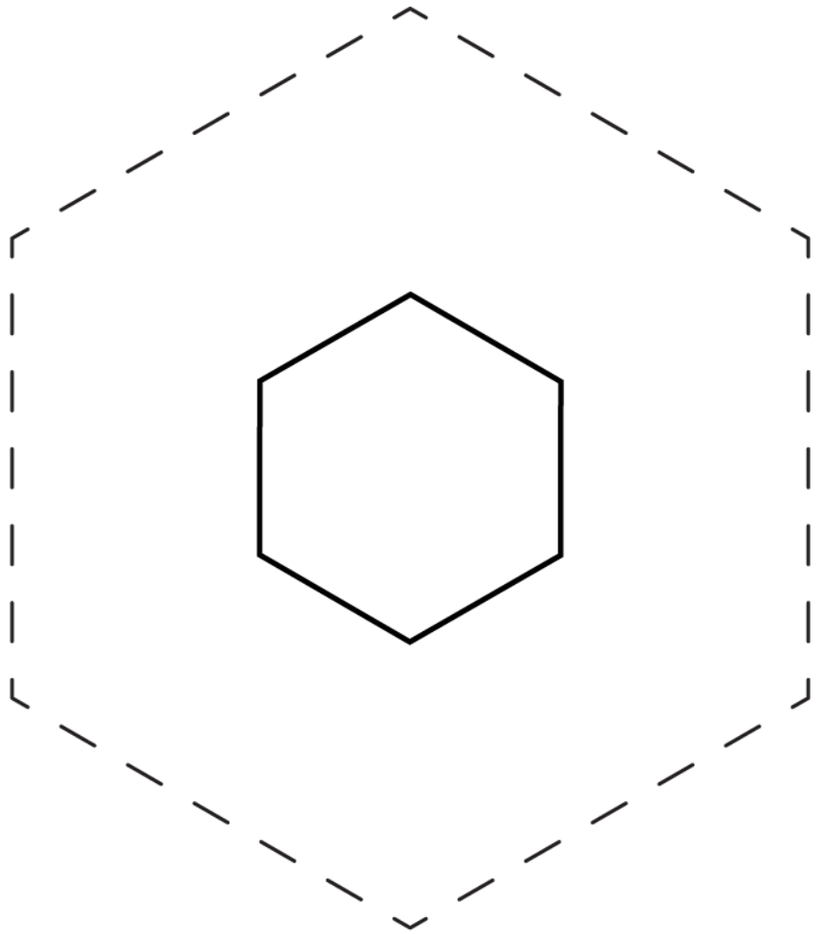
Hexagonal gridded map

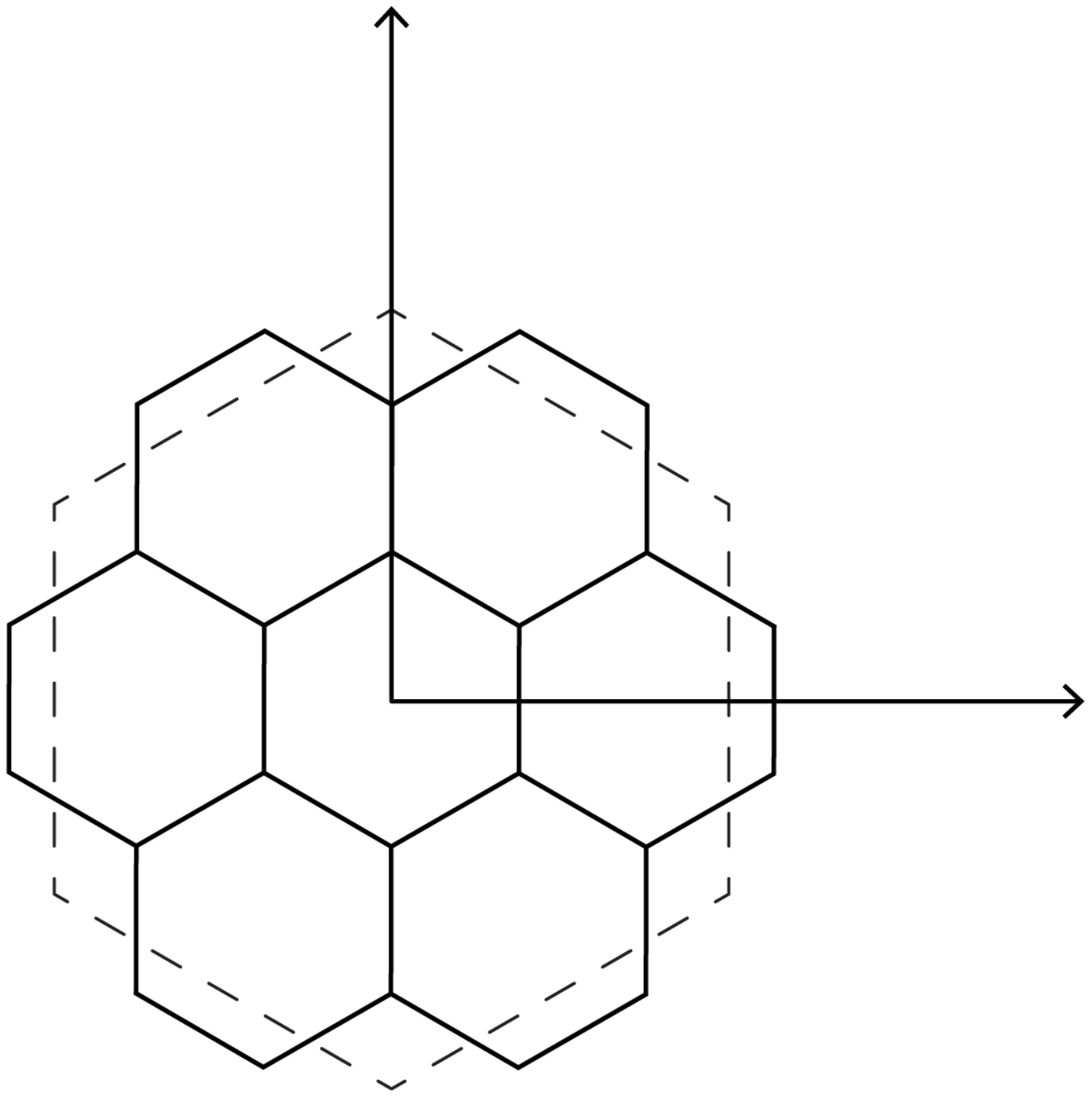
- Diverse information layers;
- Adaptive zoom;
- Provide means to:
  - explore and analyze data regarding customer-supermarket relations;
  - reveal the impact of supermarkets localization on customer preferences;
  - suggest areas of low coverage by supermarkets.

