Cartography M.Sc.

Enhancement of Density Visualization using Dot Density Maps

Mostafa Kheyrollahi

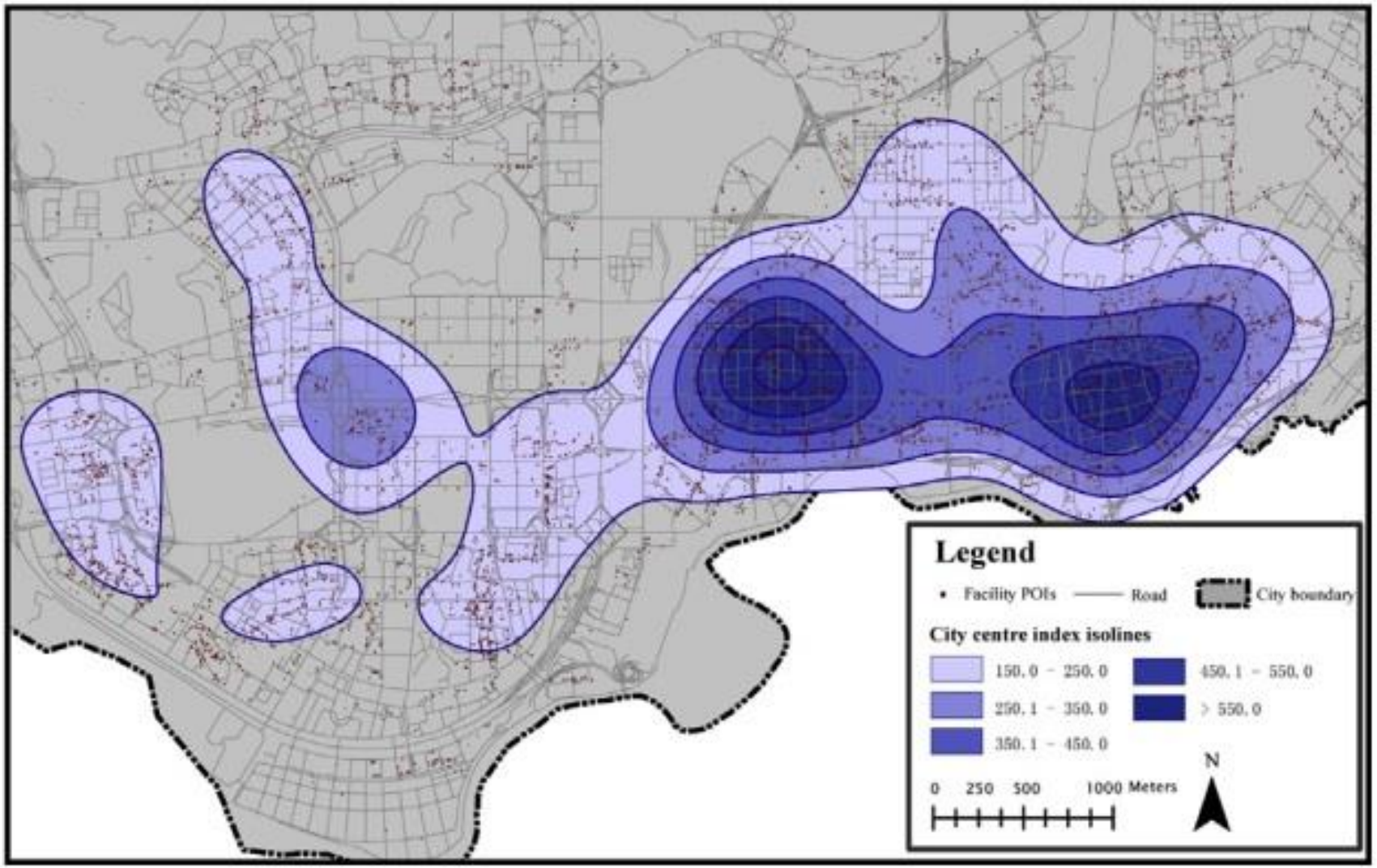
The presentation at a glance

- Problem statement
- Motivation
- Objectives
- Research Questions
- Previous research and related works
- Conventional dot map algorithm
- Conventional dot map results
- Graduated dot map algorithm
- Graduated dot map results
- User test
- Answers to research questions
- Conclusions

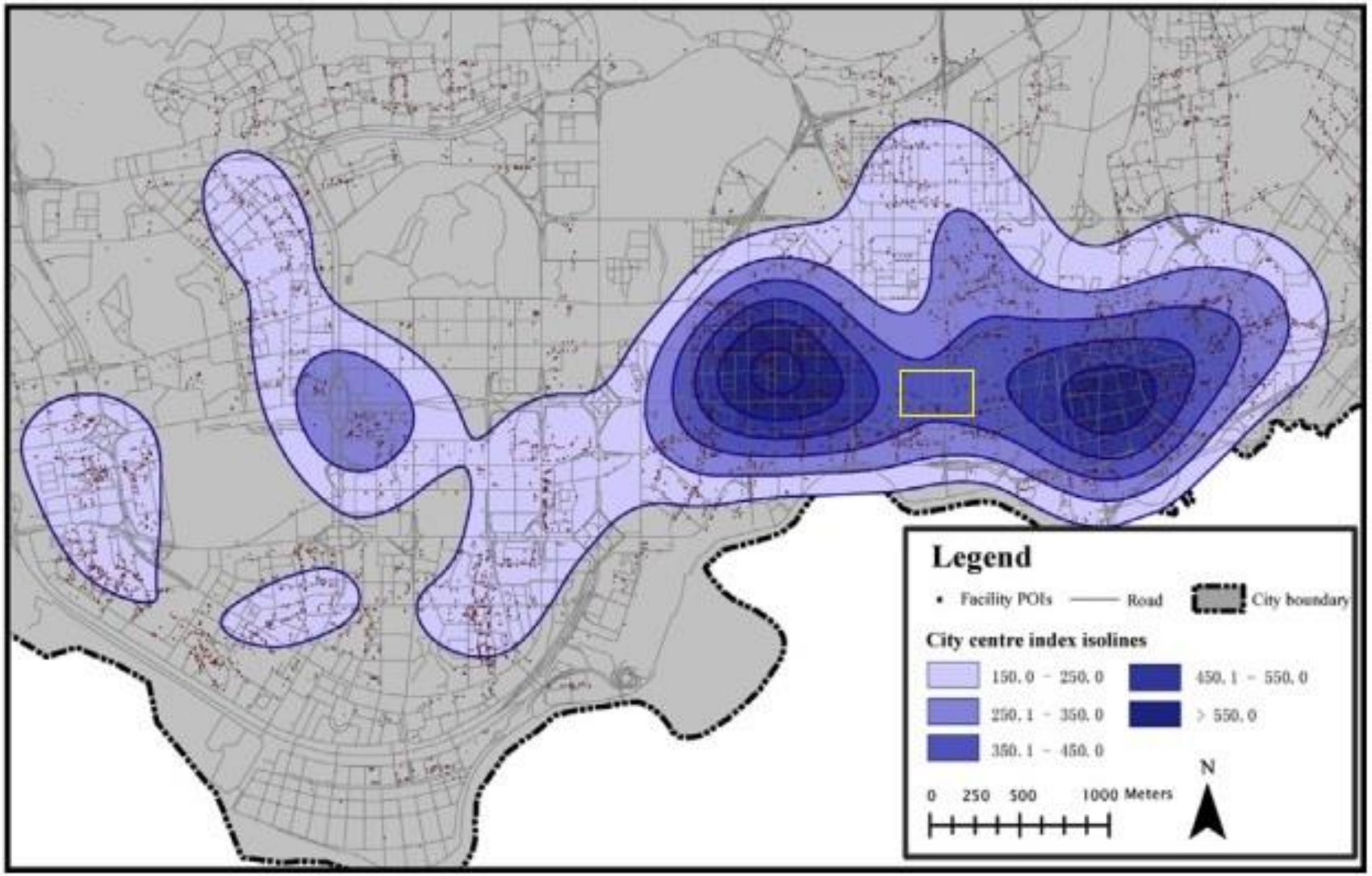
Problem statement

A heatmap product from Kernel Density Estimation (KDE) function is one of the popular methods to monitor the density of a phenomenon which is not uniform throughout the studied area. However, it has some drawbacks.

- No quantitative estimation from the colour scheme
- No perception for the actual values of a phenomenon due to cryptic result of KDE



Heatmap from running KDE on POI to determine CBD
Li et al. (2014)



Specific area in the Heatmap from running KDE on POI to determine CBD

Motivation

How to use KDE results to make another type of map to overcome the drawbacks.

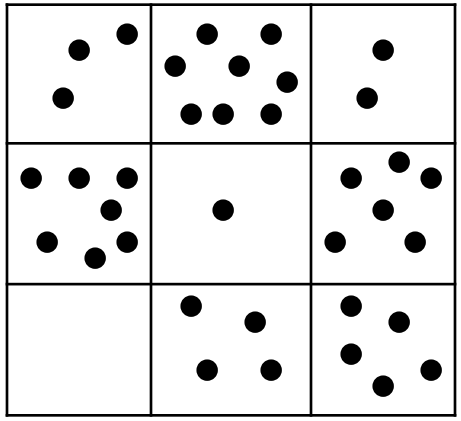
The proposed type of map is a dot map.

The map should overcome the problems to give a more quantitative estimation to the user.

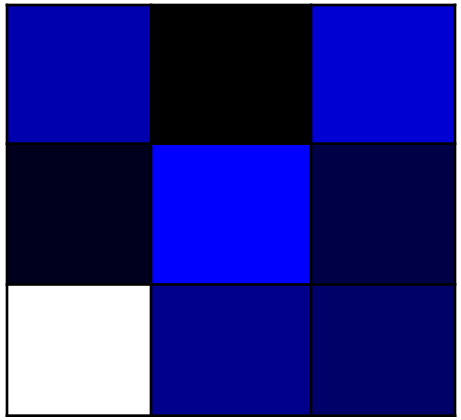
How to use the KDE product?

30	80	20
70	10	60
0	40	50

Raster



Dot map



Heat map

Objectives

- Making two types of dot map: graduated and conventional dot map
- Using the KDE output (raster) as an input to the dot mapping algorithms
- The maps should be more quantitative and not cryptic
- The maps will show the overall density much better than KDE heat maps

Research Questions

- What are the most important criteria to make a dot map?
- What are the difference between KDE heatmaps and dot maps?
- Do final dot maps provide quantitative information to the user?
- What are the Advantages and disadvantages of graduated and conventional dot maps?

Previous research and related works

1. General review of dot mapping methods and applications
2. A review of graduated dot maps
3. Related works to KDE and heatmaps from KDE

1. General review of dot mapping methods and applications

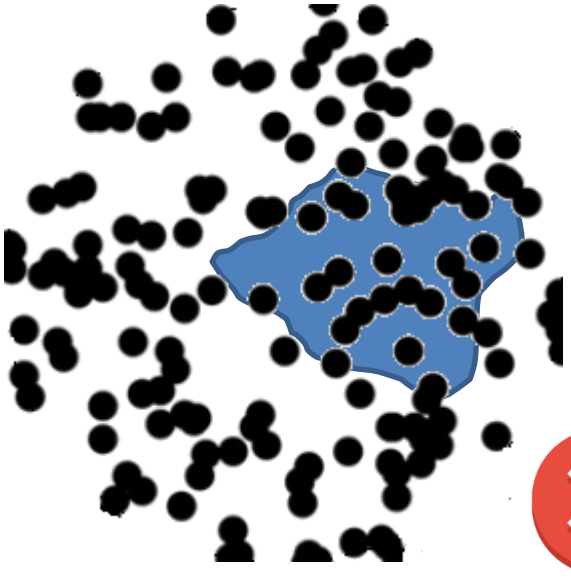
Topics:

- Principles of dot mapping
- A number of techniques of dot mapping and their pro and cons
- Research about applications of dot mapping

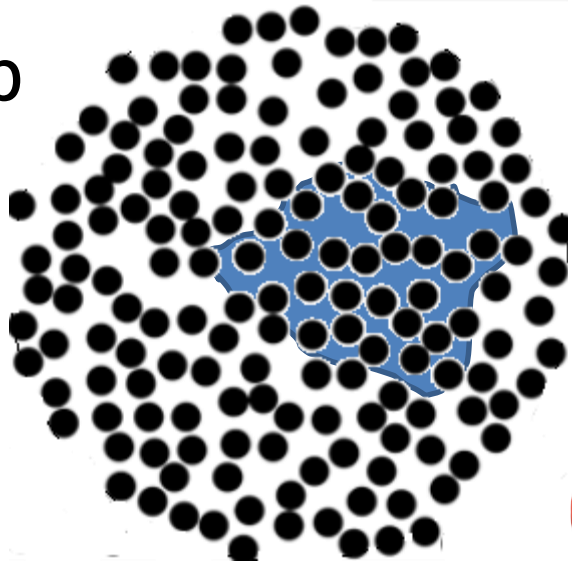
Results:

- Four main parameters of a dot map: **dot value**, **dot size**, **dot placement** and **true regions** for placement
- Size of dots has an inverse relation with the number of dots in a cell
- Increasing covered area by dots in a cell leads to more overlap of dots
- More regular dot placement is far superior to totally random dot placement

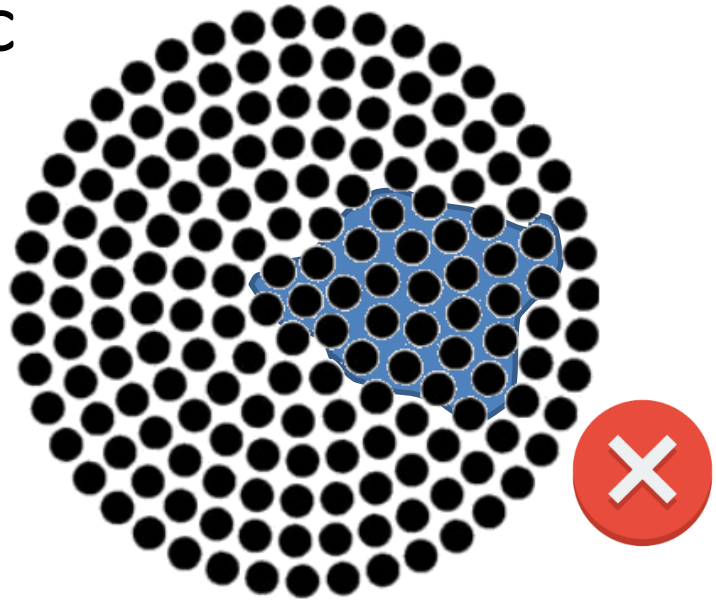
a



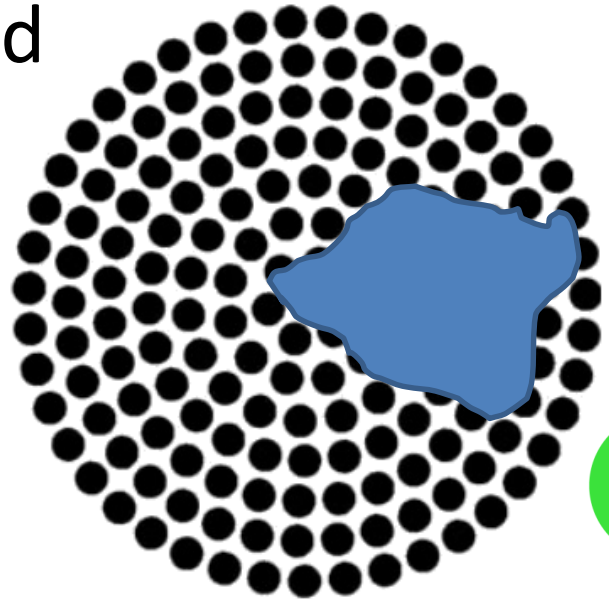
b



c



d



Dot value: **n points**
Dot size: **n (m)**

2. A review of graduated dot maps

Topics:

- Principles of graduated dot maps
- Clarifying the difference between a graduated dot map and a proportional symbol map
- Prior research on dot maps

Results:

- Difference between proportional symbol maps and graduated dot maps.
- Superiority of graduated dot maps to conventional dot maps: Reduces the number of dots, reduces overlapping and gives a more quantitative view to the user
- More accuracy of quantitative overview to a user than conventional dot maps, and proportional circle (symbol) maps
- If the only goal is to get the spatial density patterns, a conventional dot map is appropriate.