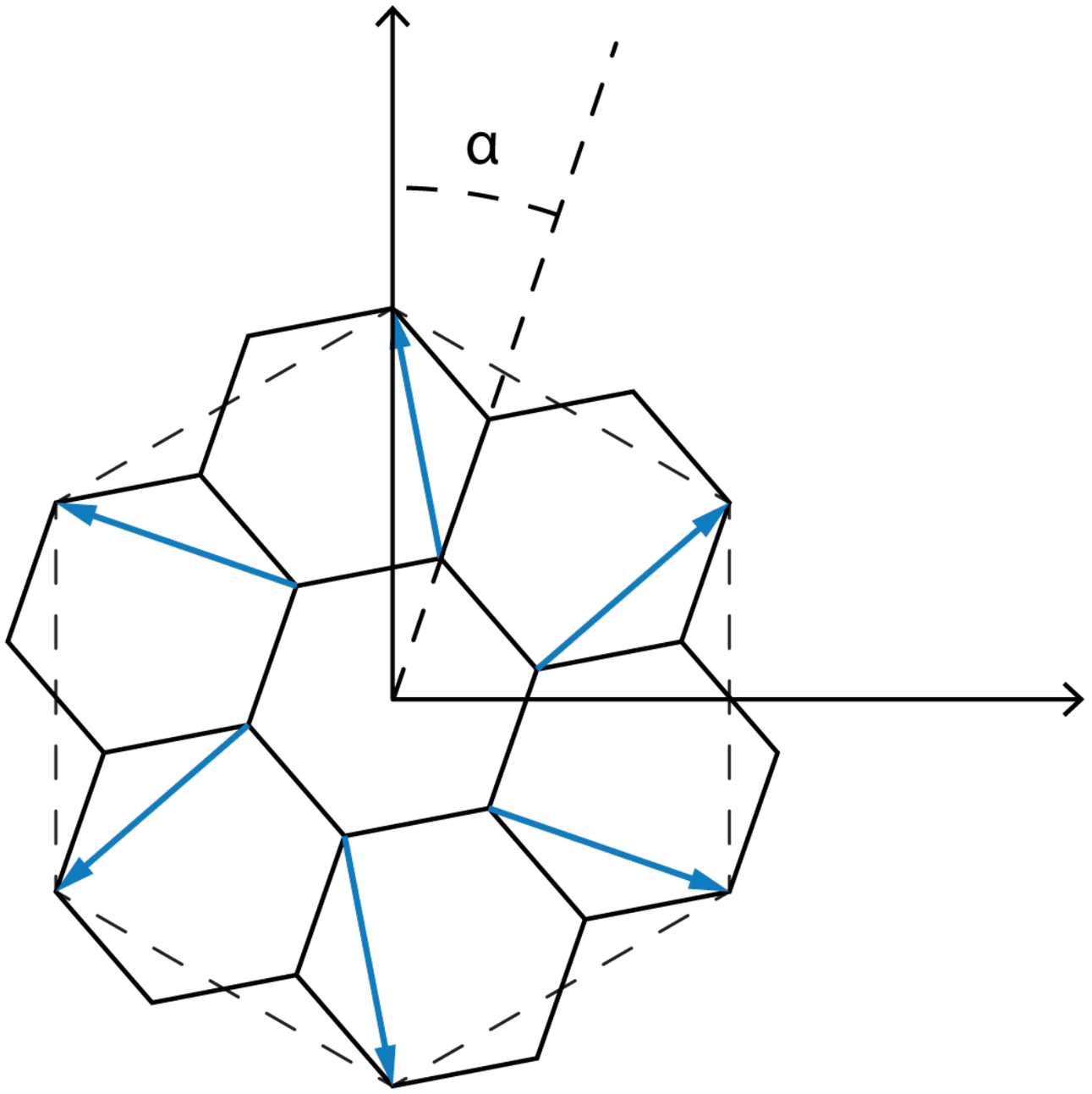
Hexagonal binning

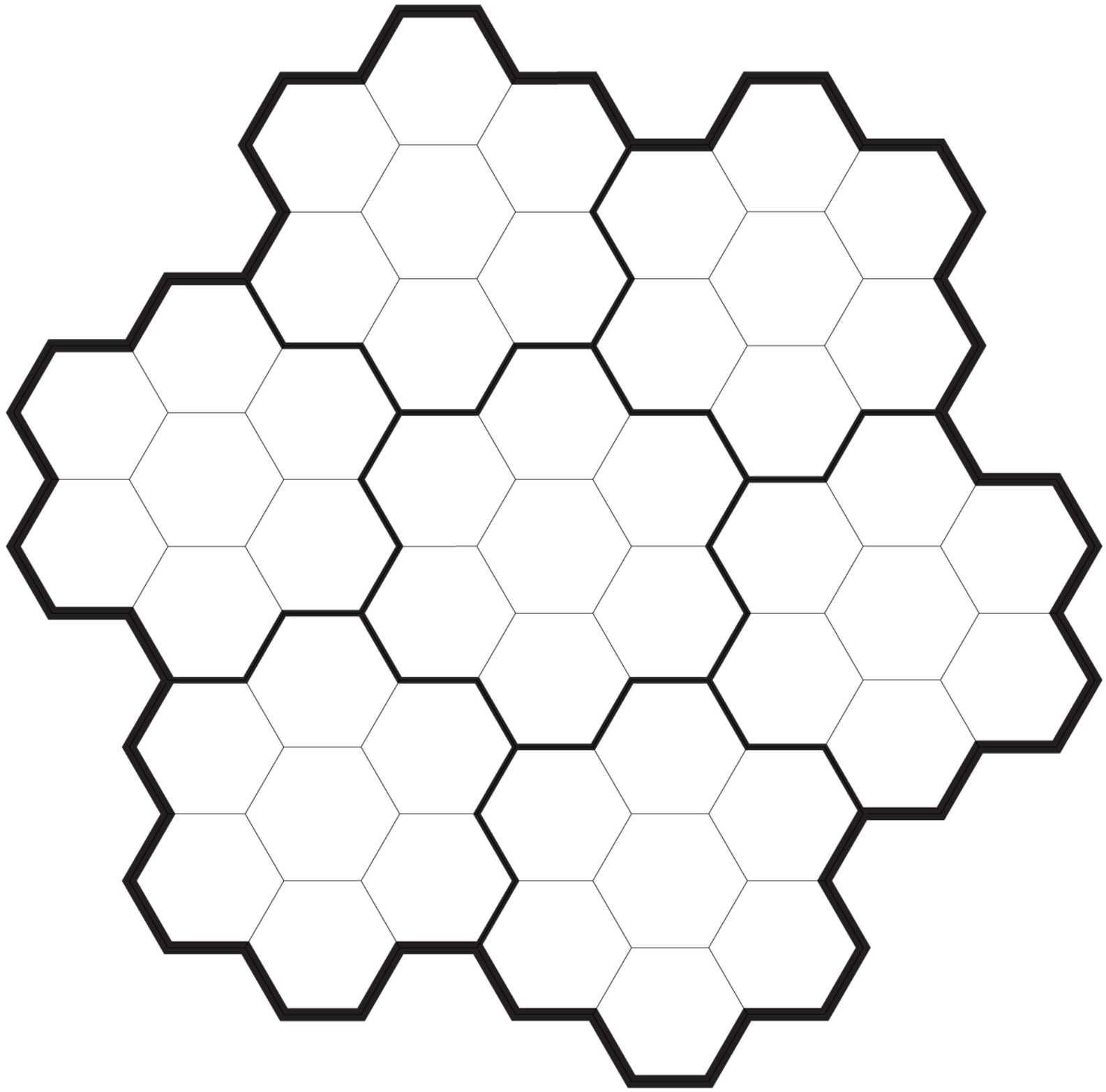
1. Start with an hexagon and build a tree (top-down)
2. Compute the geometry of polygons (bottom-up)
3. Insert the points into the three (top-down)
4. Merge branches (bottom-up)
5. Compute combined circles