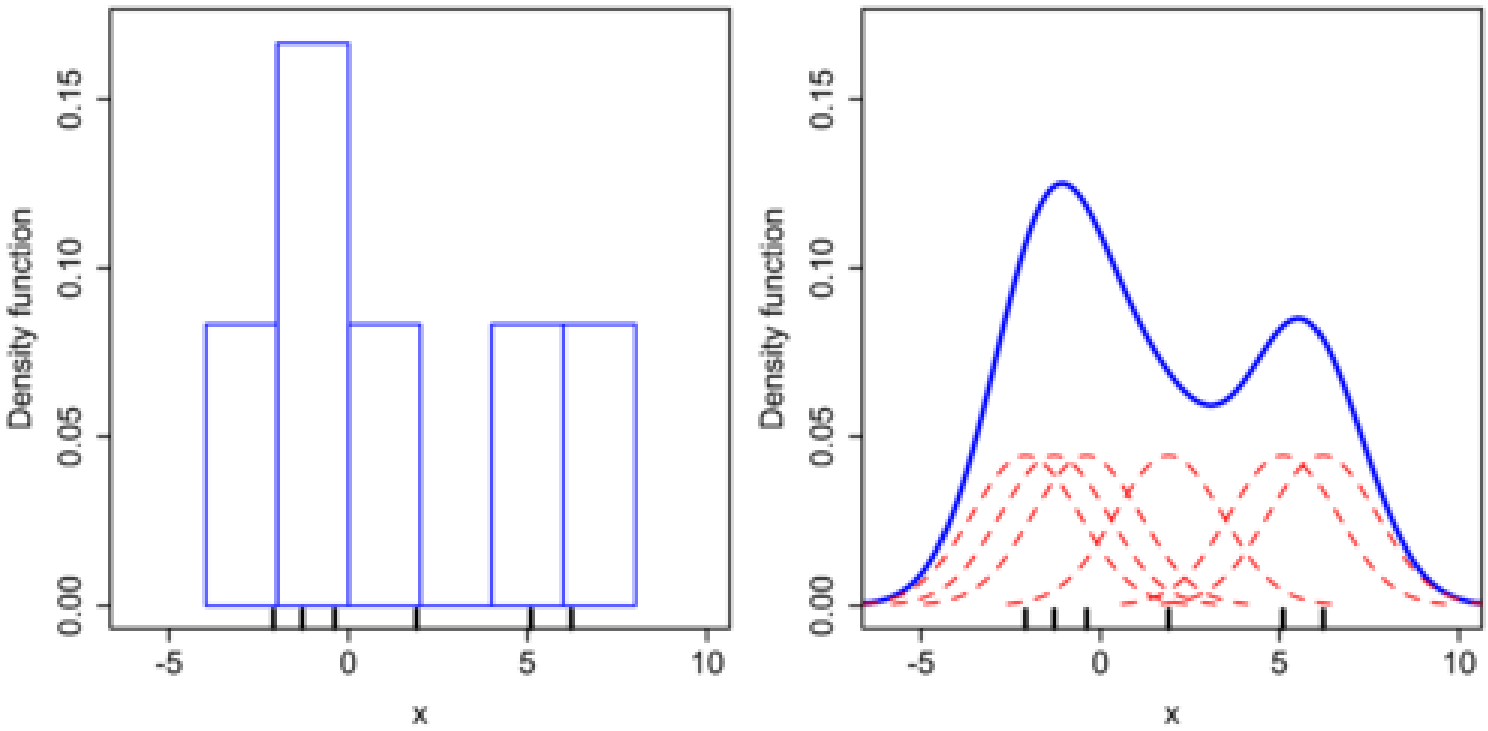
3. Related works to KDE

Topics:

- Principles of Kernel Density Estimation
- Application of KDE to monitor density in different case studies
- Presentation heatmaps as a possible output of KDE

Results:

- KDE is a non-parametric way to estimate the probability density function of a random variable.
- It is used for point or linear datasets.
- The result can be a heat map with a range of colors and some isolines.
- The KDE result is a matrix (raster) with pixels which have specific values.



KDE Diagram

https://en.wikipedia.org/wiki/Kernel_density_estimation

Methodology

Two developed algorithms are presented:

- Conventional dot mapping
- Graduated dot mappping

Conventional dot map algorithm

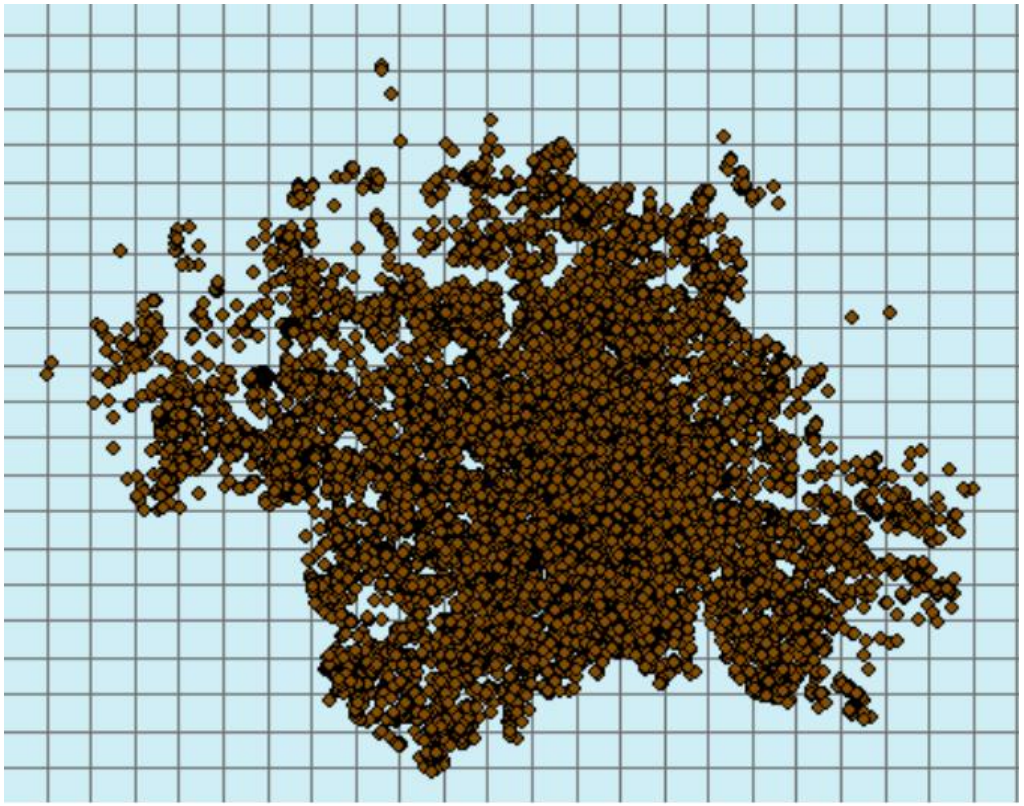
Main steps:

1- Running KDE function on the input data to get a raster and recognize the value of each pixel

- Choosing a metric coordinate system
- Choosing the function type (Gaussian)
- Determination of the bounding box of the original data
- **Setting the resolution (number of cells)**

Conventional dot map algorithm

2- Converting the raster to a vector and removing the cells which lie outside of the studied area



The vector and the original data

Conventional dot map algorithm

3- Assigning a number of equal sized dots to each cell.

- Normalization of cell values to set them between 0 and 1 (z_i)
- Setting the proportion of dot coverage (P_i) for each cell with the help of maximum (P_{max}) and **minimum proportion of dot coverage ($\boldsymbol{P_{min}}$)**

$$P_i = z_i\,(P_{max} - P_{min}) + P_{min}$$

- Calculation of the **number of dots ($\boldsymbol{N_d}$)** with the help of **dot size ($\boldsymbol{A_d}$)** and **cells' size ($\boldsymbol{A_c}$)**

$$N_d = \frac{P_i * A_c}{A_d}$$

- Calculation of the **dot value**

Conventional dot map algorithm

3- Assigning a number of equal sized dots to each cell.

A sample solution:

If z_i = 0.74

& P_{min} = 0.02

& P_{max} = 0.50

& A_c= 100,000 m^2

& A_d = 5,000 m^2 :

$$P_i = 0.74\,(0.50 - 0.02) + 0.02 = 0.37$$

$$N_d = \frac{0.37 * 100{,}000}{5000} = 7.4 \sim 7$$

Conventional dot map algorithm

4- Placement of dots

- No dot overlap or dot coalescence between and within cells
- Dots are placed separately in each cell.
- The placement is a pseudo random process.

Conventional dot map algorithm

Main variables:

- Dot size
- Resolution (cell size)
- Minimum proportion of dot coverage

They are set in an iterative process

Conventional dot map algorithm

Main variables:

Cell size  Resolution

Cell size  Dot size

Resolution  Dot size

Dot size  Dot value

Number of dots  Dot size & value

Conventional dot map algorithm

Main variables:

Minimum proportion of dot coverage sets the number of dots in the emptiest cell(s) to avoid zero number of dots in those cells.

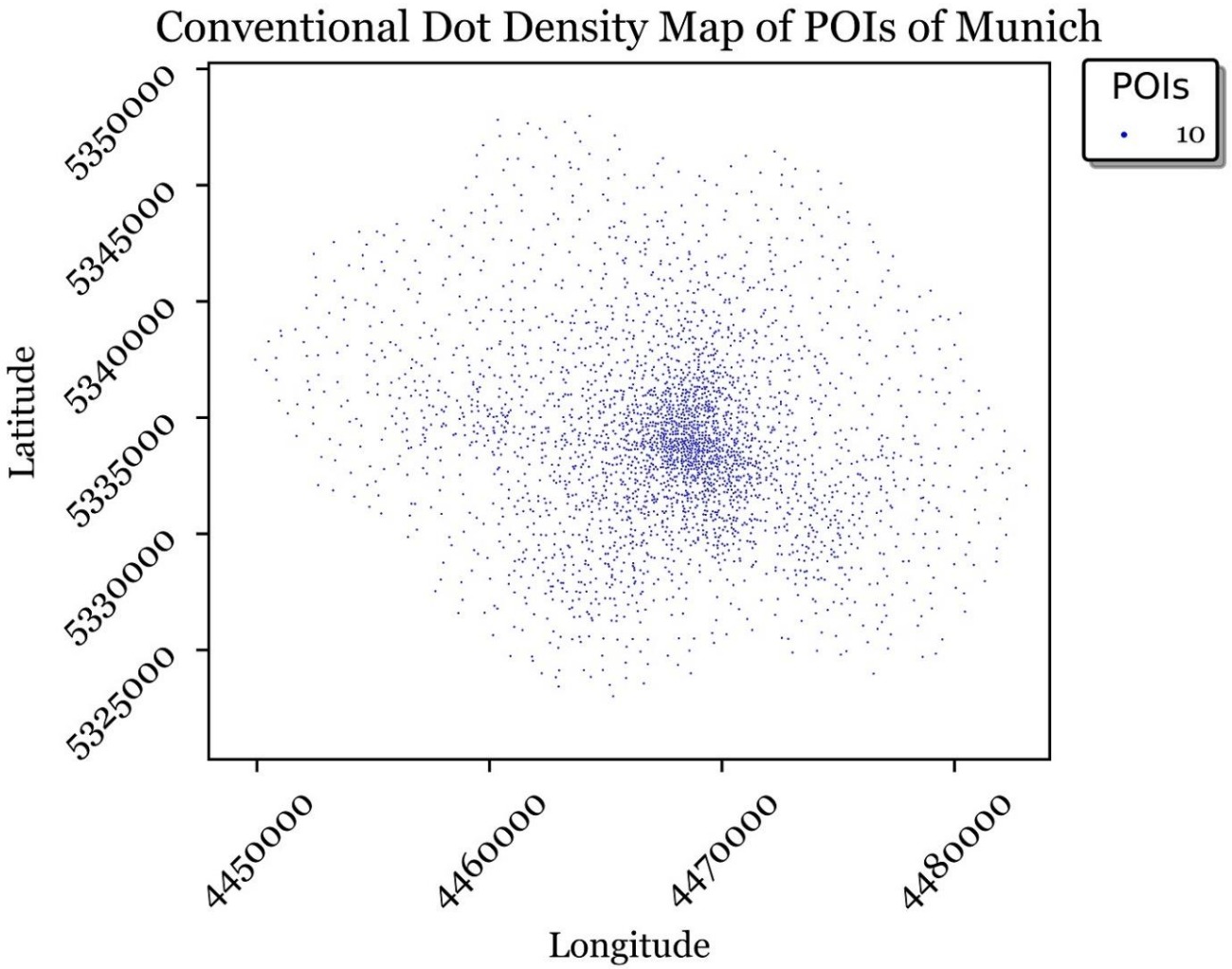
It is usually between 0.01 to 0.1.

Conventional dot map algorithm

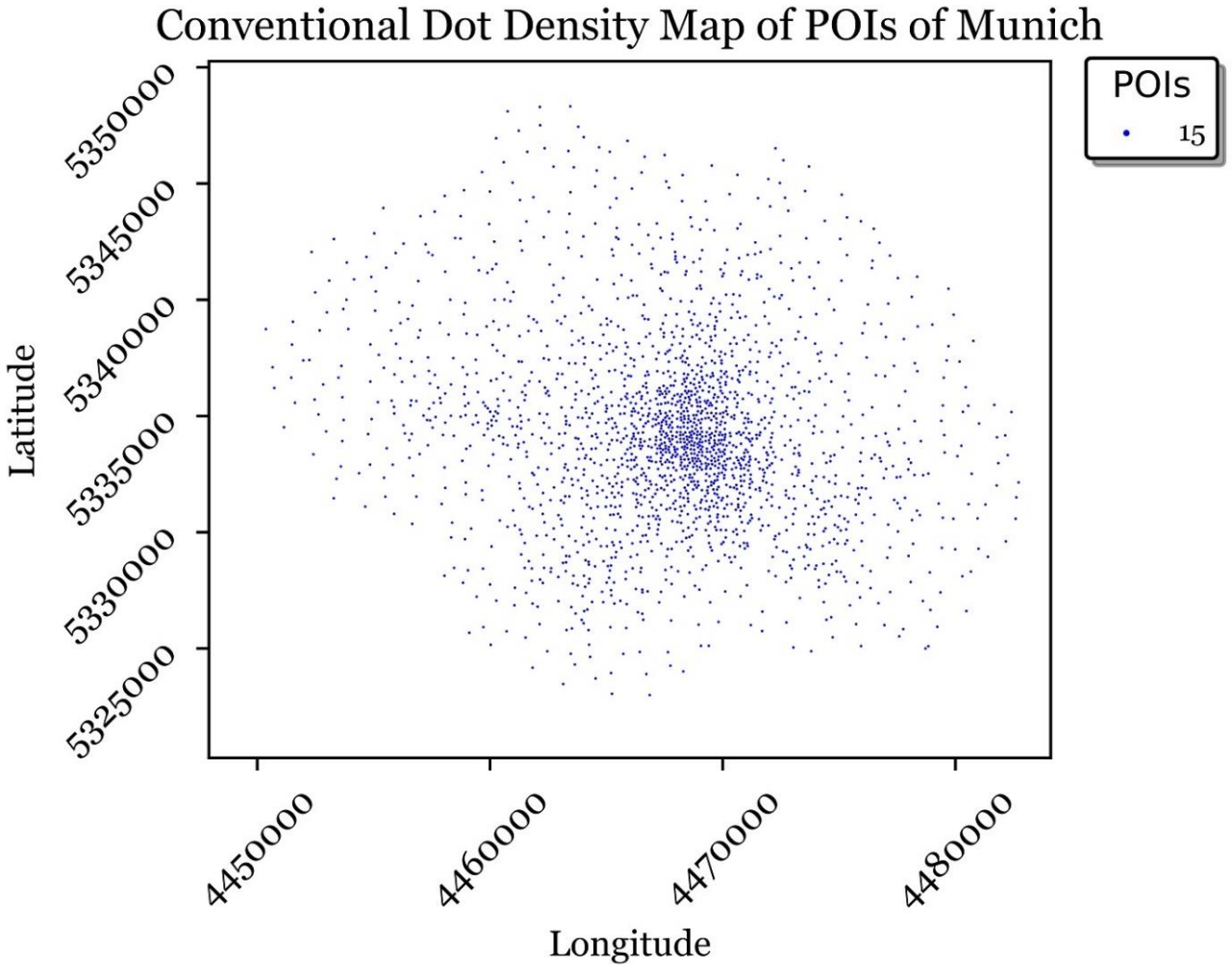
Used data:

- The Points of Interest (POIs) dataset of Munich obtained from OpenStreetMap
- Tweets of the Oktoberfest in Munich obtained through Twitter API

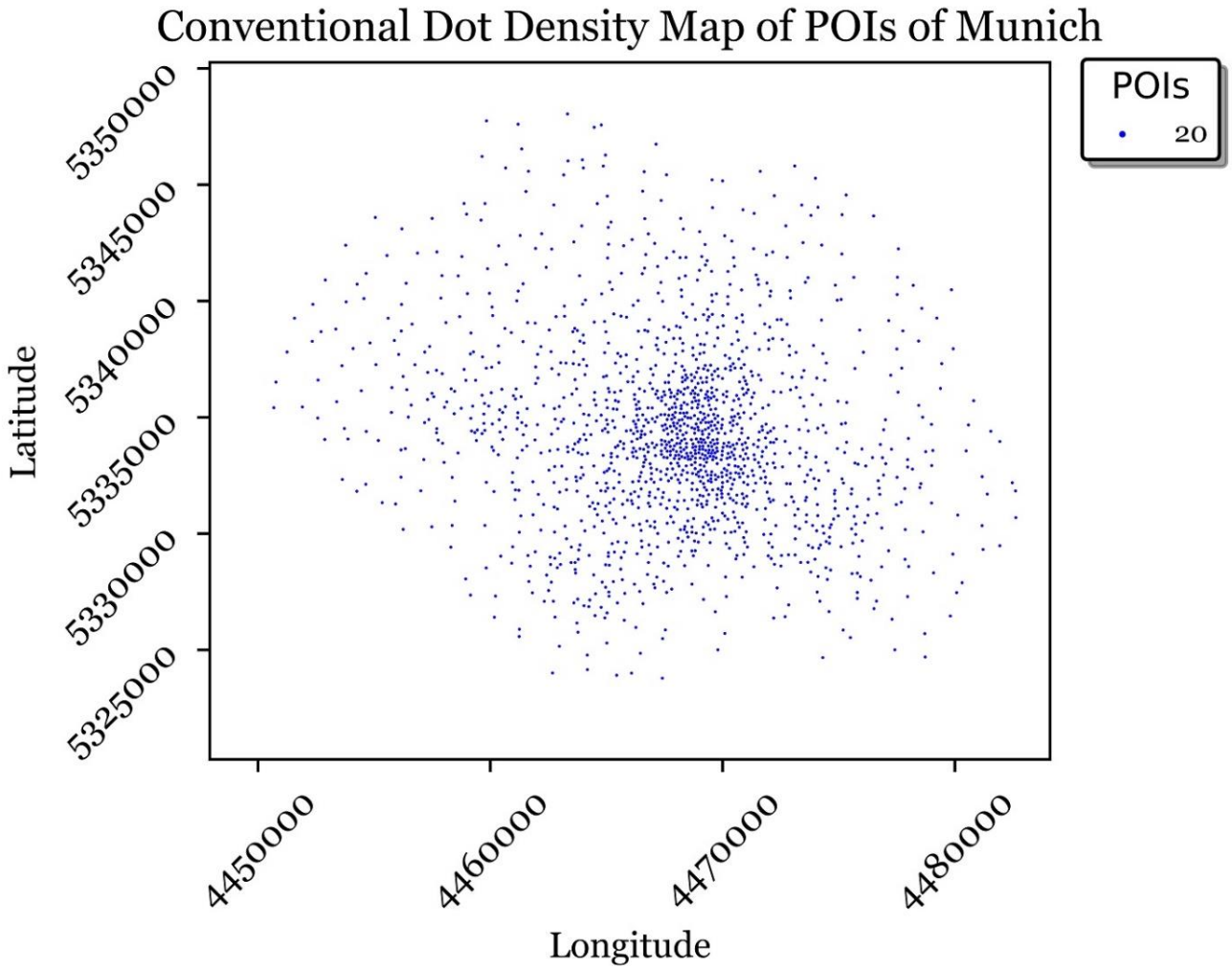
Conventional dot map results



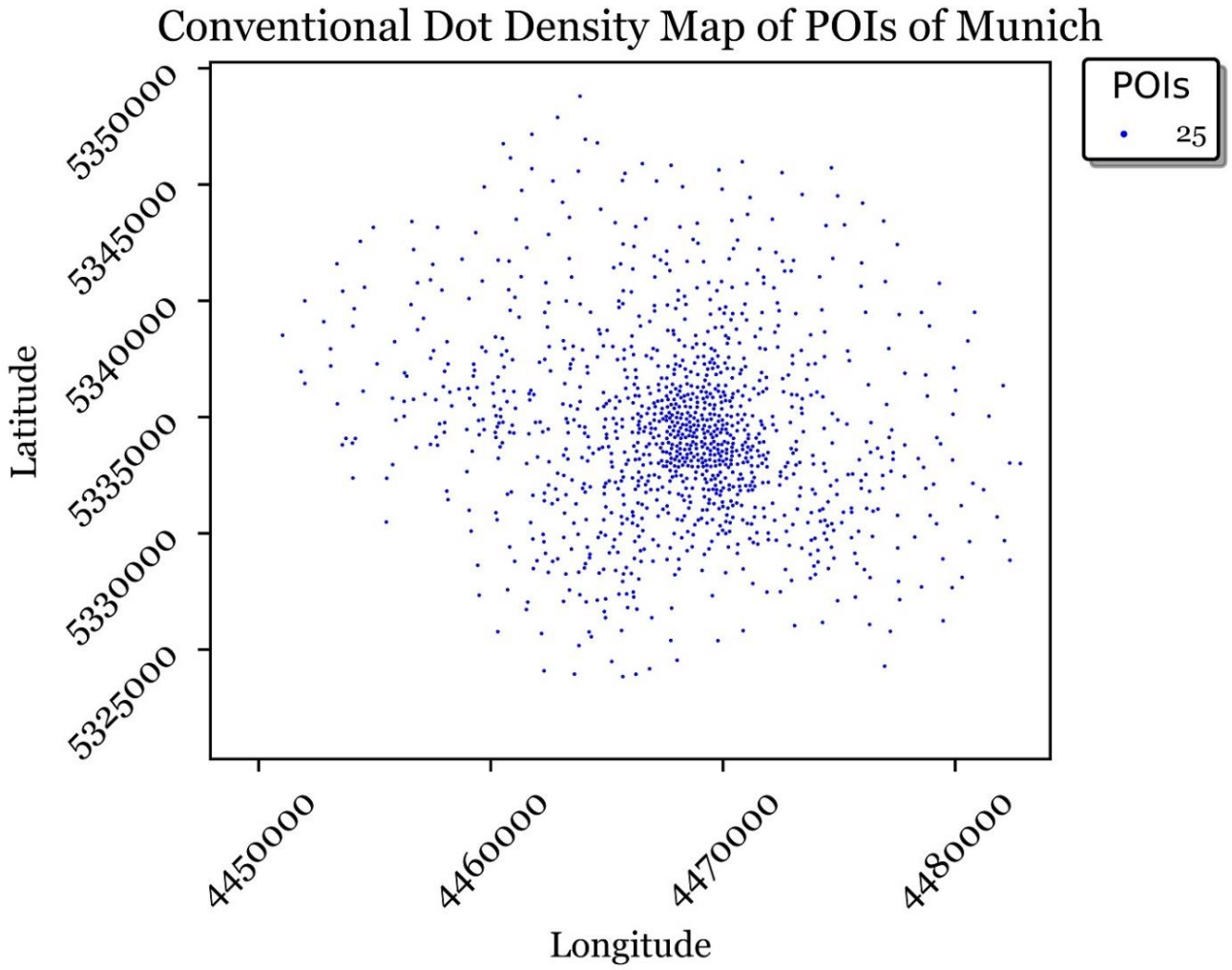
Conventional dot map results



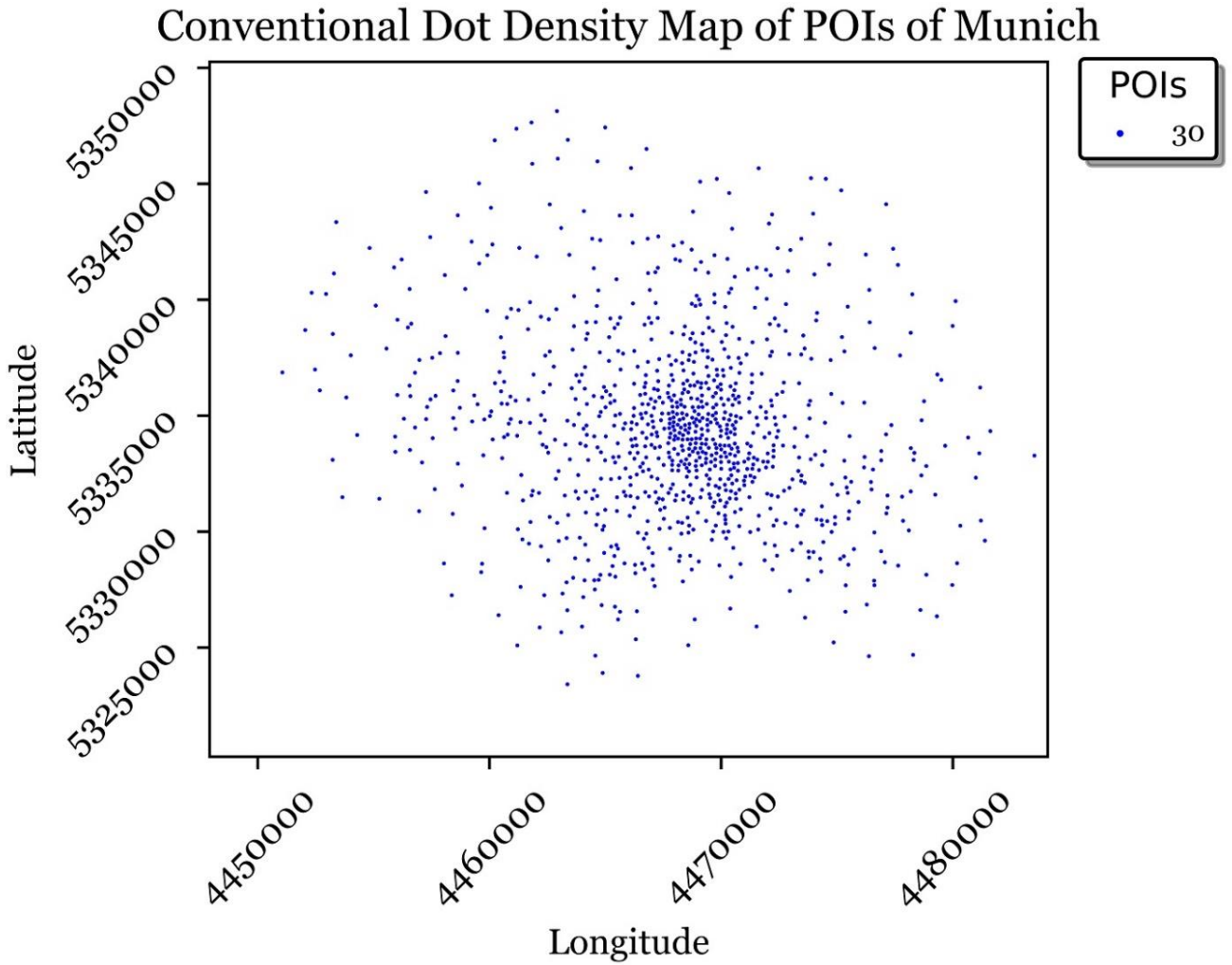
Conventional dot map results



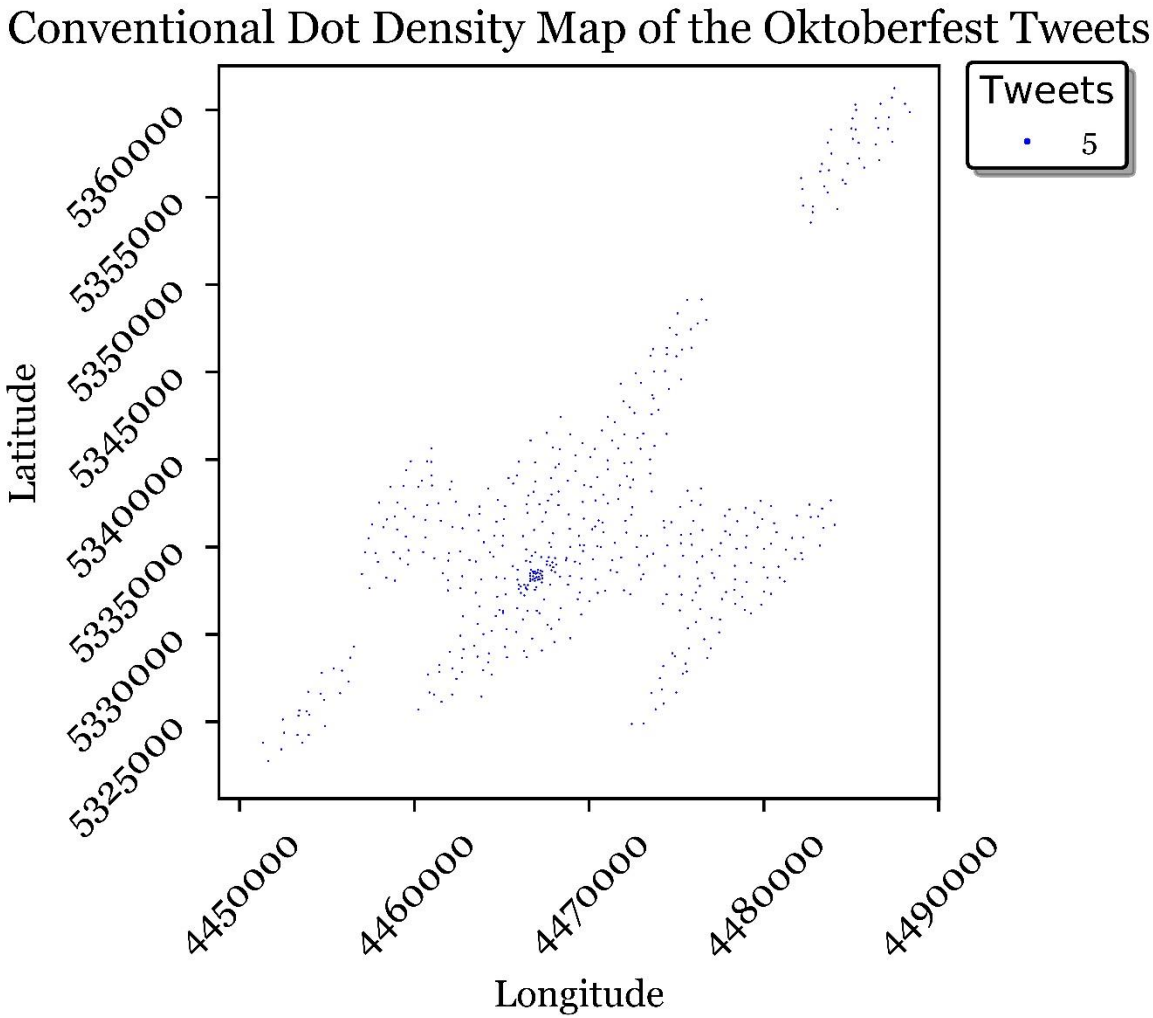
Conventional dot map results



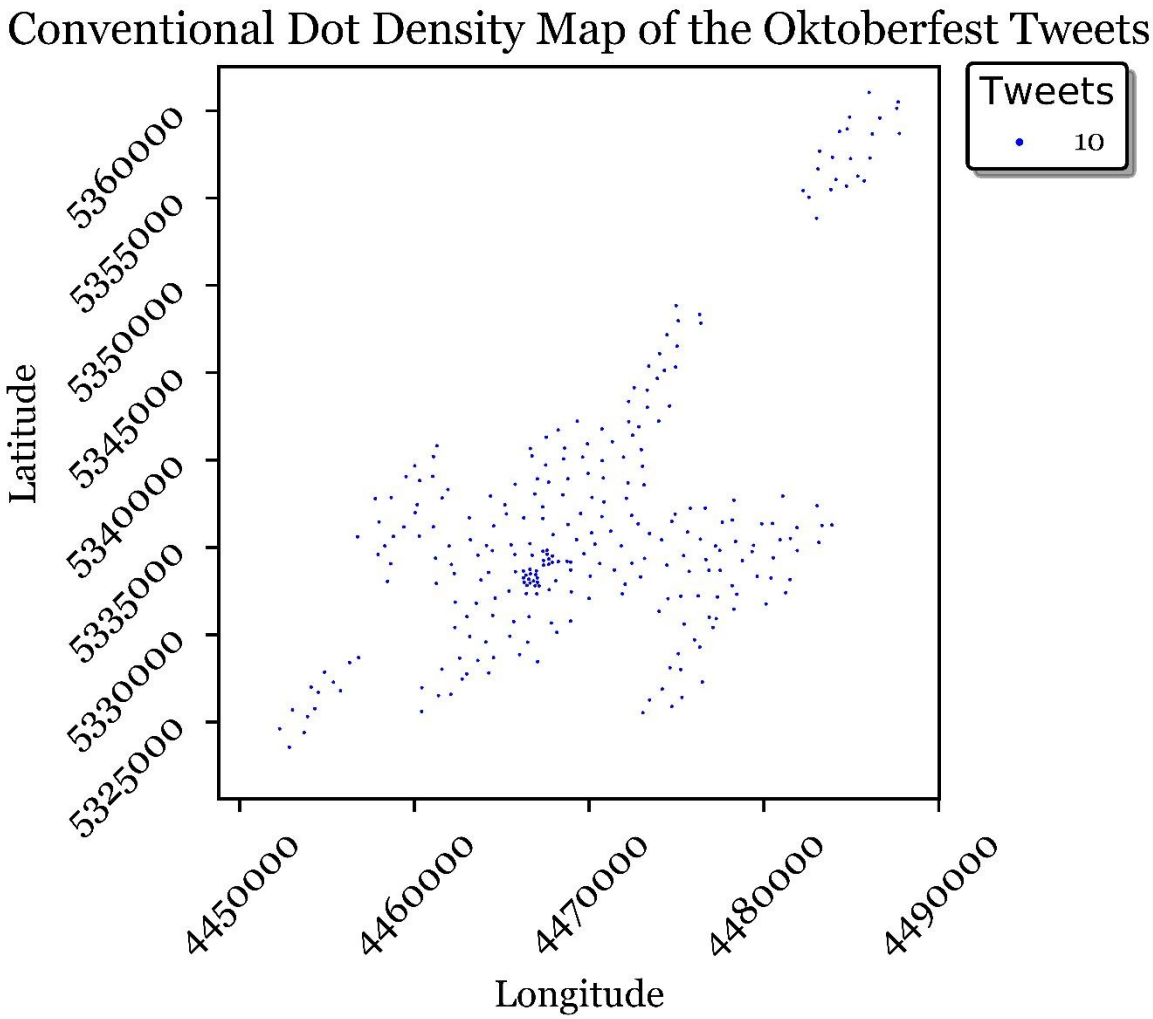
Conventional dot map results



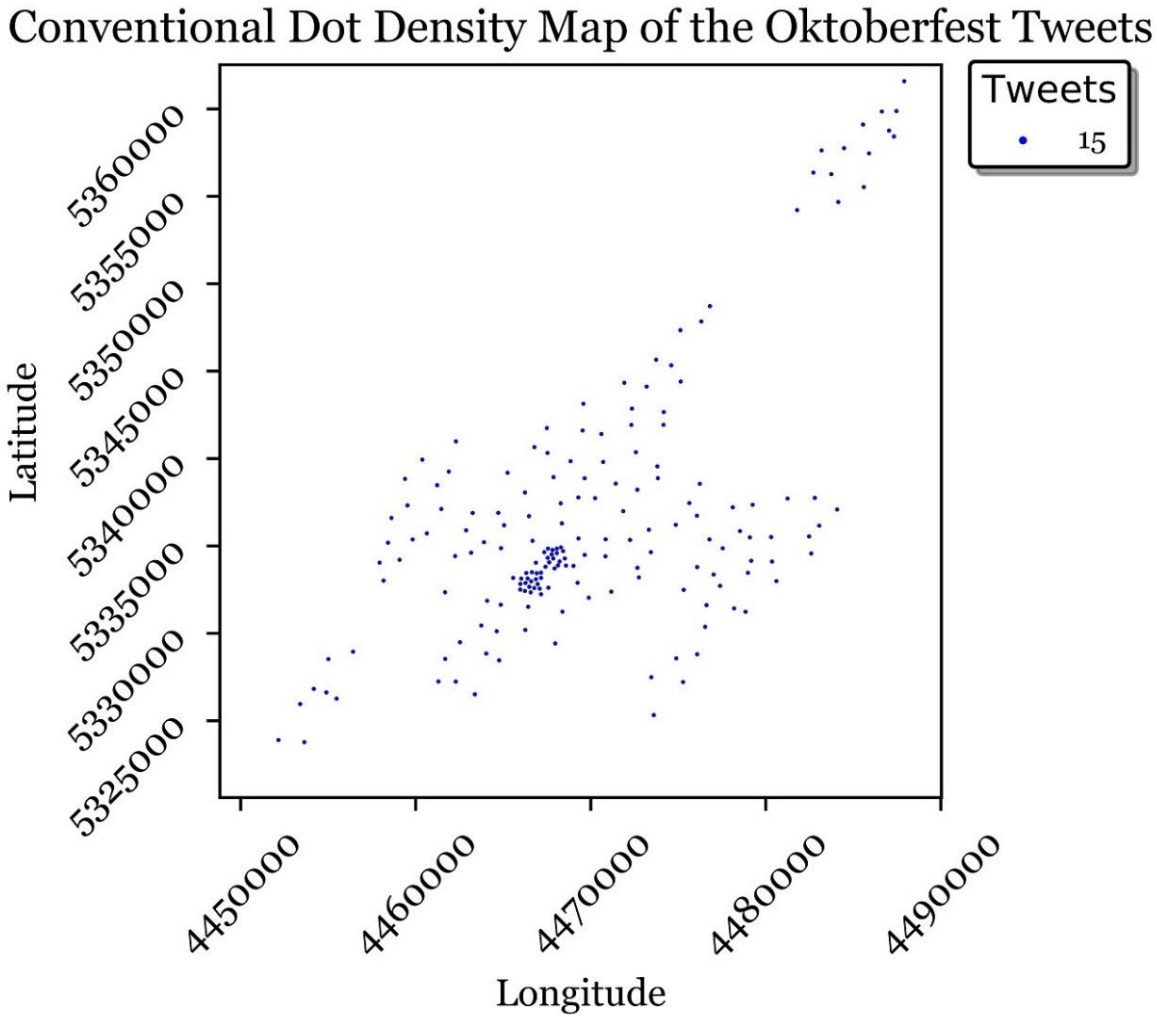
Conventional dot map results



Conventional dot map results



Conventional dot map results



Conventional dot map results

Summary of variables

Data	Resolution	Dot size	Dot value	Pmin
POI	42 * 42	45	10	0.01
POI	35 * 35	56	15	0.01
POI	29 * 29	65	20	0.01
POI	25 * 25	79	25	0.01
POI	23 * 23	85	30	0.01
Tweets	52 * 52	60	60	0.02
Tweets	39 * 39	100	100	0.02
Tweets	30 * 30	120	120	0.02

Summary of variables in the created conventional dot maps

Graduated dot map algorithm

Main steps:

1- Running KDE function on the input data to get a raster and recognize the value of each pixel