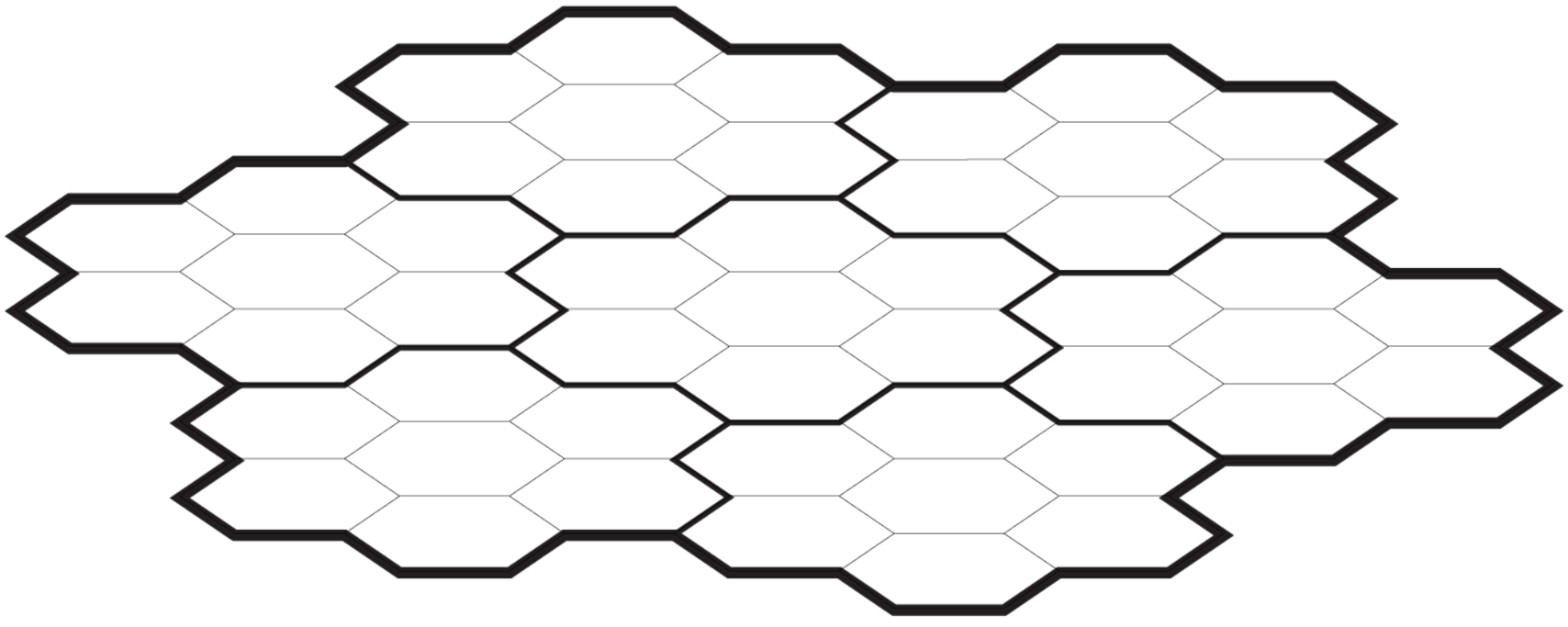


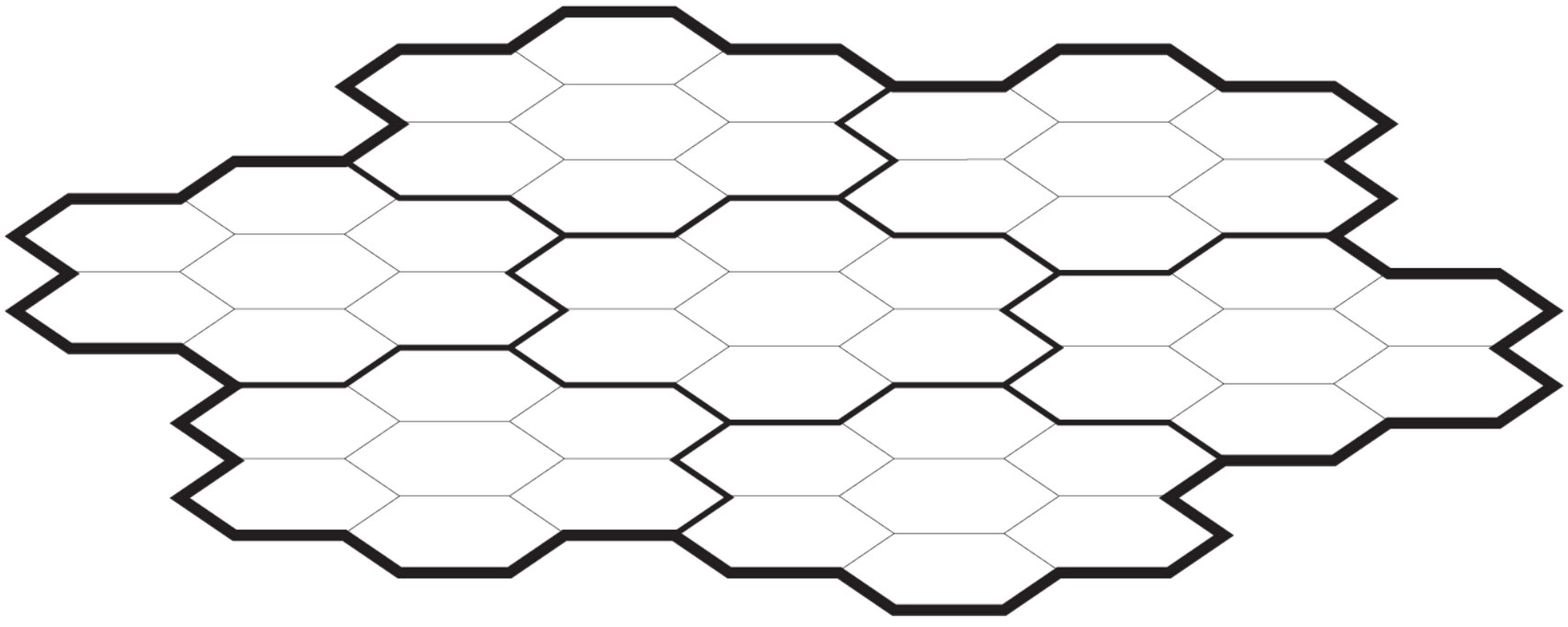


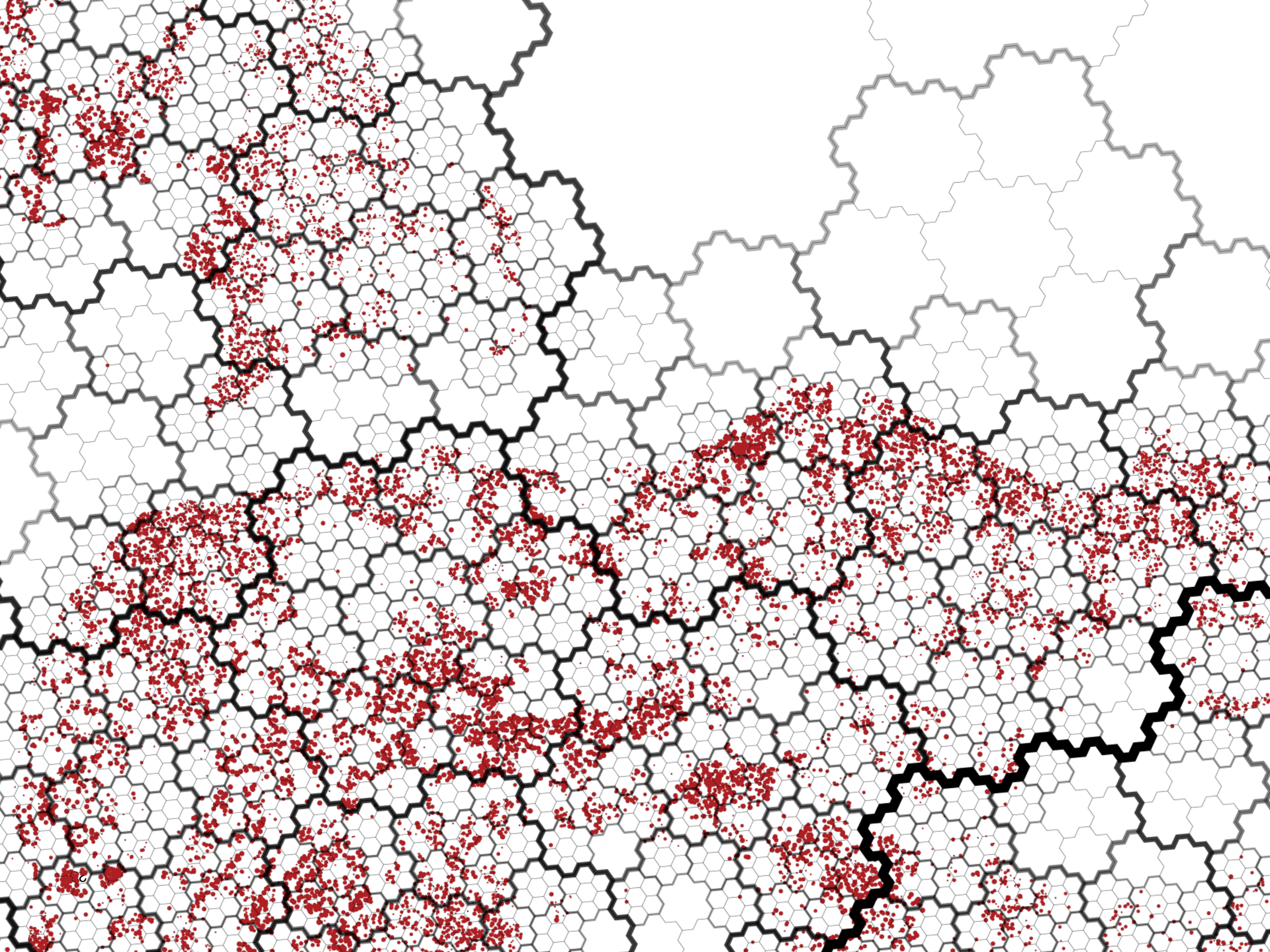



α

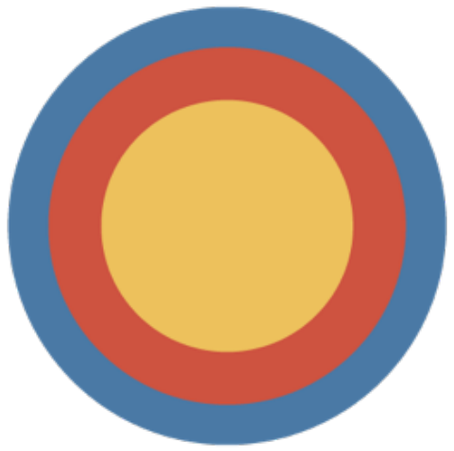








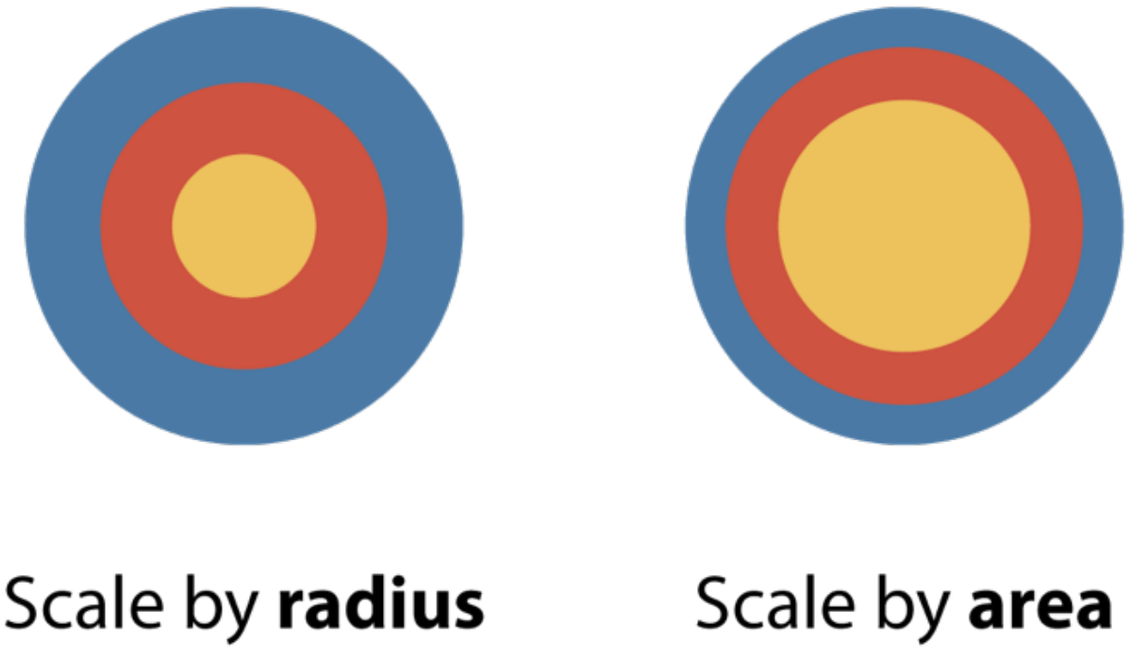
Composite Circles



$$A_{yellow} = A_{red} = A_{blue} \approx 33\%$$

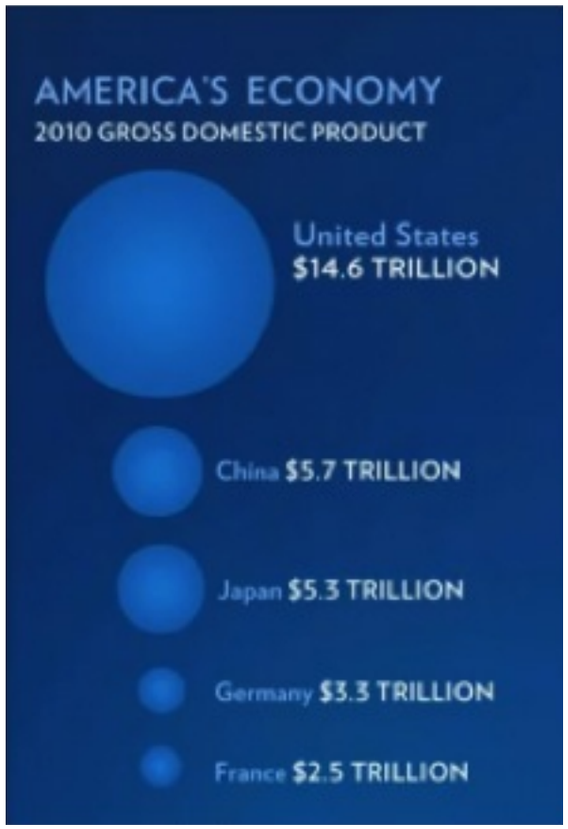
Three component proportions

Composite Circles



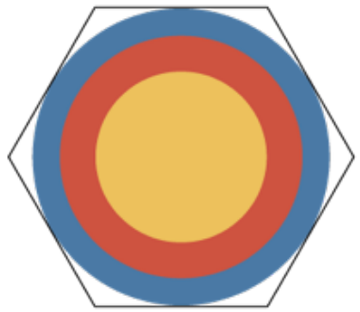
Comparison of mapping scales

Composite Circles



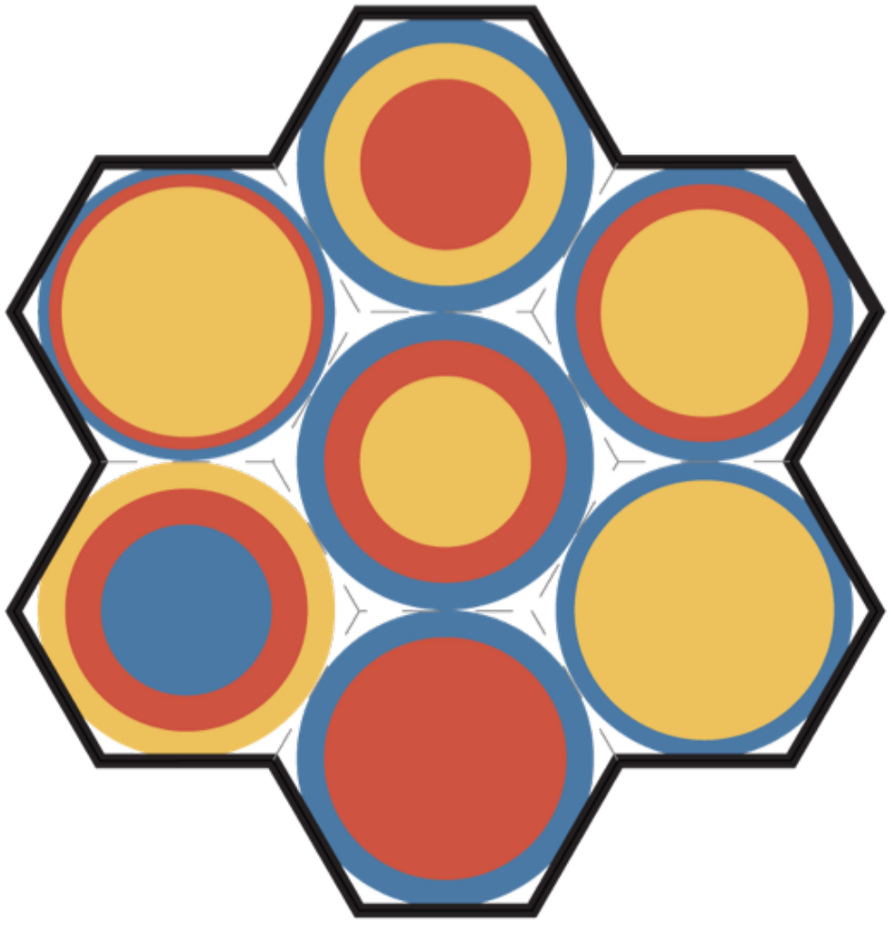
Comparison of mapping scales

Composite Circles

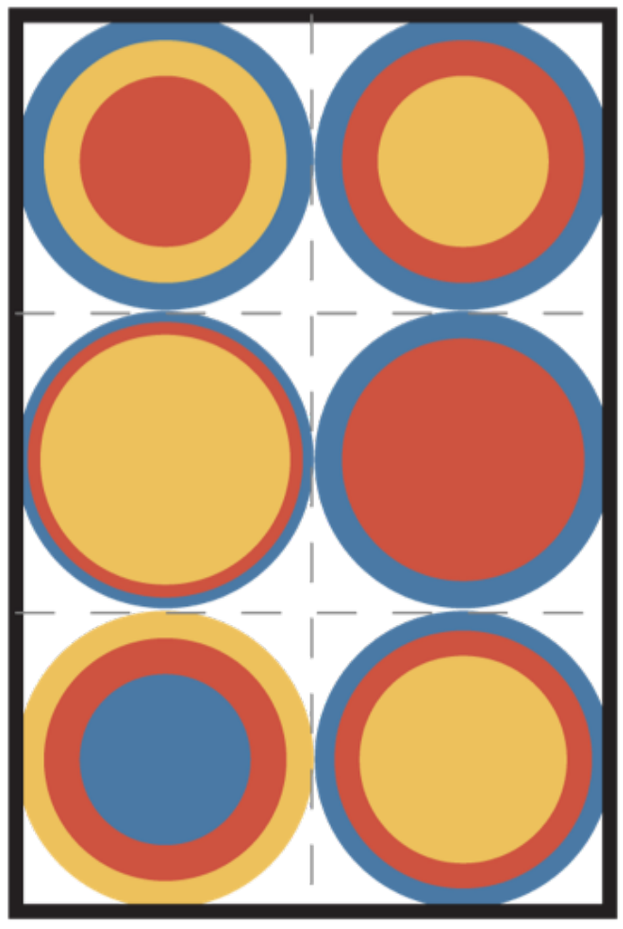


Positioning of the circle inside a hexagon

Composite Circles



Hexagonal Grid



Rectangular Grid

Comparison of both grids

Frequency | **Proximity** | Preference

Data regarding customers' visits to nearby stores.