- Choosing a metric coordinate system
- Choosing the function type (Gaussian)
- Determination of the bounding box of the original data
- **Setting the resolution (number of cells)**

2- Converting the raster to a vector and removing the cells which lie outside of the studied area

Graduated dot map algorithm

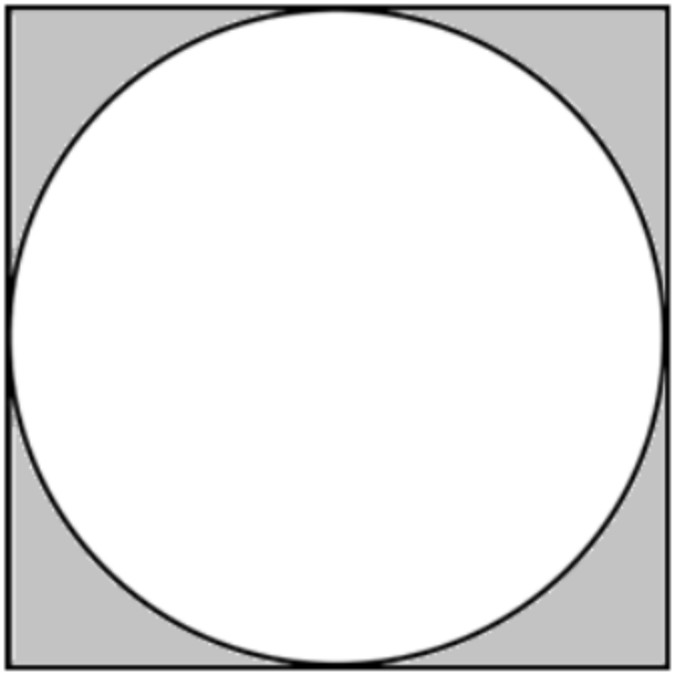
3- Normalization and classification of values

- Normalization of cell values to set them between 0 and 1
- Classification of normalized values to a limited **number of classes** (e.g., 6)

Graduated dot map algorithm

4- Calculation of dot sizes and dot placement

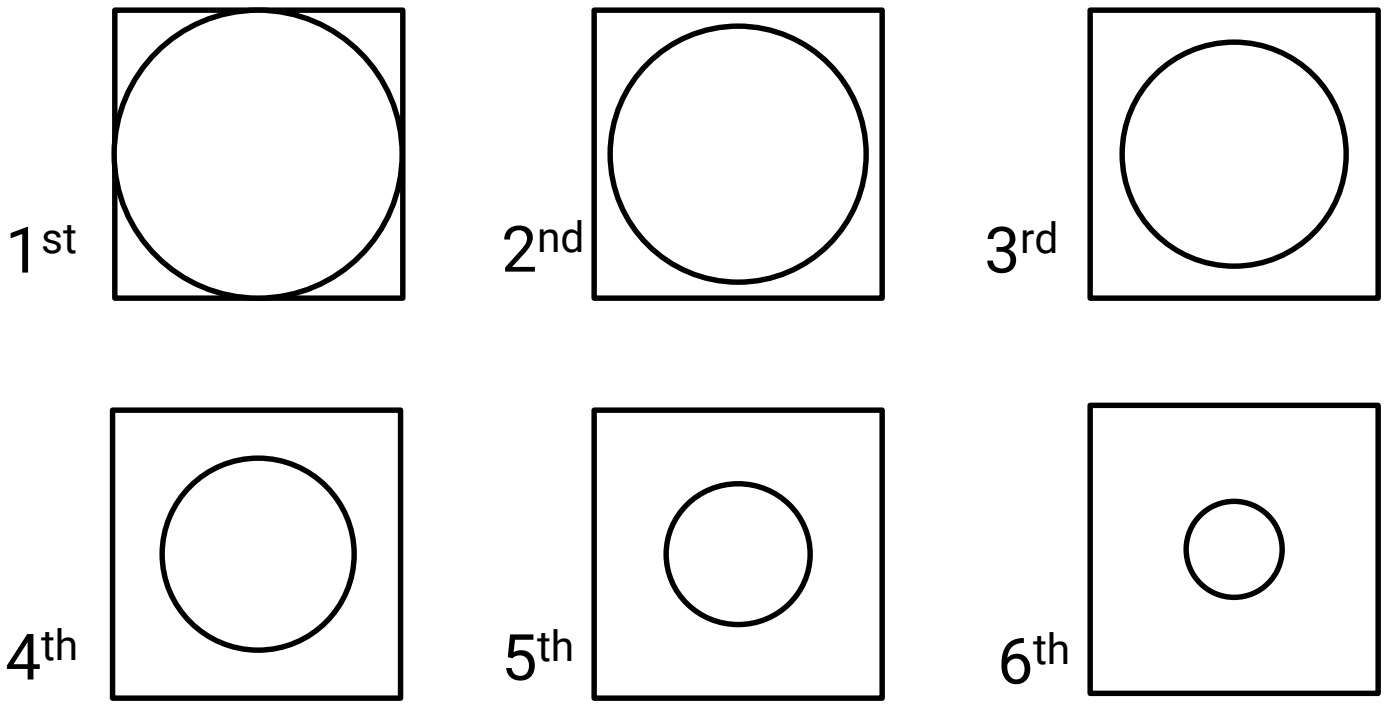
- Calculation of the largest size by the inscribed circle in a cell which is the size of the highest class.



Graduated dot map algorithm

4- Calculation of dot sizes

- Calculation of the size of the dots of the other classes based on the largest size



Graduated dot map algorithm

5- Calculation of the value of each class

Division of the number of points of the original data which lie in the territory of each class by the number of the dots of that class.