Customers (•) that:

- Do not shop in the nearest supermarket
- Shop one or less times per week in the nearest supermarket
- Shops more than one time per week in the nearest supermarket

May 2012

Frequency | **Proximity** | Preference

Data regarding customers' visits to nearby stores.

Customers (•) that:

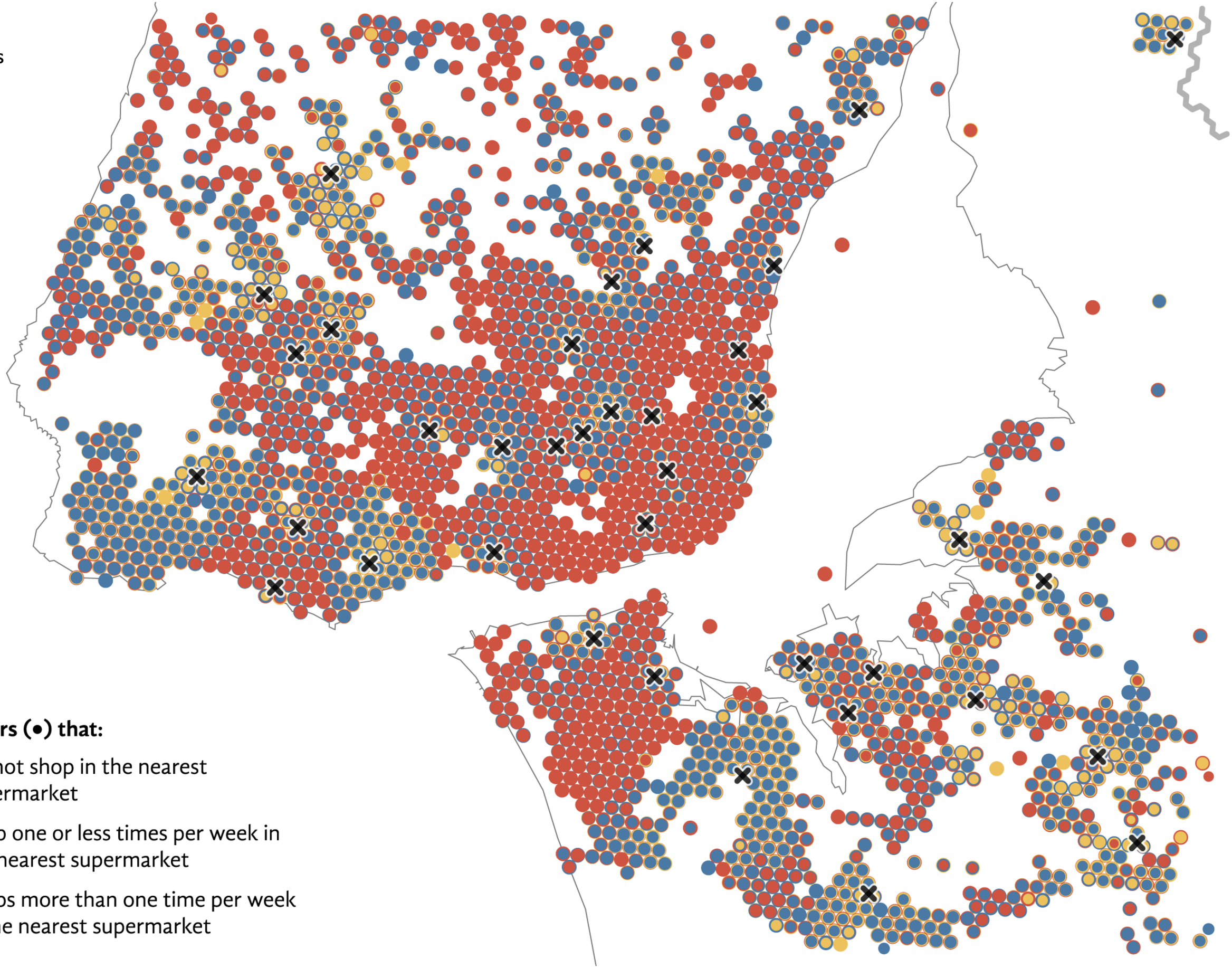
- Do not shop in the nearest supermarket
- Shop one or less times per week in the nearest supermarket
- Shops more than one time per week in the nearest supermarket

CIRCLES

Customers (●) that:

- Do not shop in the nearest supermarket
- Shop one or less times per week in the nearest supermarket
- Shops more than one time per week in the nearest supermarket

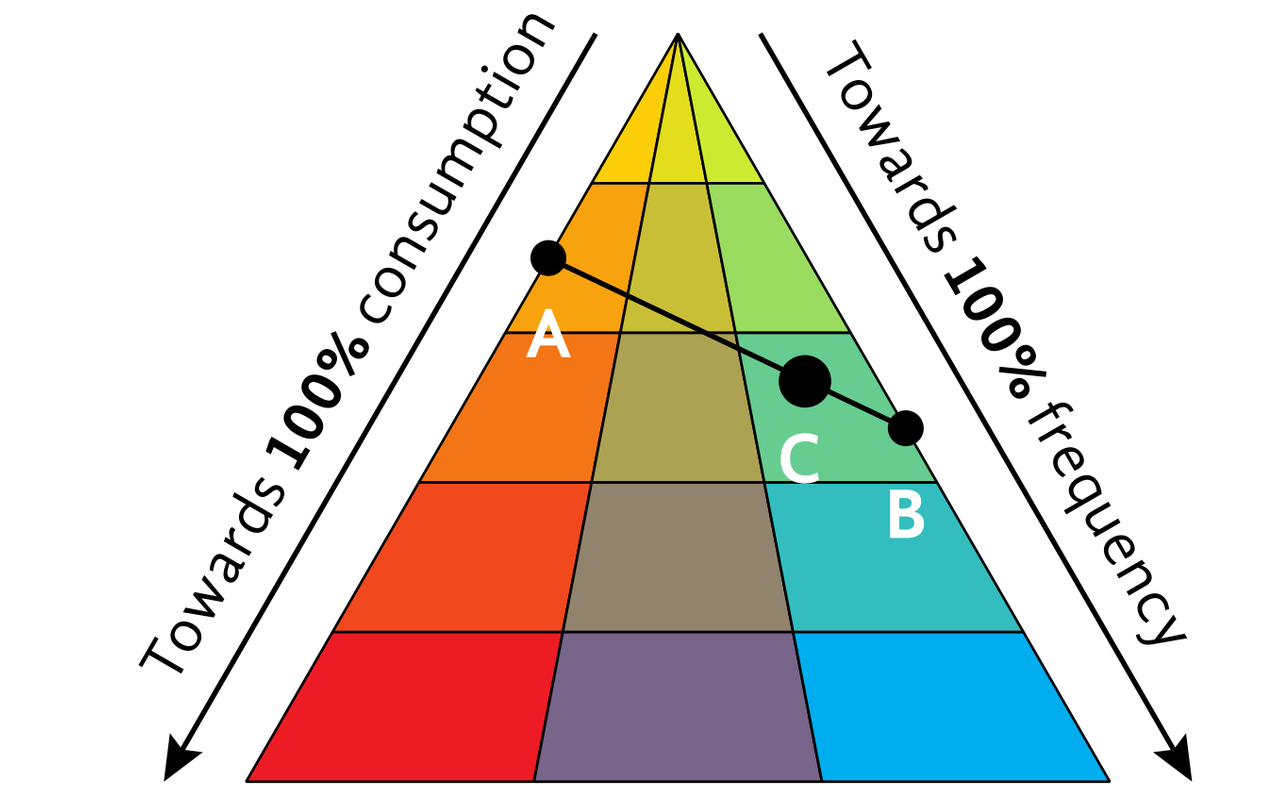
CIRCLES



Customers (●) that:

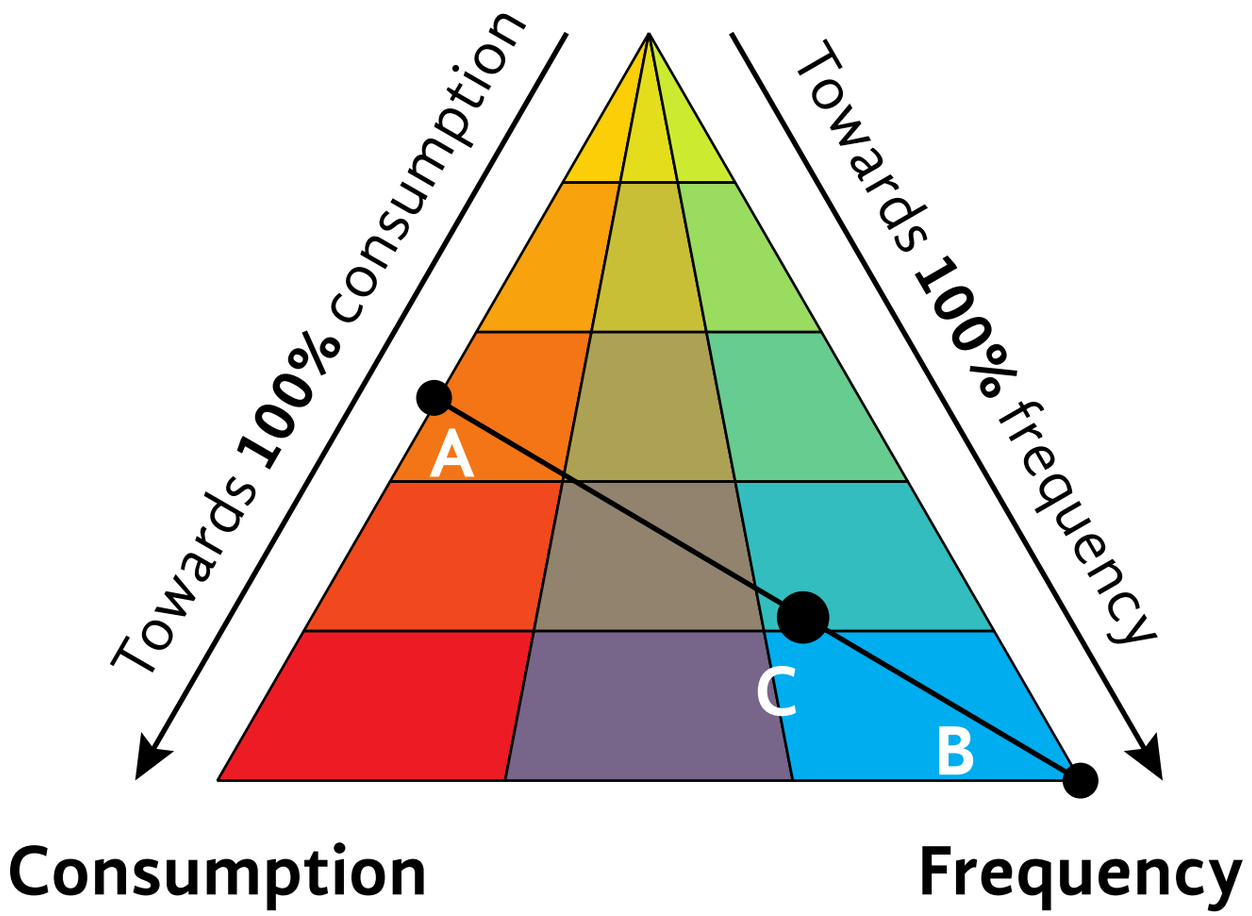
- Do not shop in the nearest supermarket
- Shop one or less times per week in the nearest supermarket
- Shops more than one time per week in the nearest supermarket

Bivariate color scheme



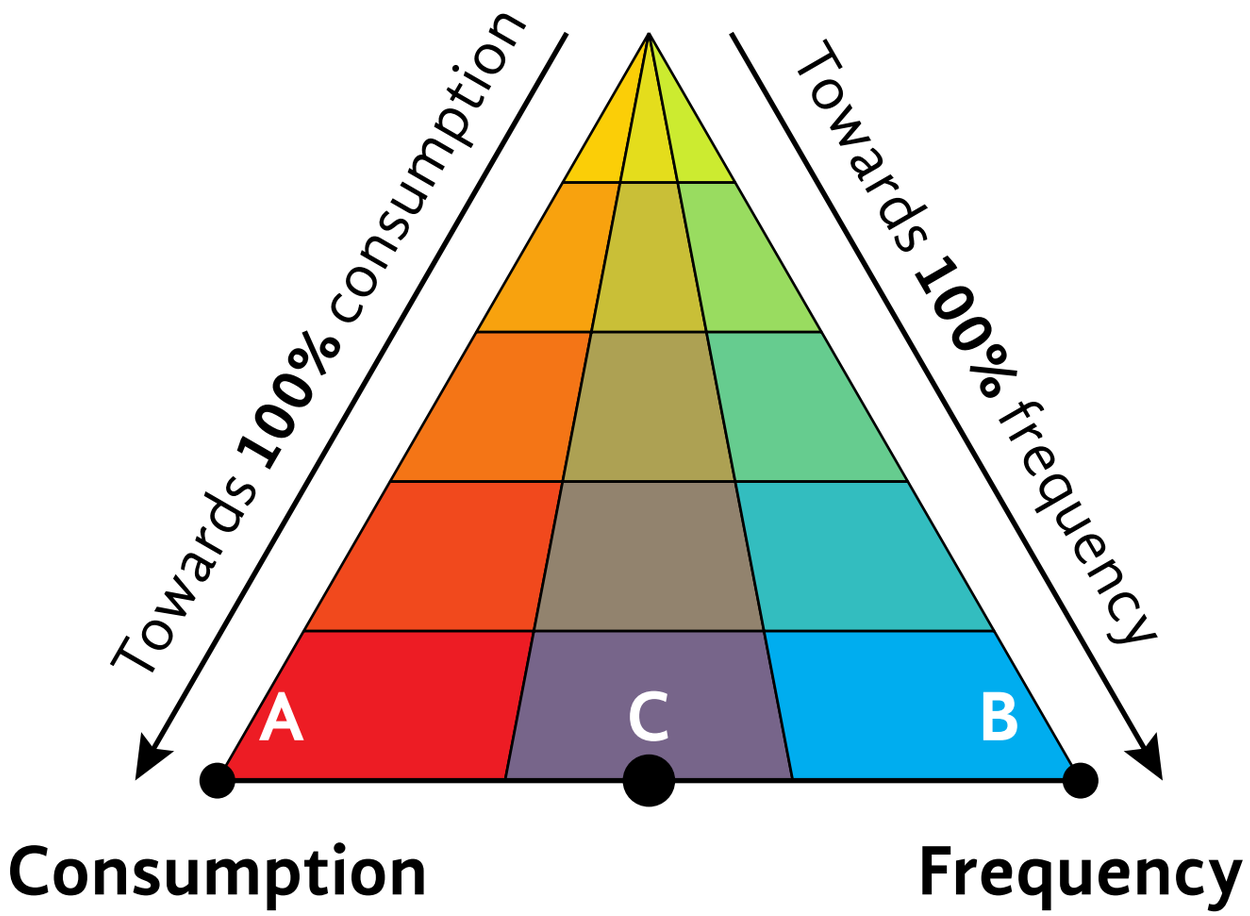
Triangular color scheme

Bivariate color scheme



Triangular color scheme

Bivariate color scheme



Triangular color scheme

Julho 2012

Frequency | **Proximity** | Preference

Data regarding customers' visits to nearby stores.

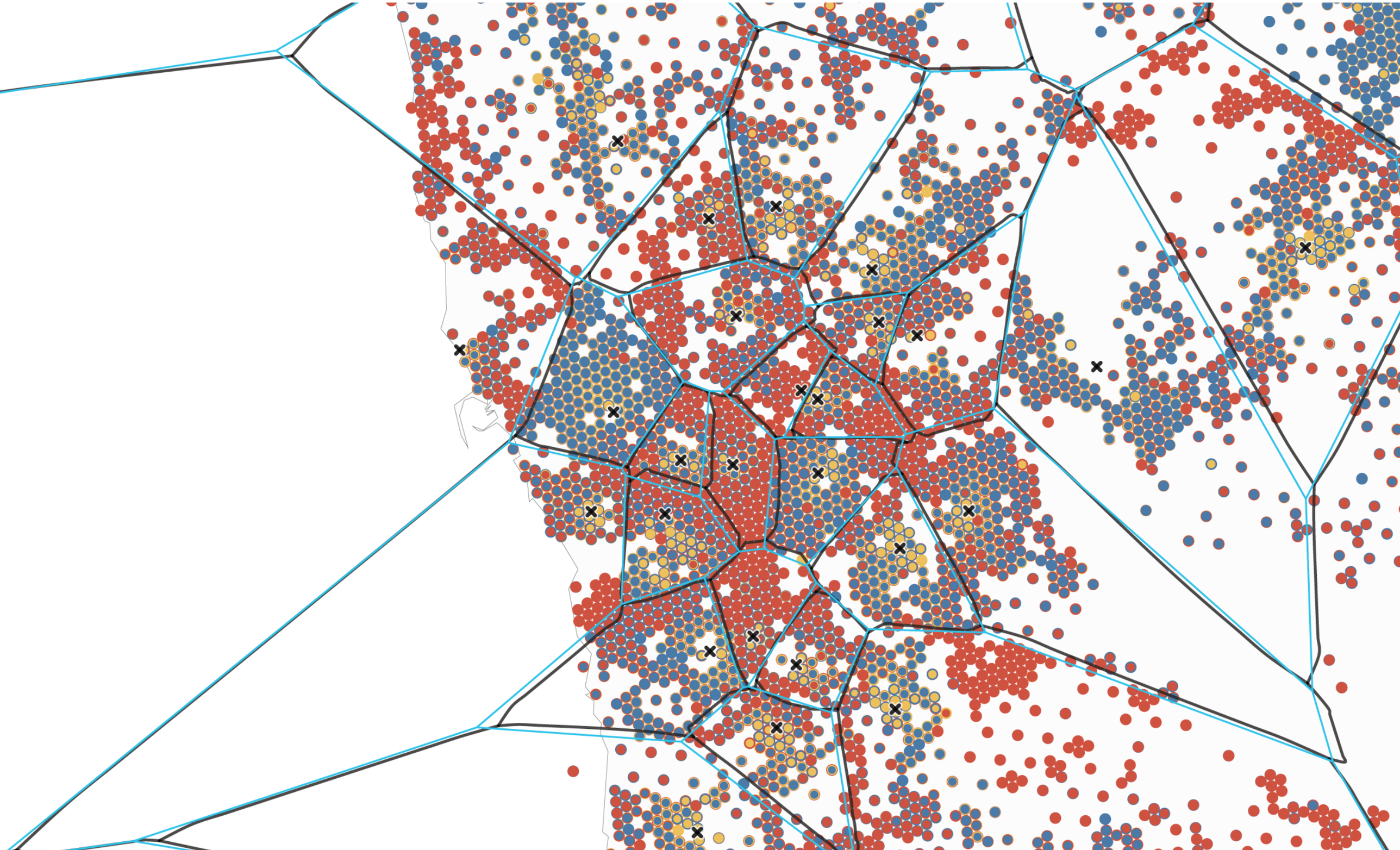
Customers (•) that:

- Do not shop in the nearest supermarket
- Shop one or less times per week in the nearest supermarket
- Shops more than one time per week in the nearest supermarket