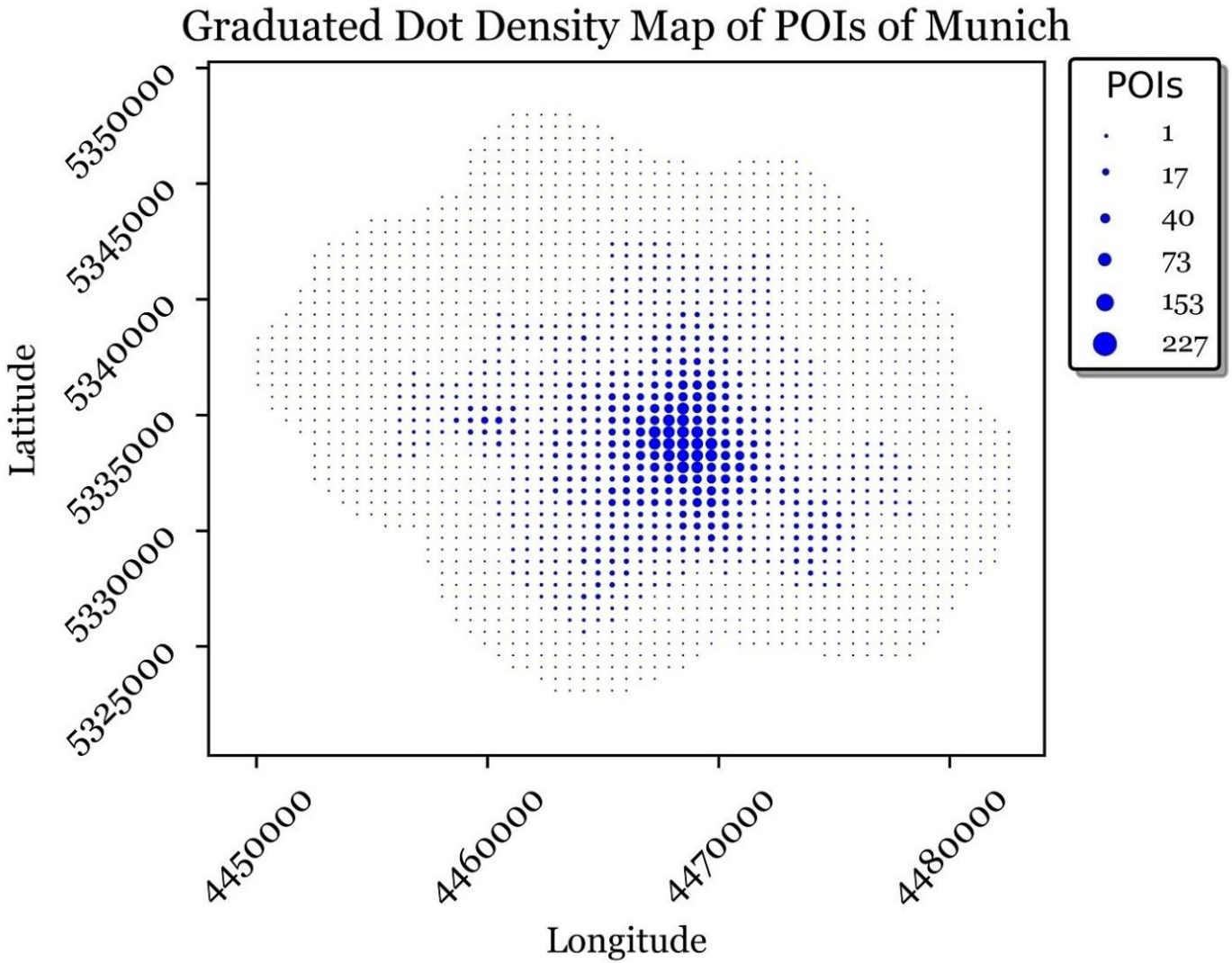
Graduated dot map algorithm

Main variables:

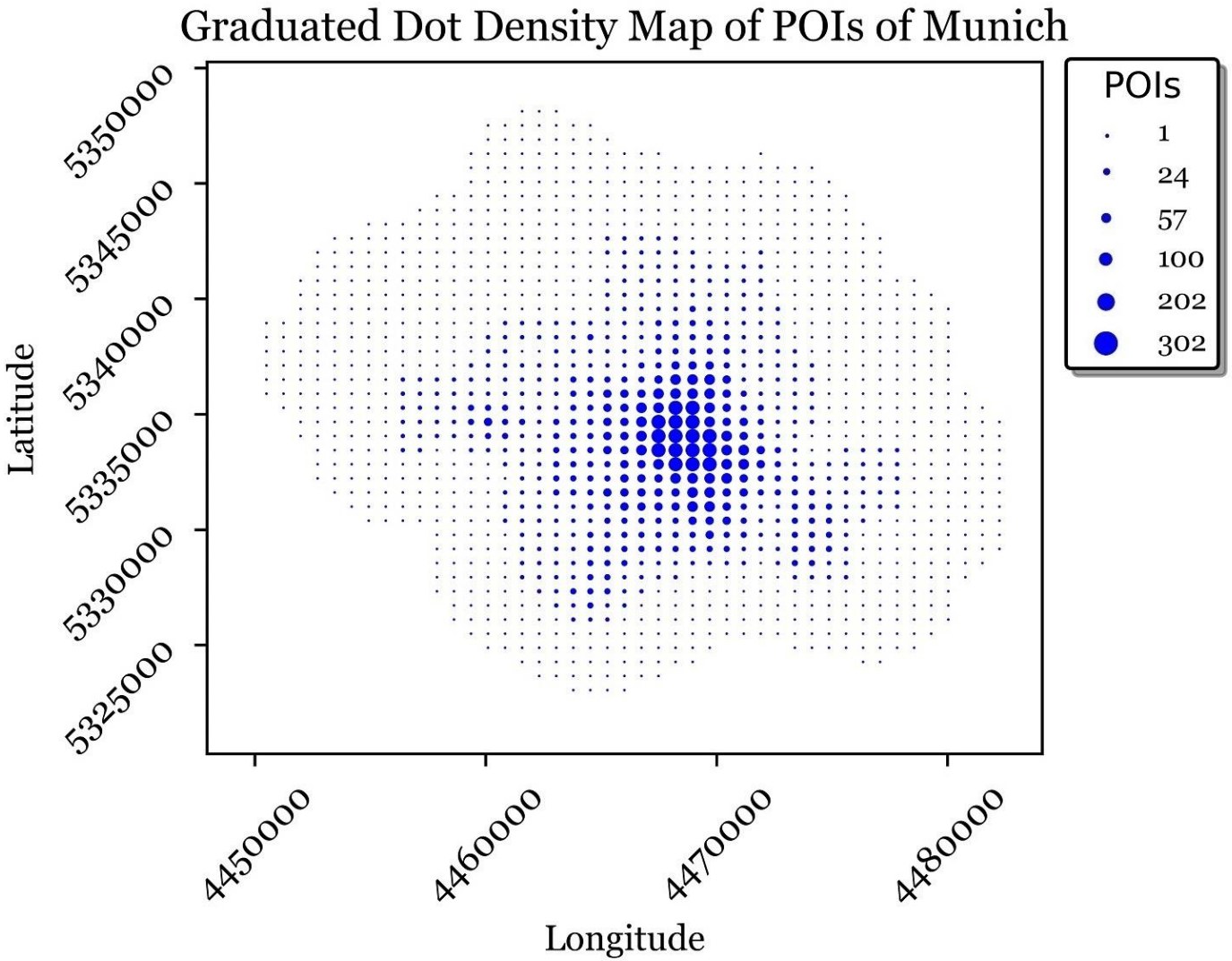
- Resolution (cell size)
- Number of classes

They are set in an iterative process

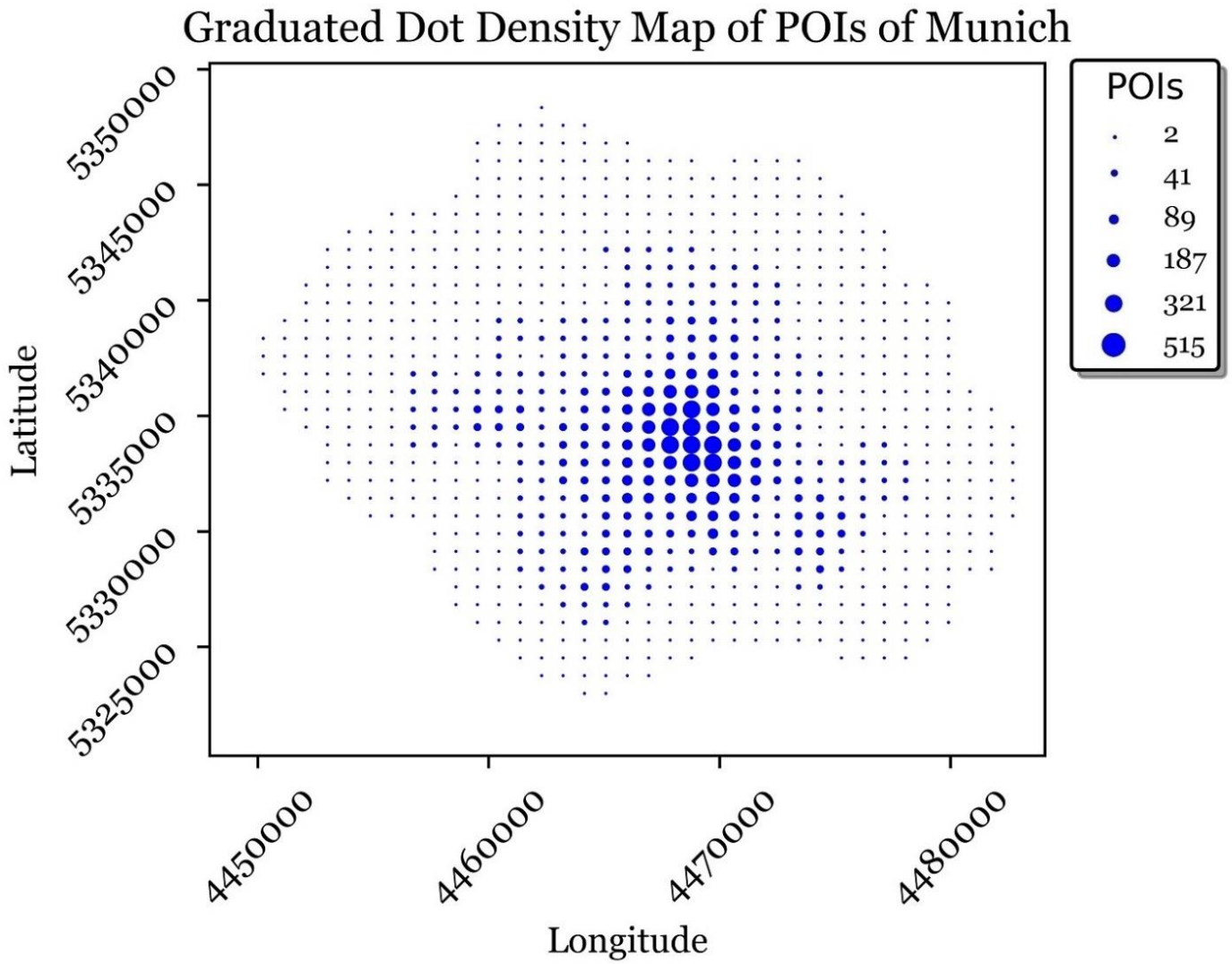
Graduated dot map results



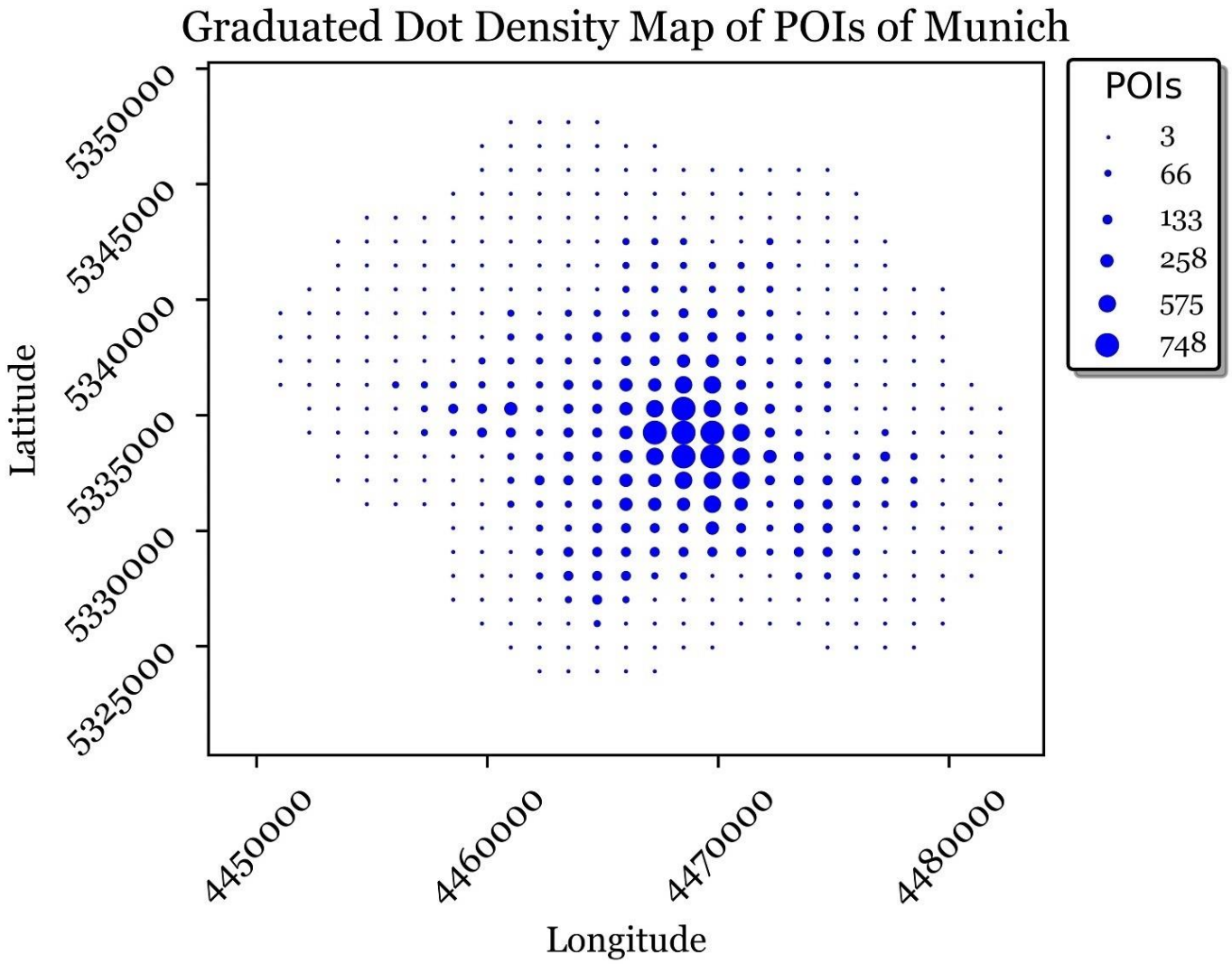
Graduated dot map results



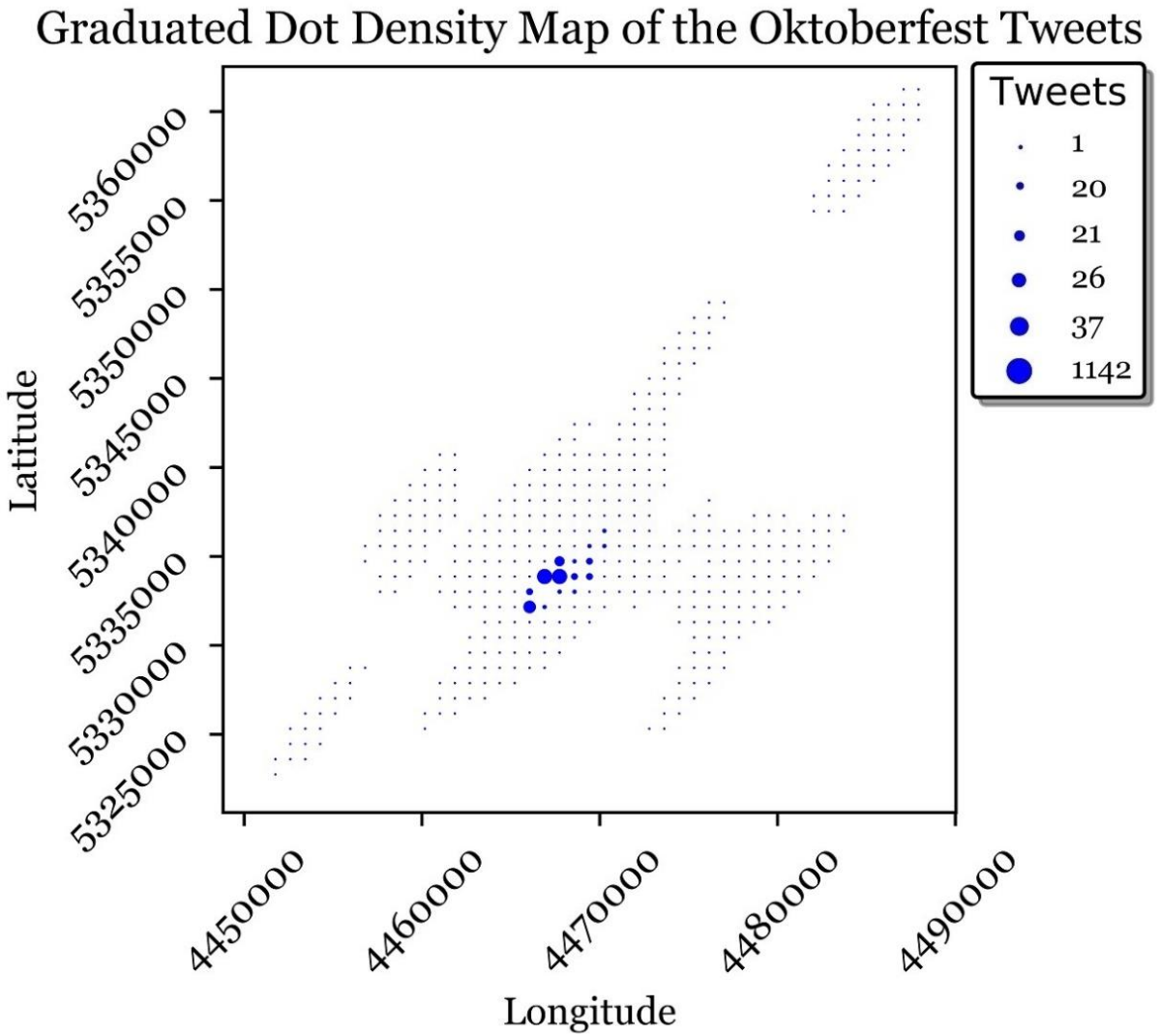
Graduated dot map results



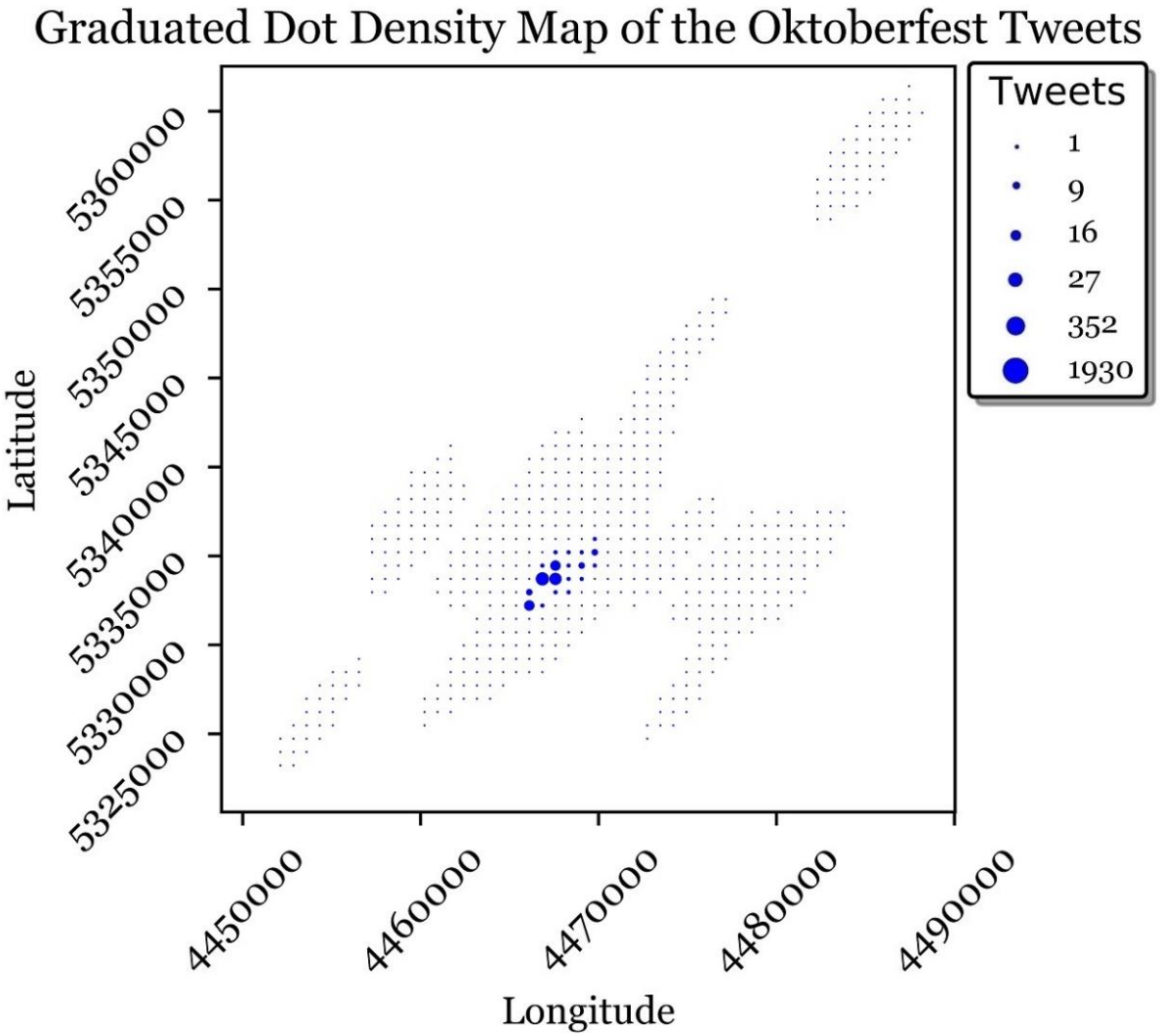
Graduated dot map results



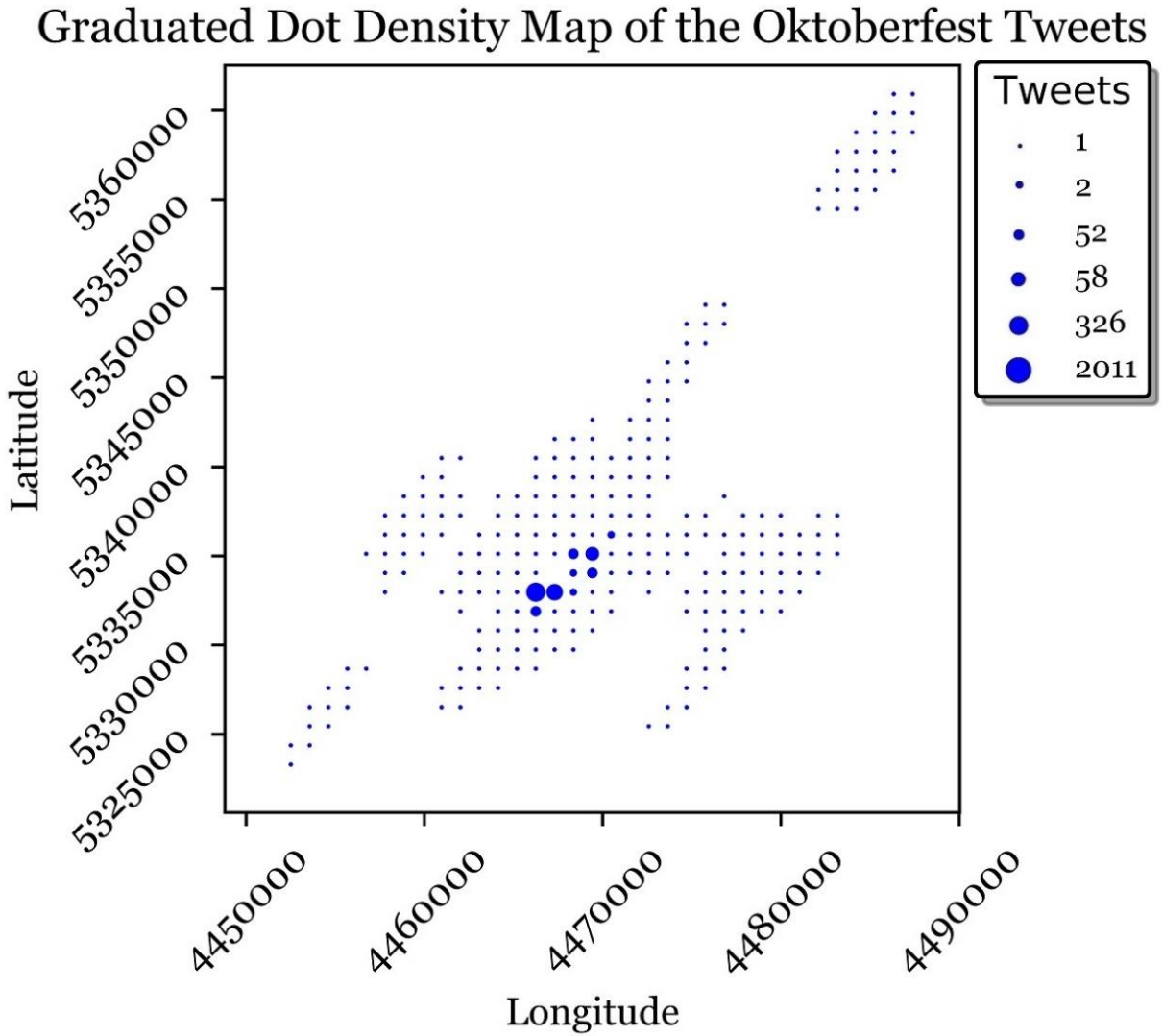
Graduated dot map results



Graduated dot map results



Graduated dot map results



Graduated dot map results

Summary of variables

Data	Resolution	Classes
POI	60 * 60	6
POI	50 * 50	6
POI	40 * 40	6
POI	30 * 30	6
Tweets	57 * 57	6
Tweets	50 * 50	6
Tweets	40 * 40	6

Summary of variables in the created Graduated dot maps

User test

1. Comparison of heatmaps and dot maps
2. Comparison of the conventional and graduated dot map
3. Estimation of the quantity in a specific region
4. Choosing the best type of map

User test

Comparison of heatmaps and dot maps
Comparison of the conventional and graduated dot map

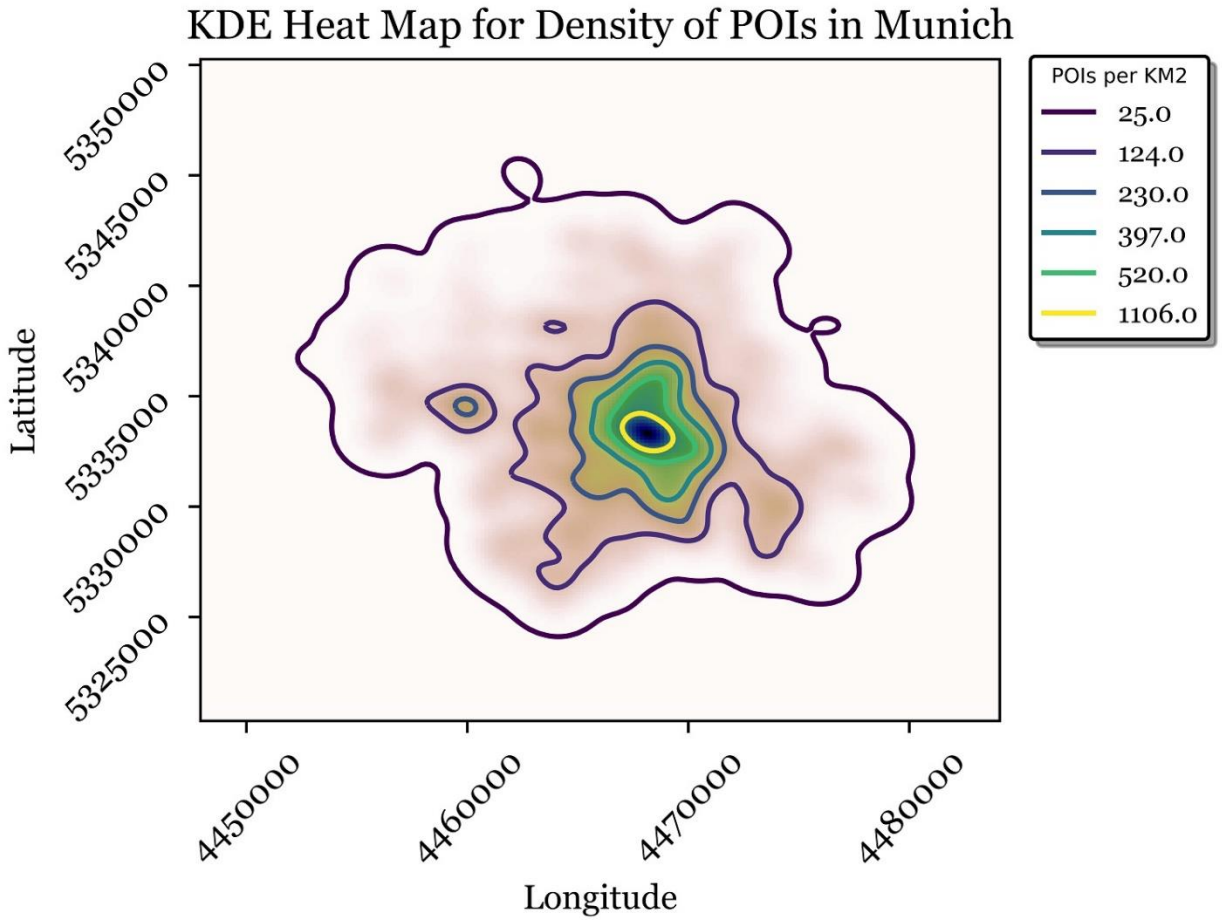
Questions:

1. Which map is better to perceive the density?

2. Which map provides a better quantitative estimation of the density in any part of the map?

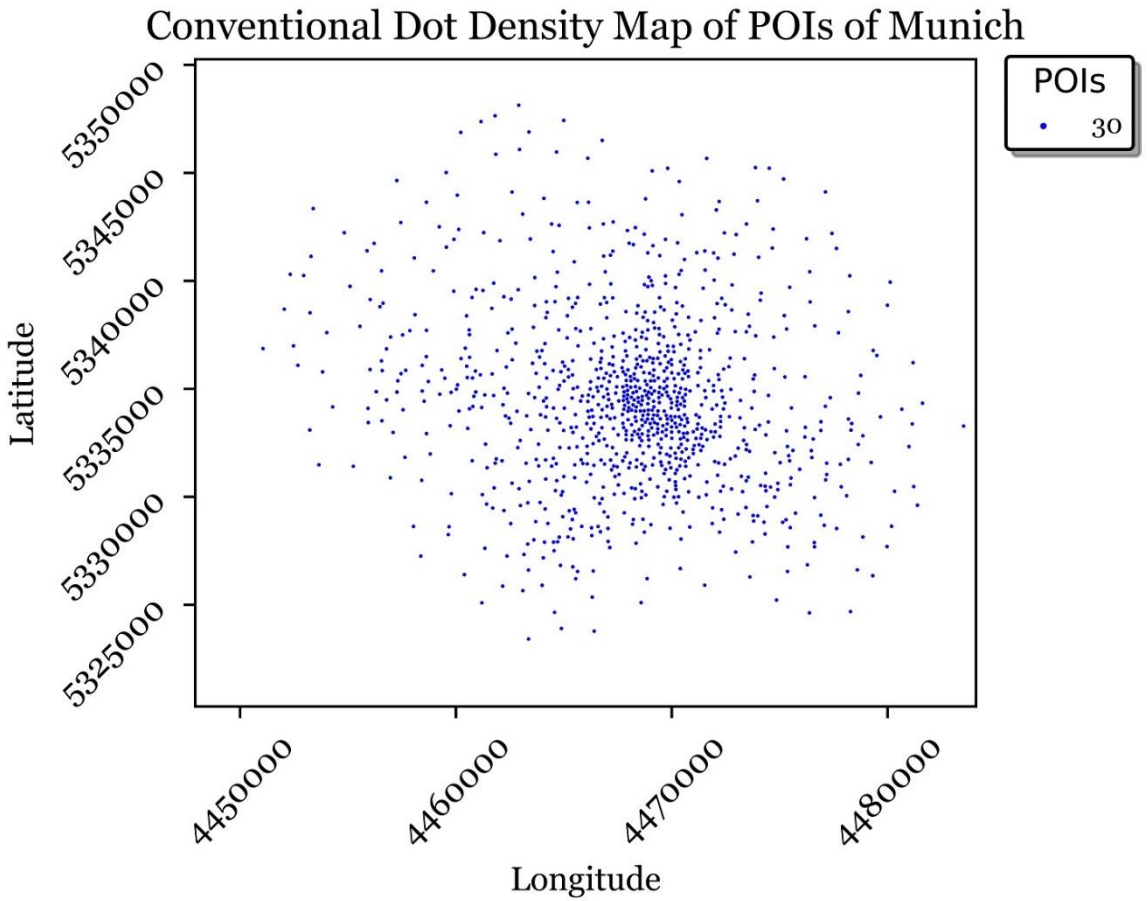
User test

Comparison of heatmaps and conventional dot maps



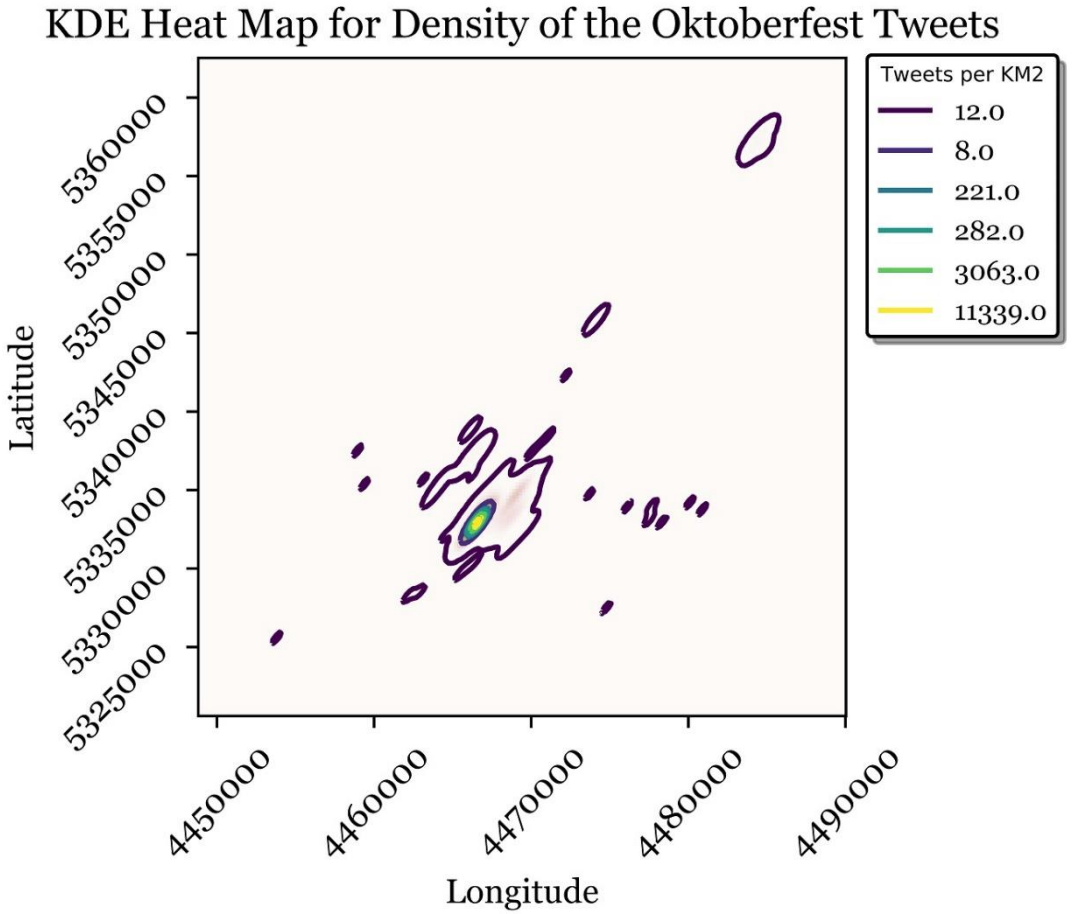
User test

Comparison of heatmaps and conventional dot maps



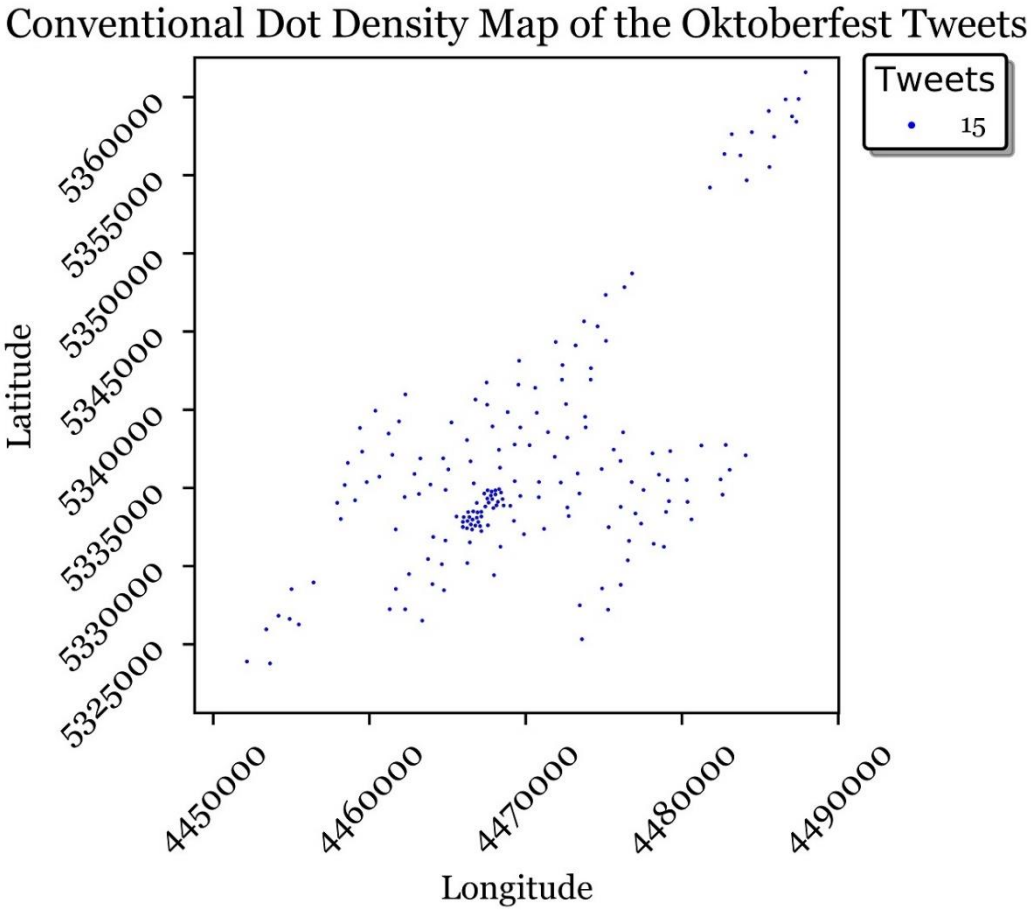
User test

Comparison of heatmaps and conventional dot maps



User test

Comparison of heatmaps and conventional dot maps



User test

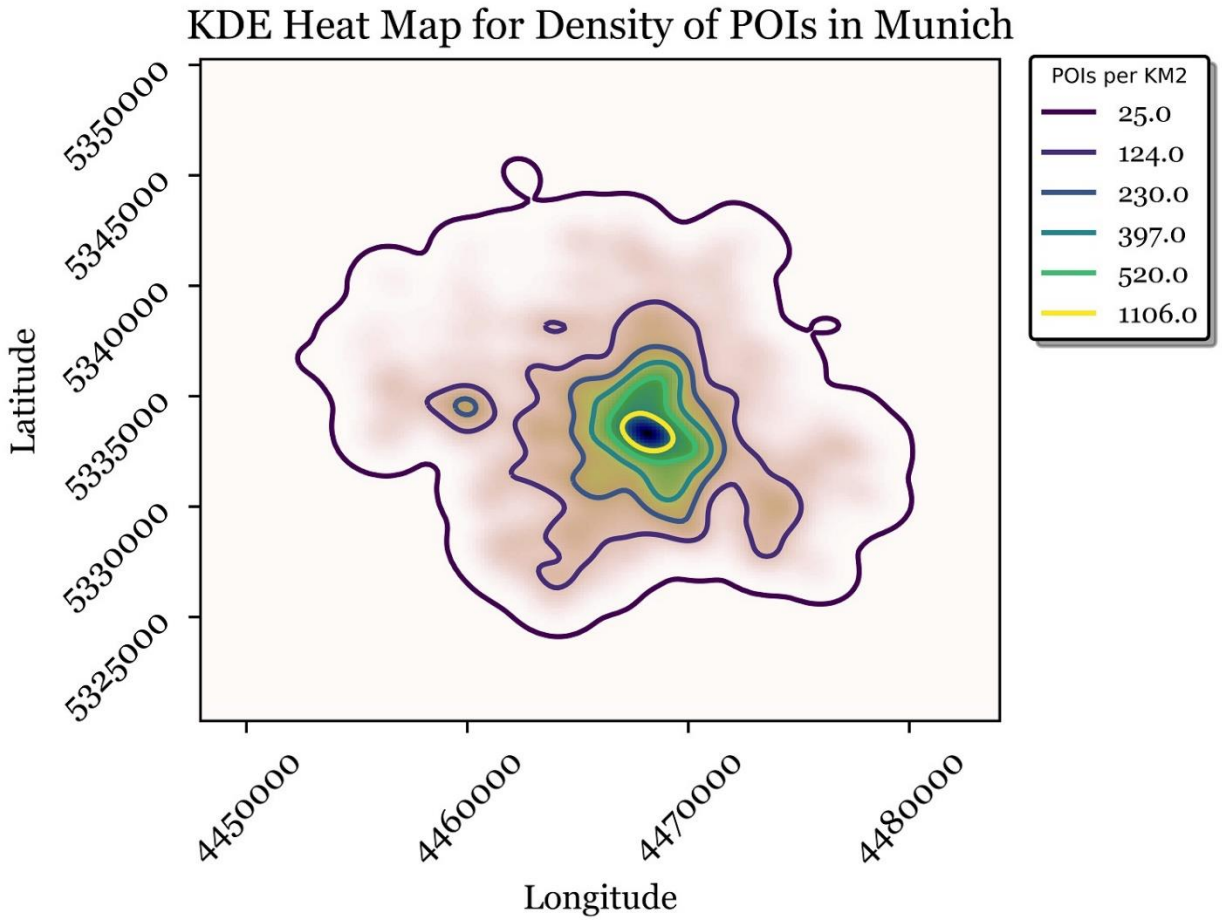
Comparison of heatmaps and conventional dot maps

Data	Objective	Dot map	Equal	Heatmap
POI	Density overview	10	1	3
	Quantification	7	4	3
Tweets	Density overview	10	2	2
	Quantification	10	2	2