Information layers

- Territorial boundaries
- Choropleth map

Territorial boundaries



Zoom-in

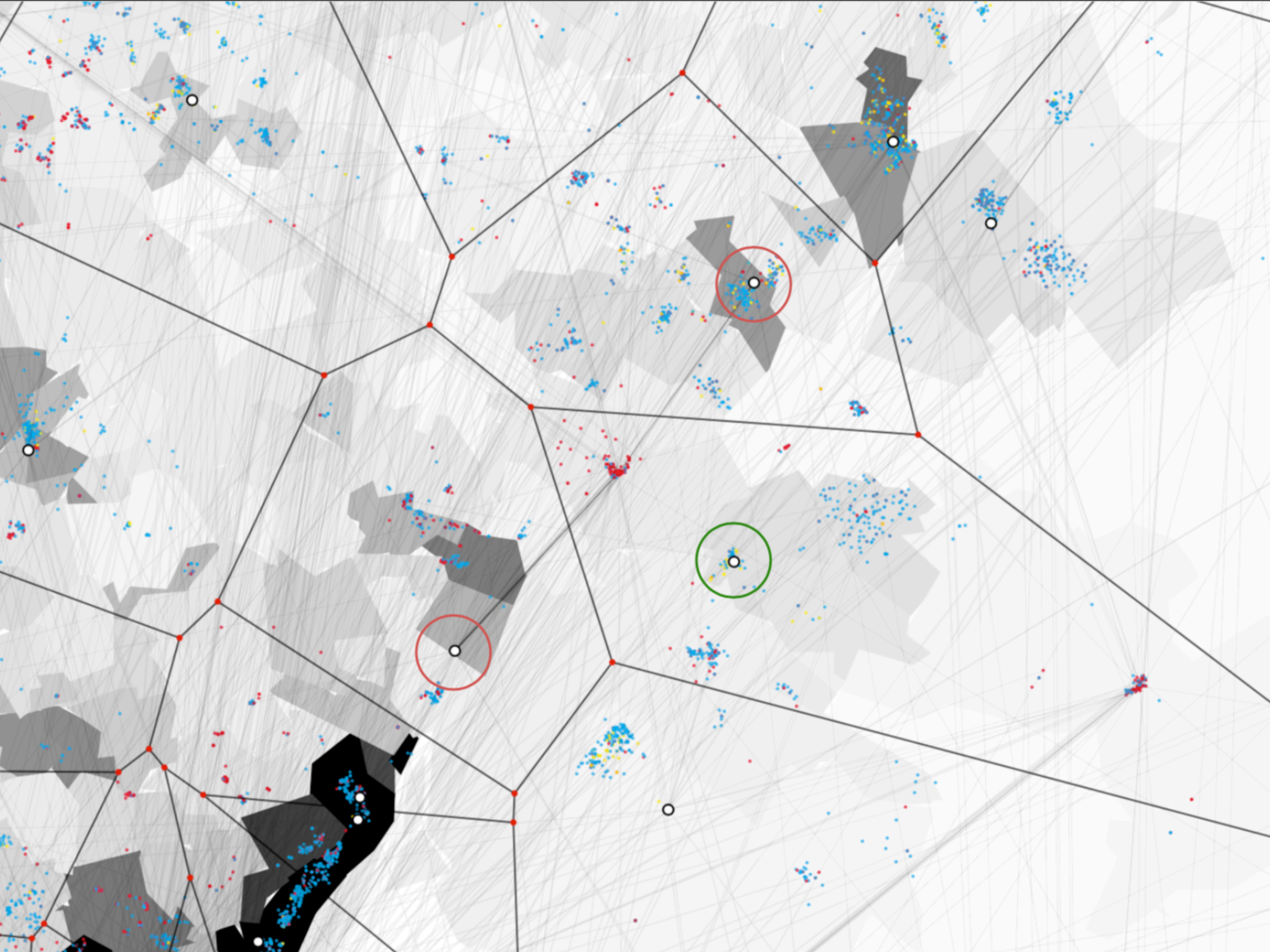
1120, 267,

733, 430, 275,

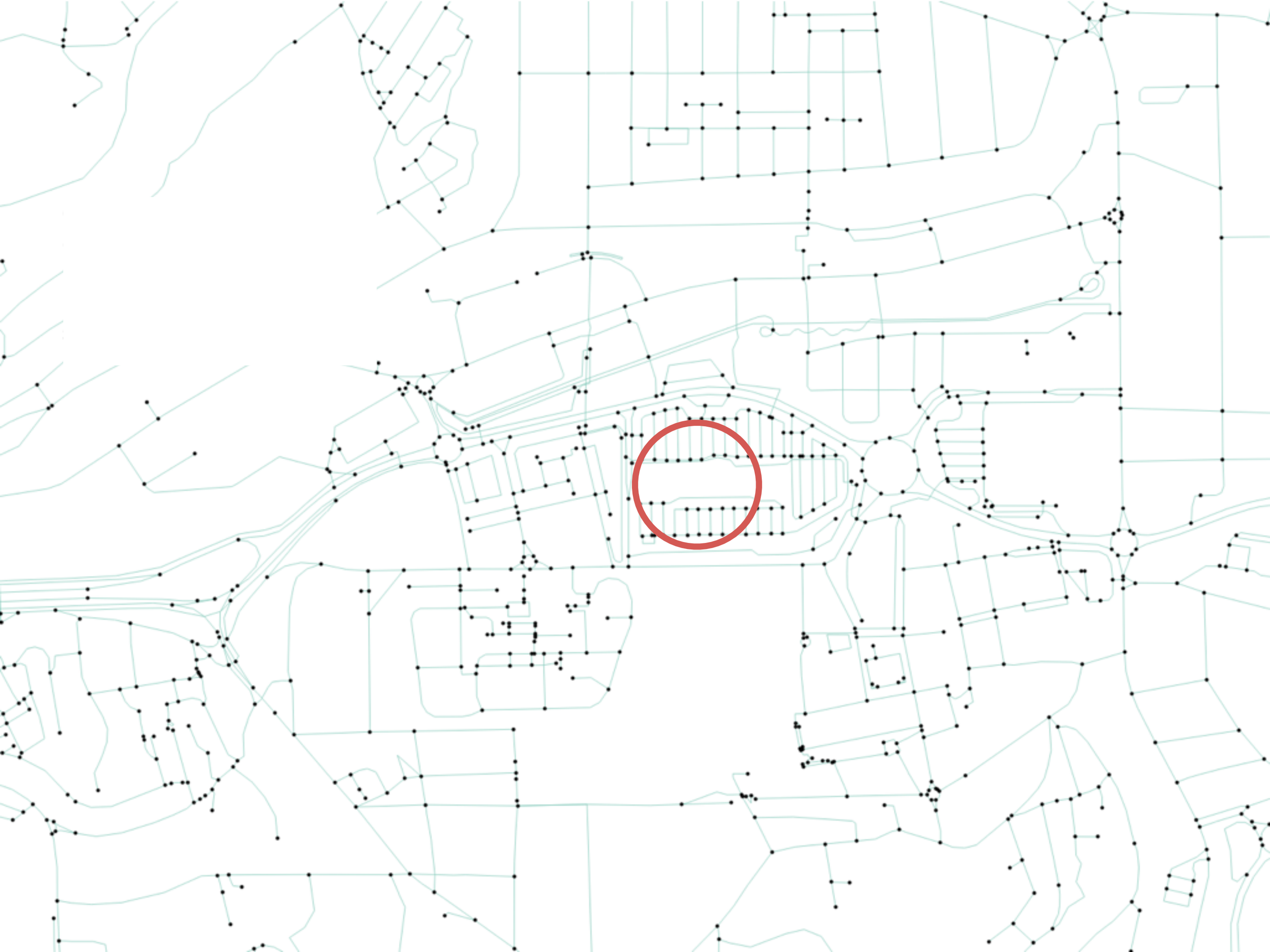
465, 484, 2754,

8, 767, 358,

451, 46, 293,



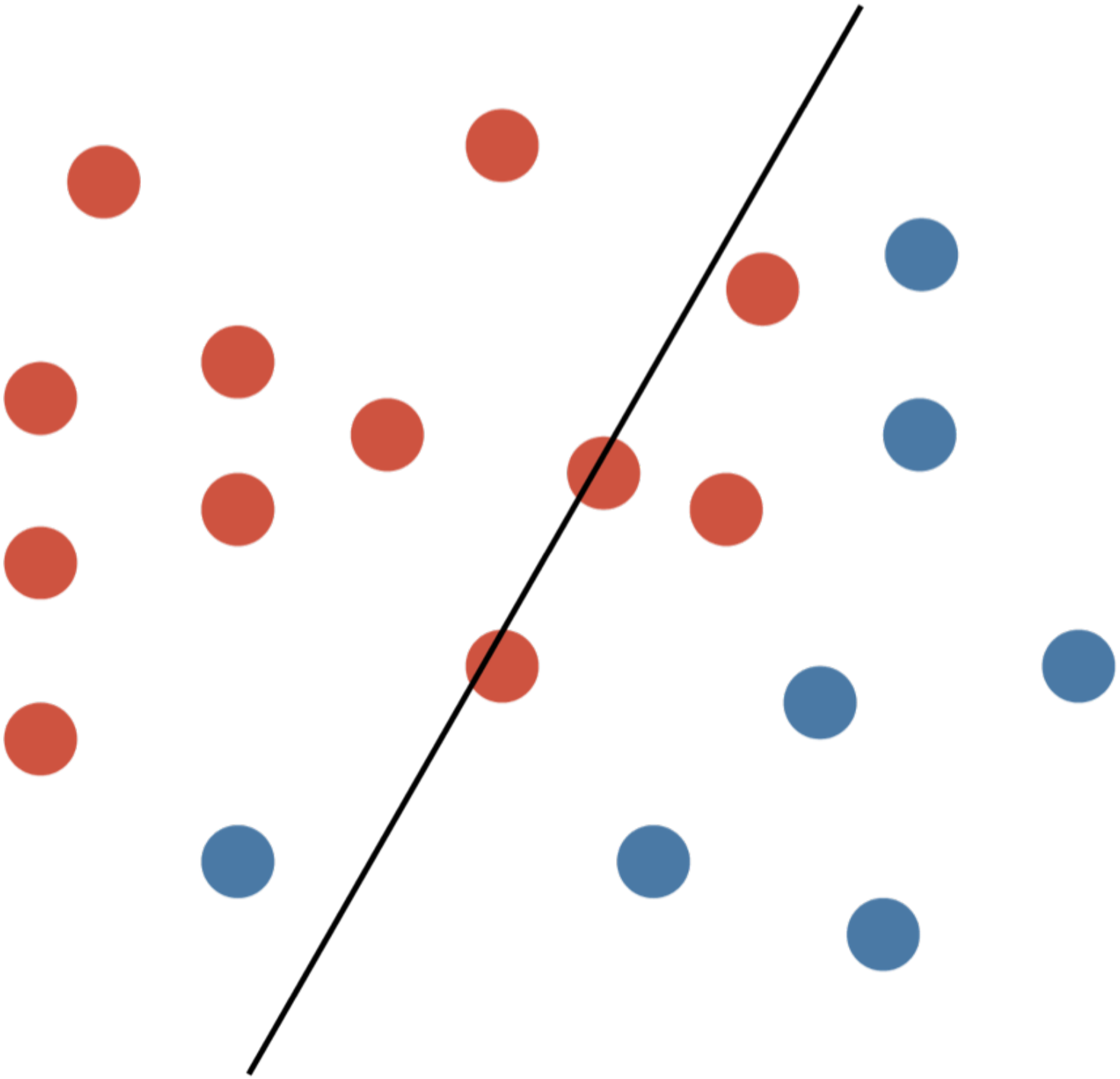
Ferrel
Atouguia da Baleia
Amoreira
IP6
Rio Maior
IC2
A15
Tremês
Evgheni
Tajo
Continente Modelo
Cadaval
Lourinhã
Ribamar
Serra de Montejunto
Alcoentre
Carnes Do Continente-indústria E...
Continente Santarém
Alpiarça
A13
Continente Modelo Almeirim
Almotolias
Continente Modelo do Cartaxo
Fazendas de Almeirim
A dos Cunhados
Abrigada
Aveiras de Cima
Continente Modelo
Azambuja
Alenquer
Marinhais
IC10
Glória do Ribatejo
Sobral de Monte Agraço
Carregado
Continente Modelo Salvaterra de Magos
Arruda dos Vinhos
Continente Modelo Vila Franca de Xira
Benavente
Fajarda
Coruche
Gradil
Mafra
Malveira
A10
Continente Modelo Porto Alto
Cavaleiros
Continente Bom Dia Alverca
Continente Modelo Alverca
Santo Estevão
Reserva Natural do Estuário do Tejo
Continente
Infantado
IC3
Lave
Cortiçadas de Lavre
Continente Bom Dia



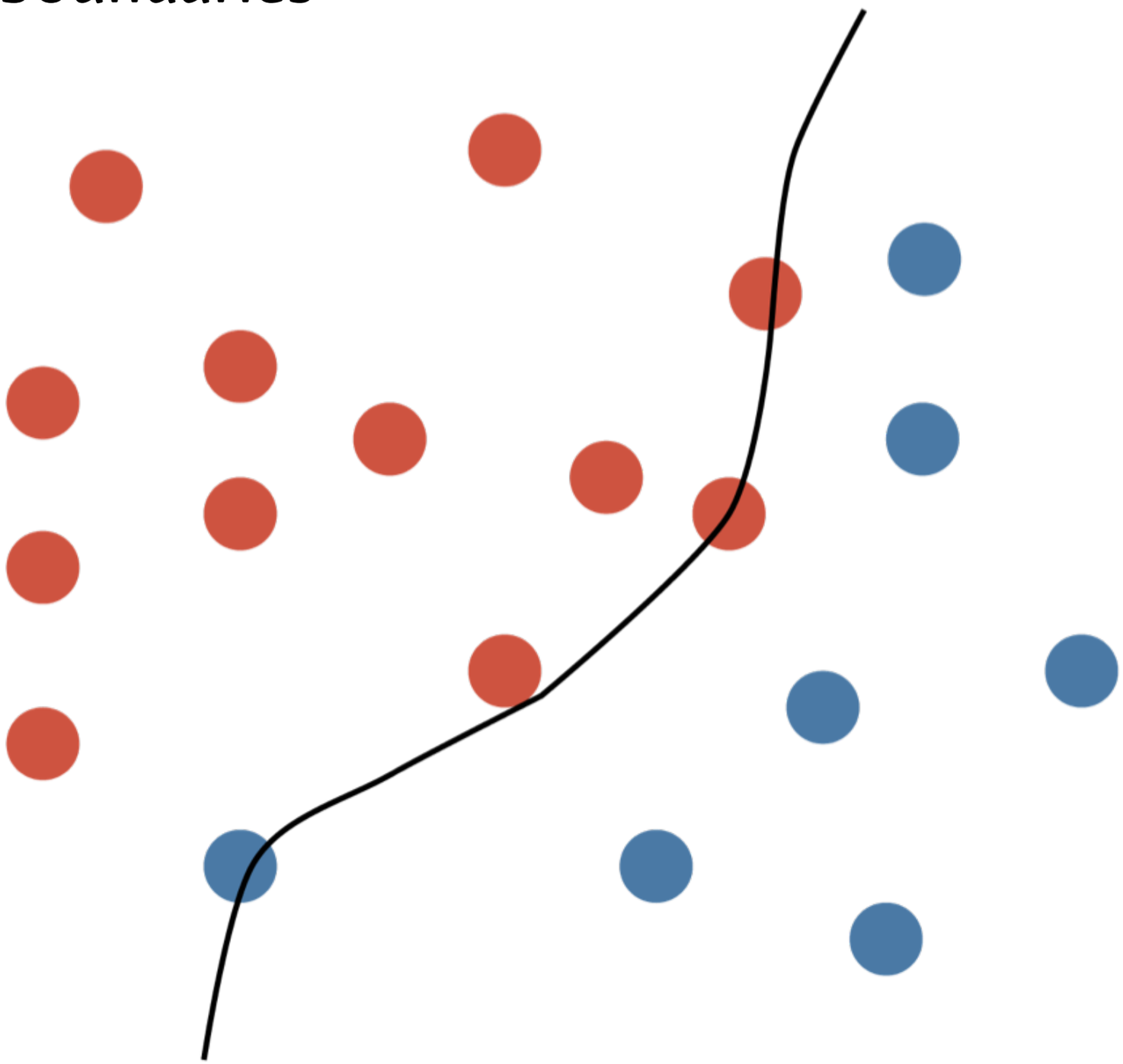
Weighted distortion of the Voronoi

1. Subdivide edges of the boundaries;
2. Each vertex is attracted by the points that do not belong to the area being computed;
3. The vertices are pushed back to their original position.

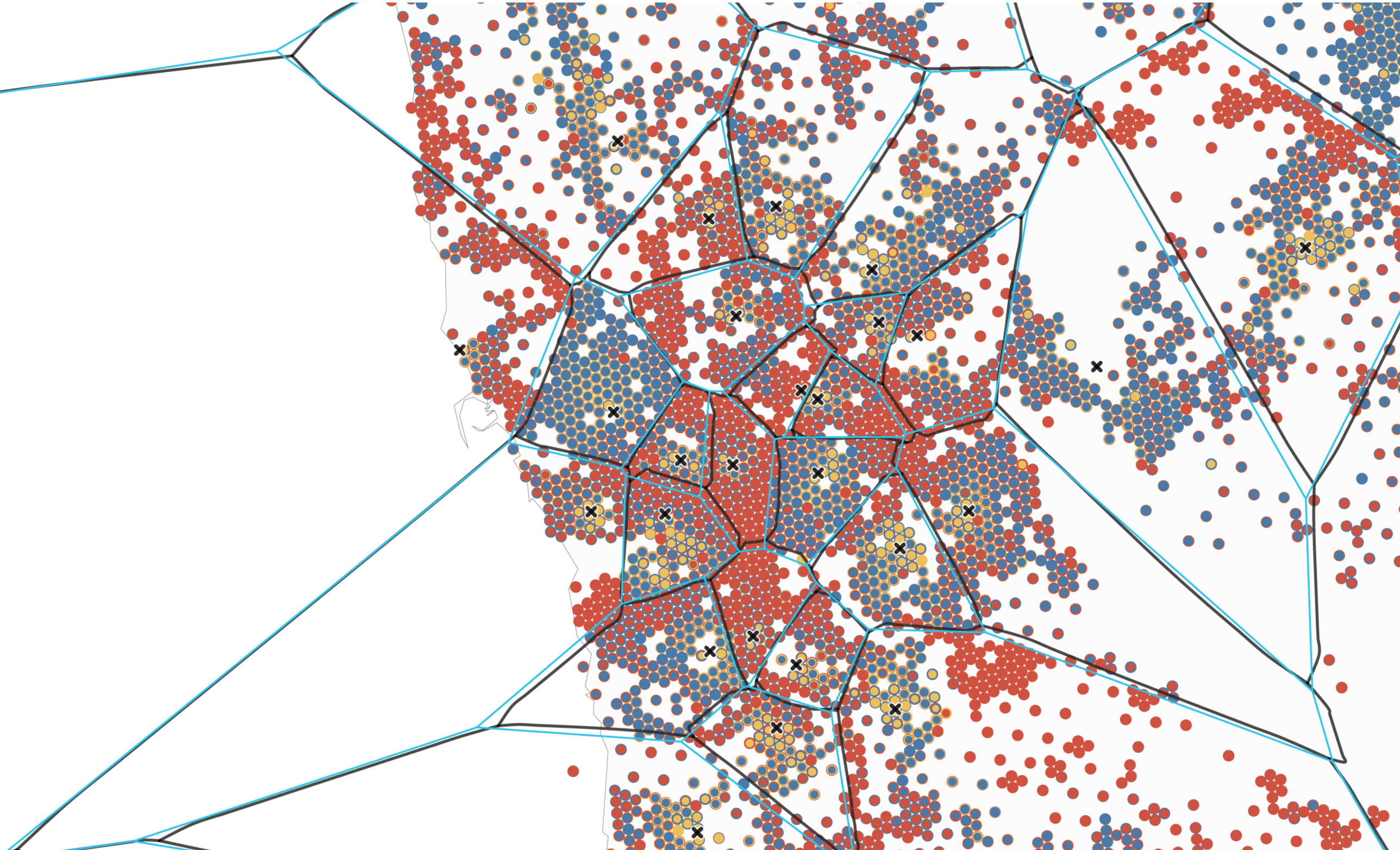
Territorial boundaries



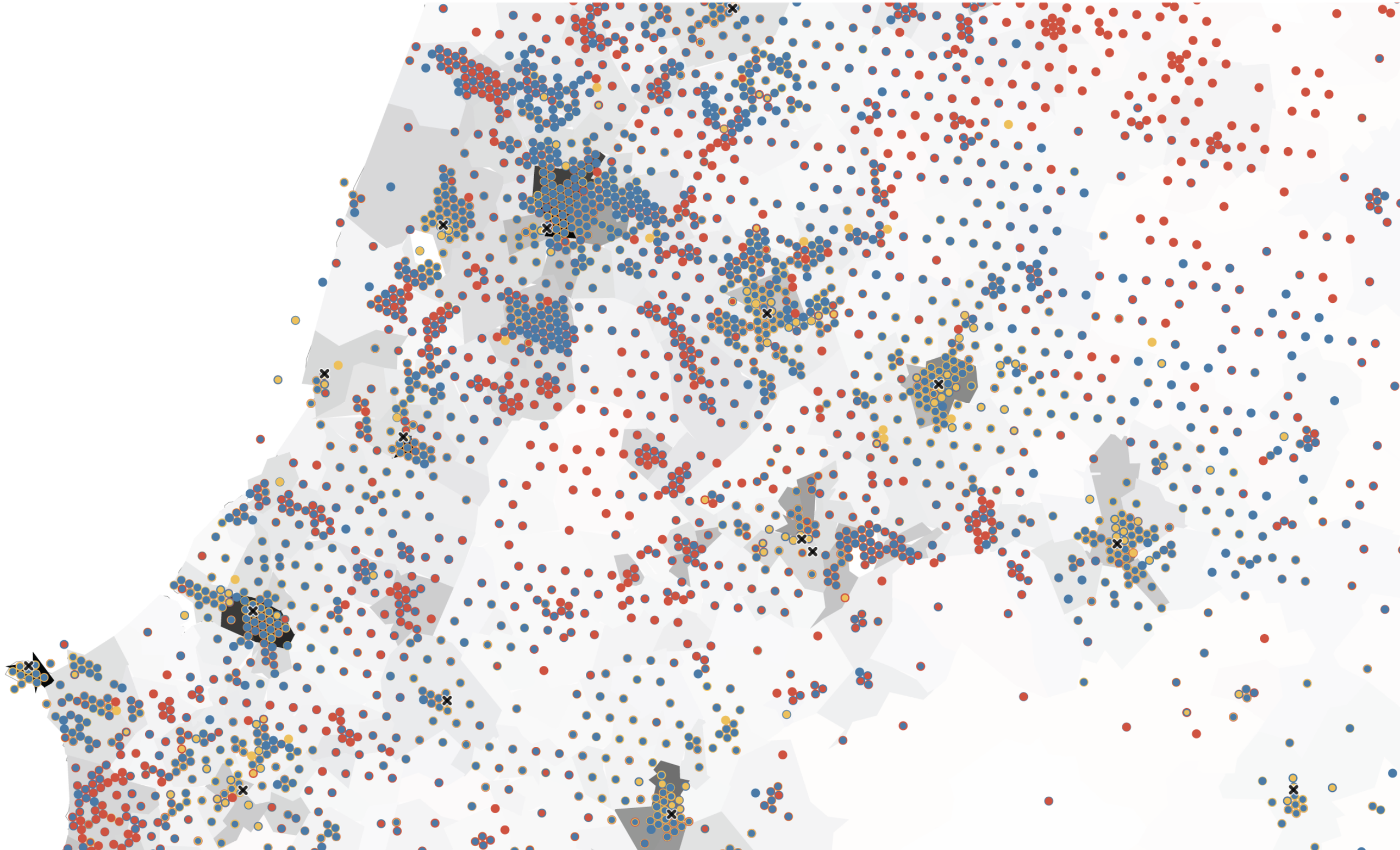
Territorial boundaries



Territorial boundaries

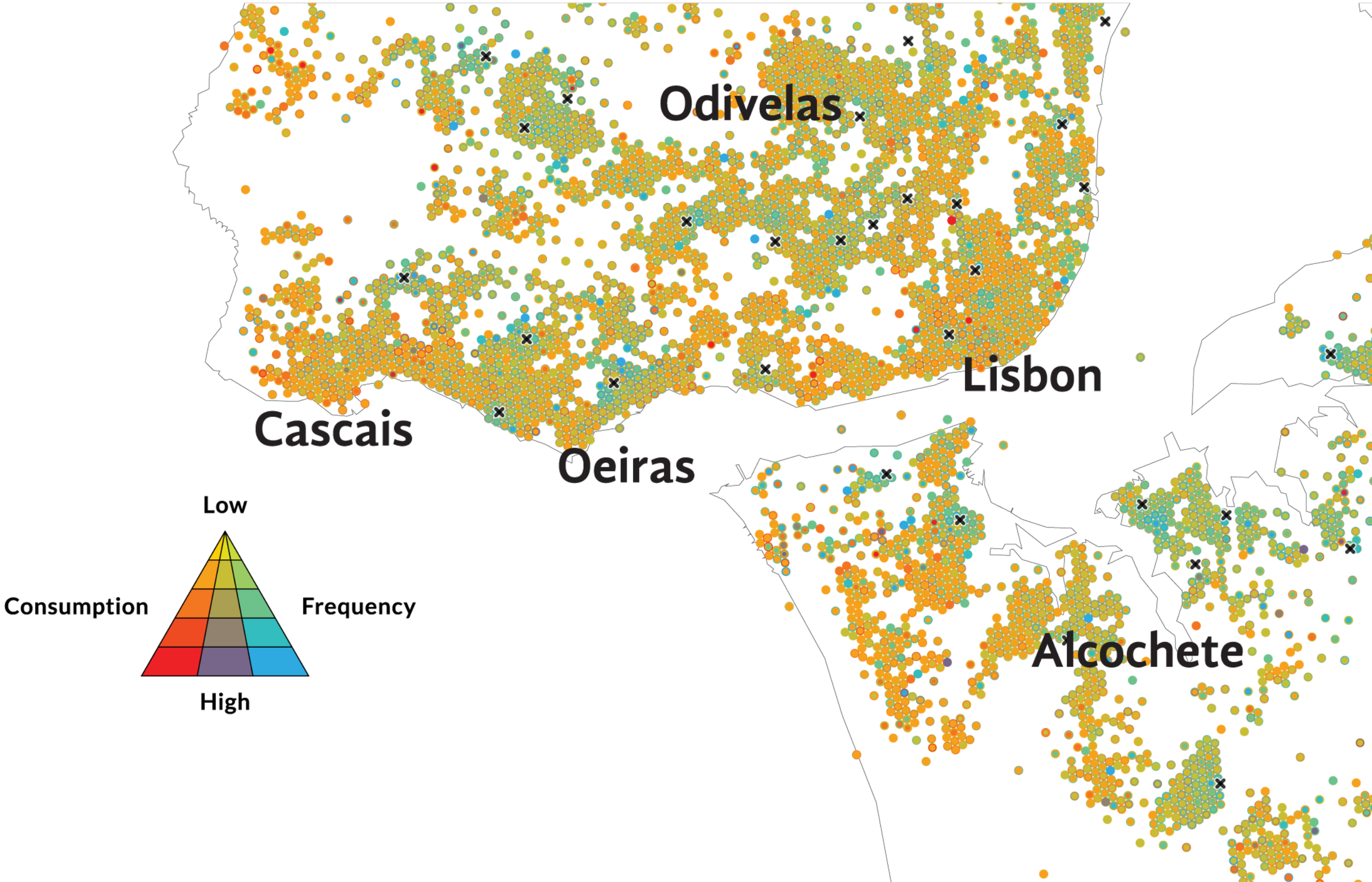


Choropleth map

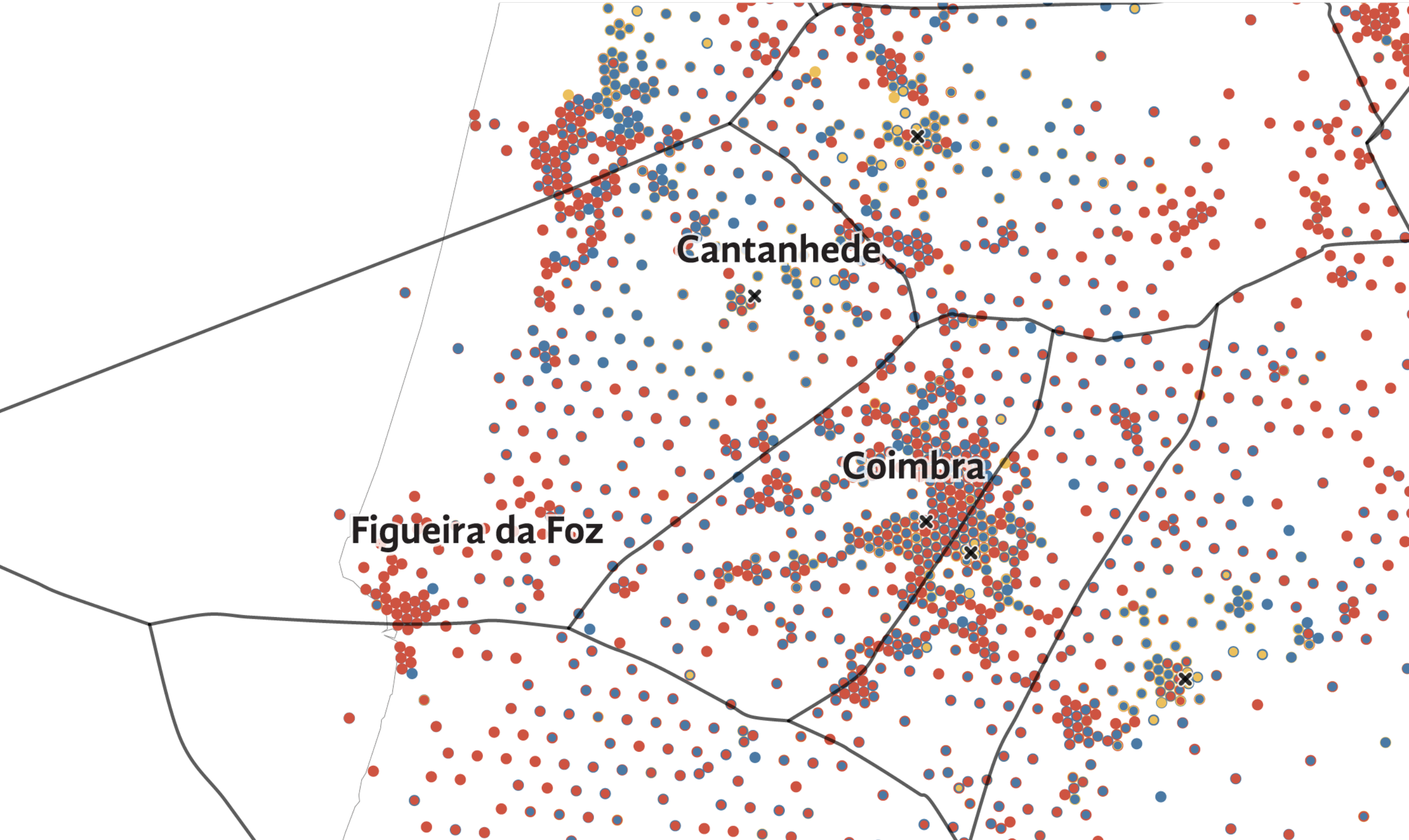


Findings

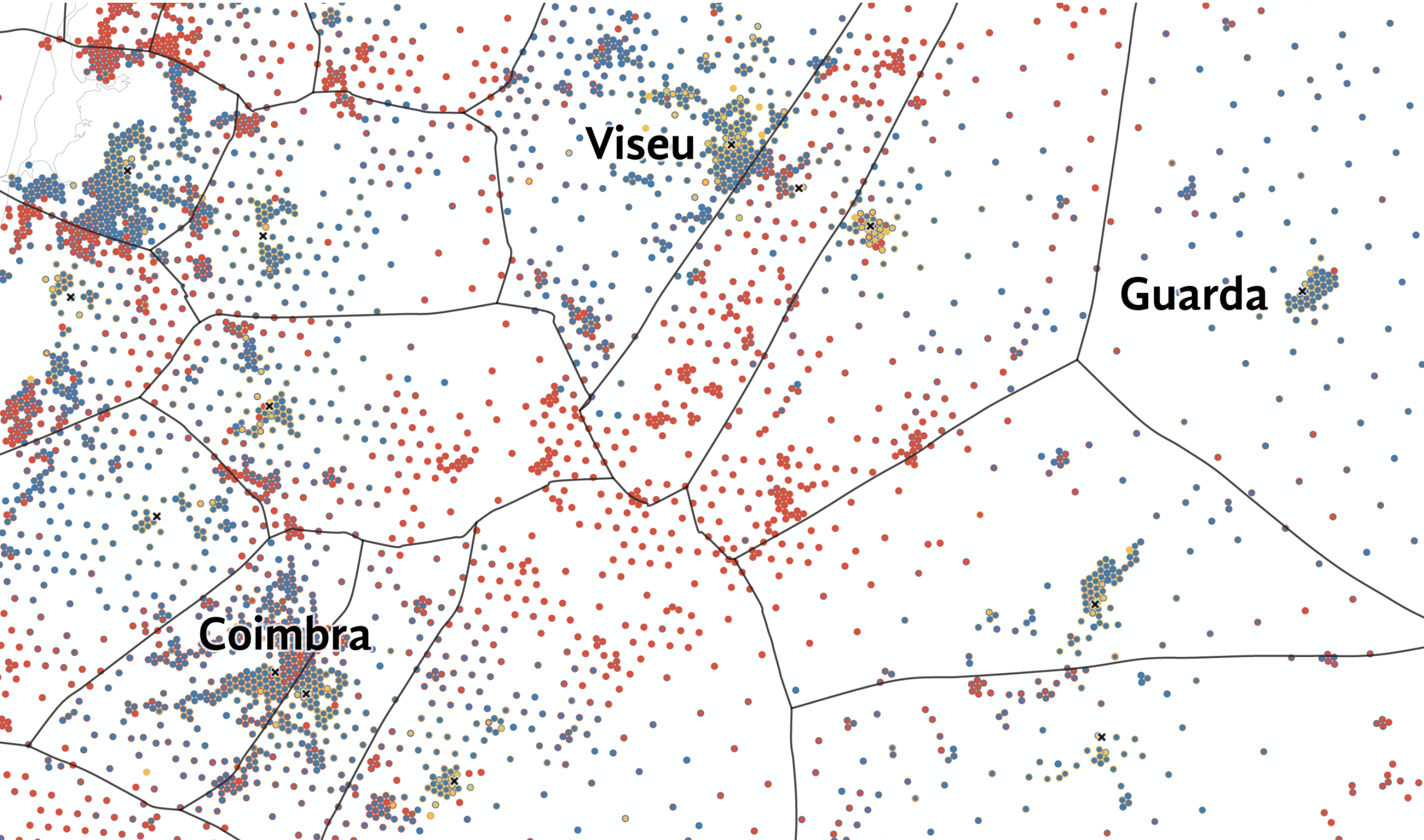
Consumption vs frequency



Frequency of visits 1



Frequency of visits 2



Prototype Demo

Choose a month to visualize

Dezembro 2013
Fevereiro 2014
Setembro 2013
Outubro 2012
Dezembro 2012
Outubro 2013
Novembro 2012
Setembro 2012
Agosto 2012
Novembro 2013
Julho 2012
Março 2014
Abril 2014
Junho 2013
Janeiro 2013
Agosto 2013
Julho 2013
Abril 2013
Fevereiro 2013
Maio 2012
Maio 2013
Junho 2012
Março 2013
Janeiro 2014

Lessons learned

- Actively use visualization throughout the project;
- Hexagonal grids, although not popular, are good alternative to traditional squared grids;
- Even small graphics that provide additional information can be powerful tools;
- When working with geo-referenced data take into considerations all the imaginable characteristics of our surroundings;
- Working with Big data requires appropriate technological solutions;

Thank you

Bibliography

Tools

- Processing (https://processing.org/)
- QGIS (https://qgis.org/en/site/)
- PostgreSQL (https://www.postgresql.org/)
- pgRouting (https://pgrouting.org/)

Literature

- Isabel Meirelles, “Design for Information: An Introduction to the Histories, Theories, and Best Practices Behind Effective Information Visualizations”, 2013
- Borden Dent, “Cartography: Thematic Map Design”, 1990
- Fritz C. Kessler, Hugh H. Howard, Robert B. McMaster, and Terry A. Slocum, “Thematic Cartography and Geovisualization”, 1999