Comparison of the preferences of 14 users

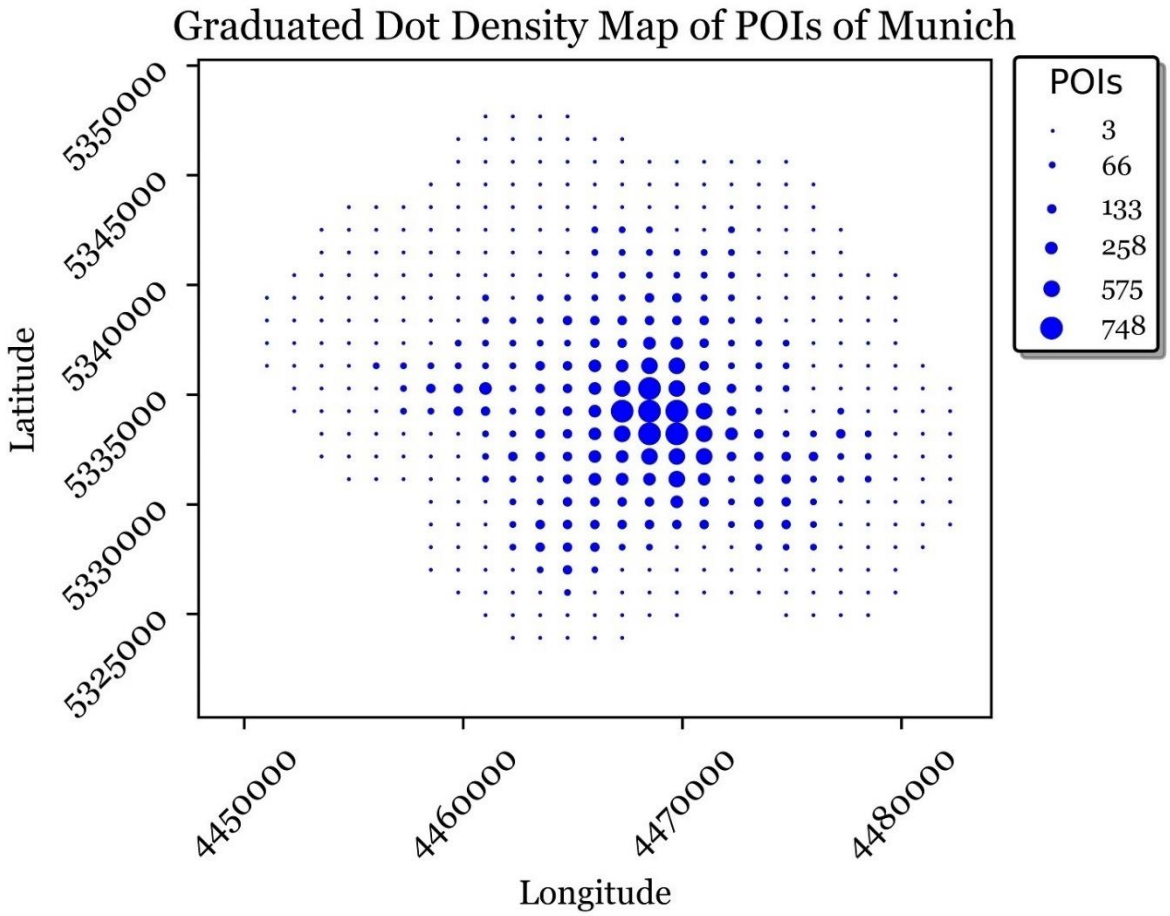
User test

Comparison of heatmaps and graduated dot maps



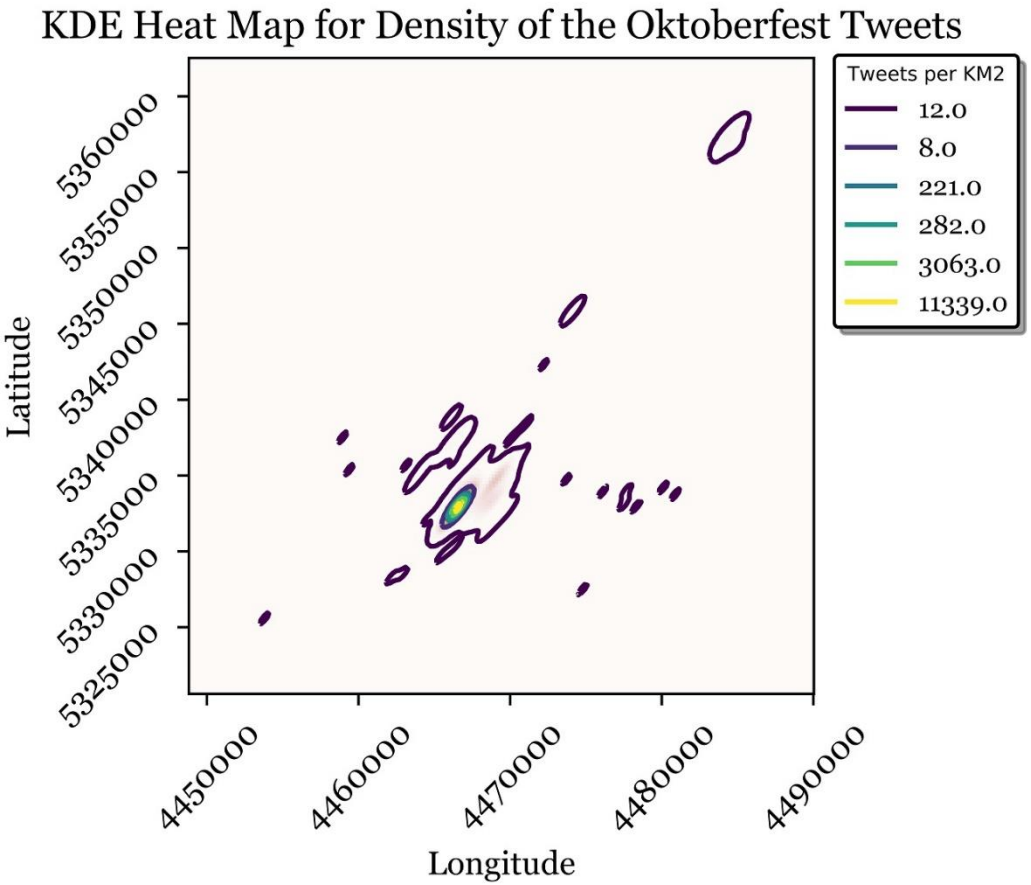
User test

Comparison of heatmaps and graduated dot maps



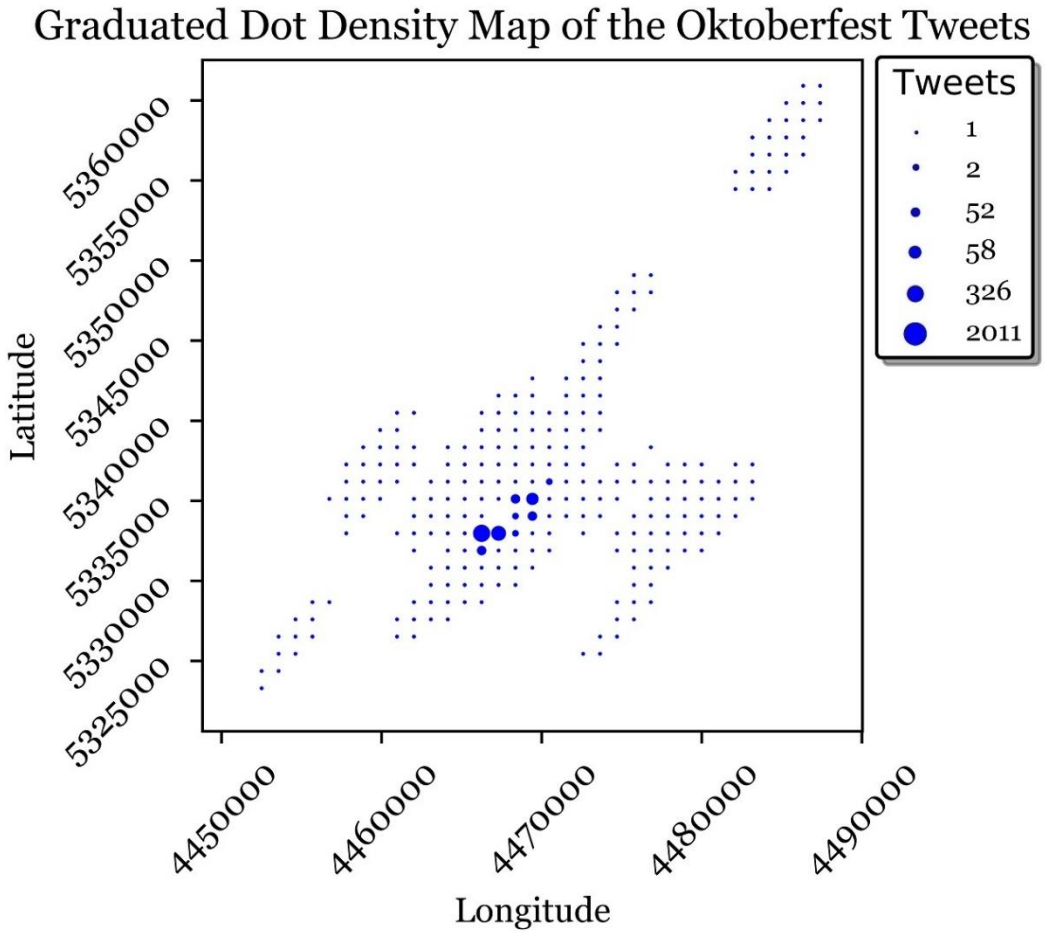
User test

Comparison of heatmaps and graduated dot maps



User test

Comparison of heatmaps and graduated dot maps



User test

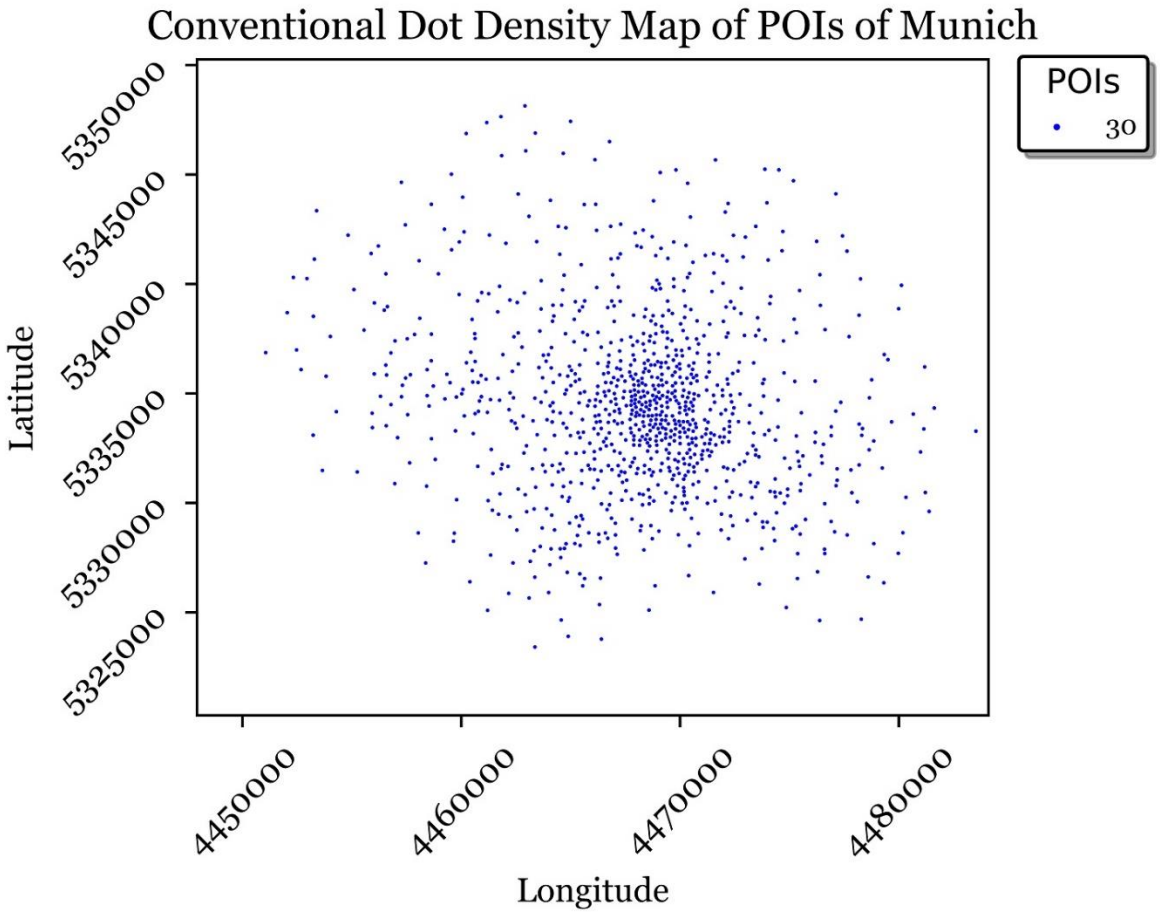
Comparison of heatmaps and graduated dot maps

Data	Objective	Dot map	Equal	Heatmap
POI	Density overview	11	1	2
	Quantification	10	1	3
Tweets	Density overview	10	3	1
	Quantification	9	2	3

Comparison of the preferences of 14 users

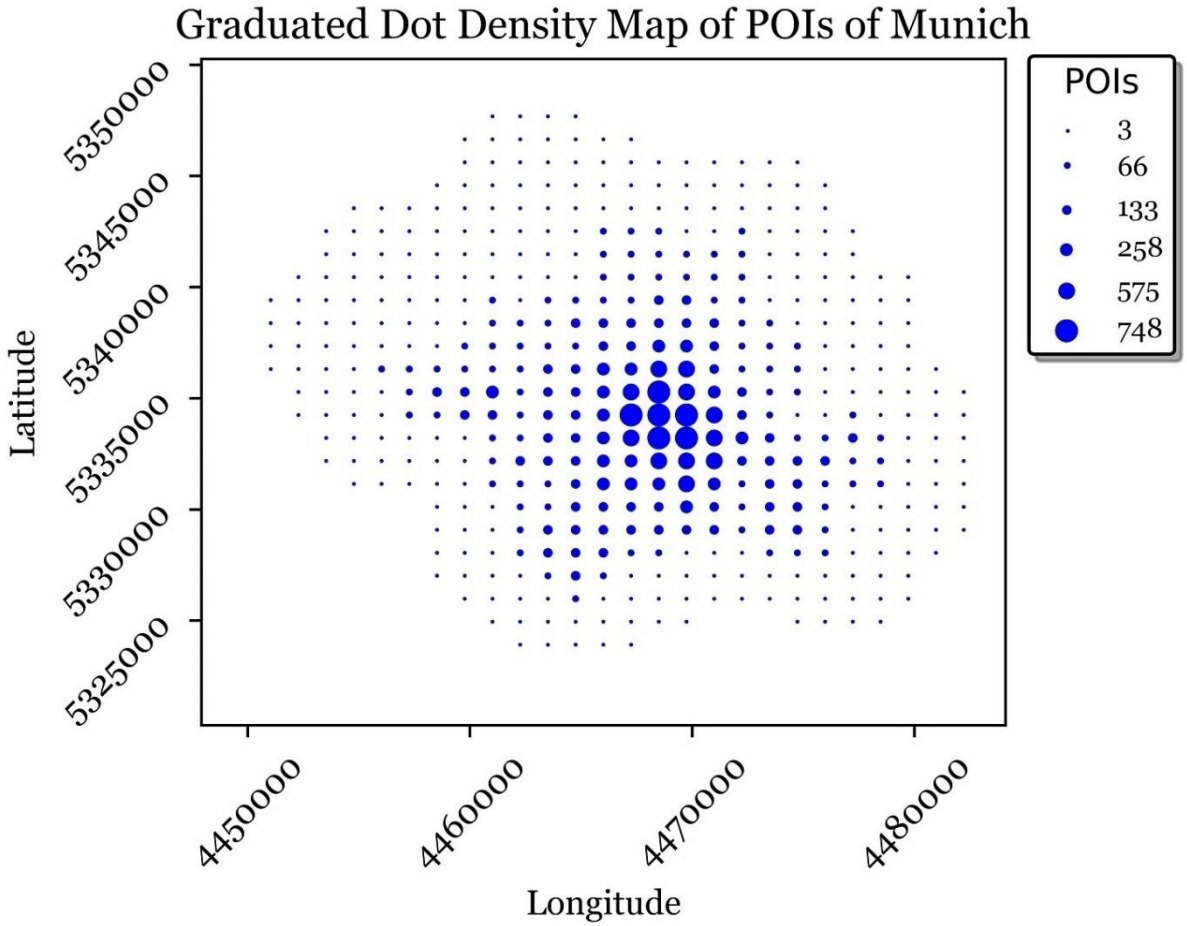
User test

Comparison of conventional and graduated dot maps



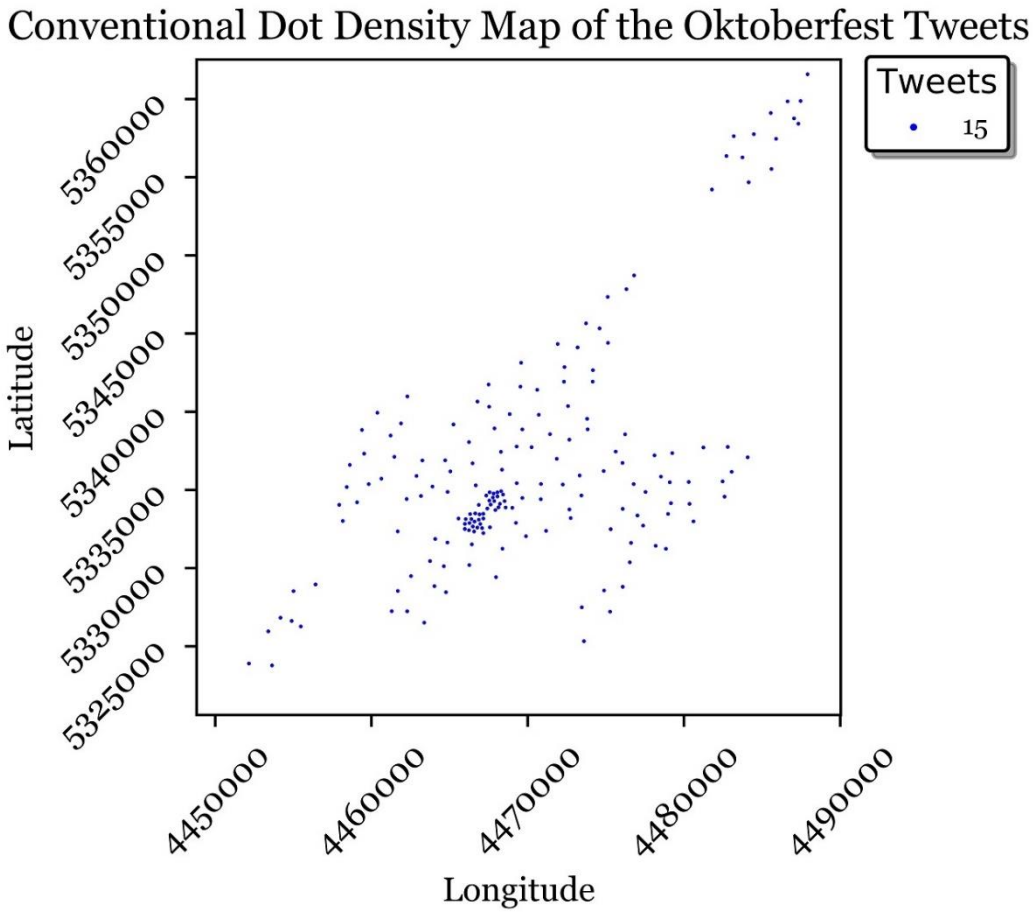
User test

Comparison of conventional and graduated dot maps



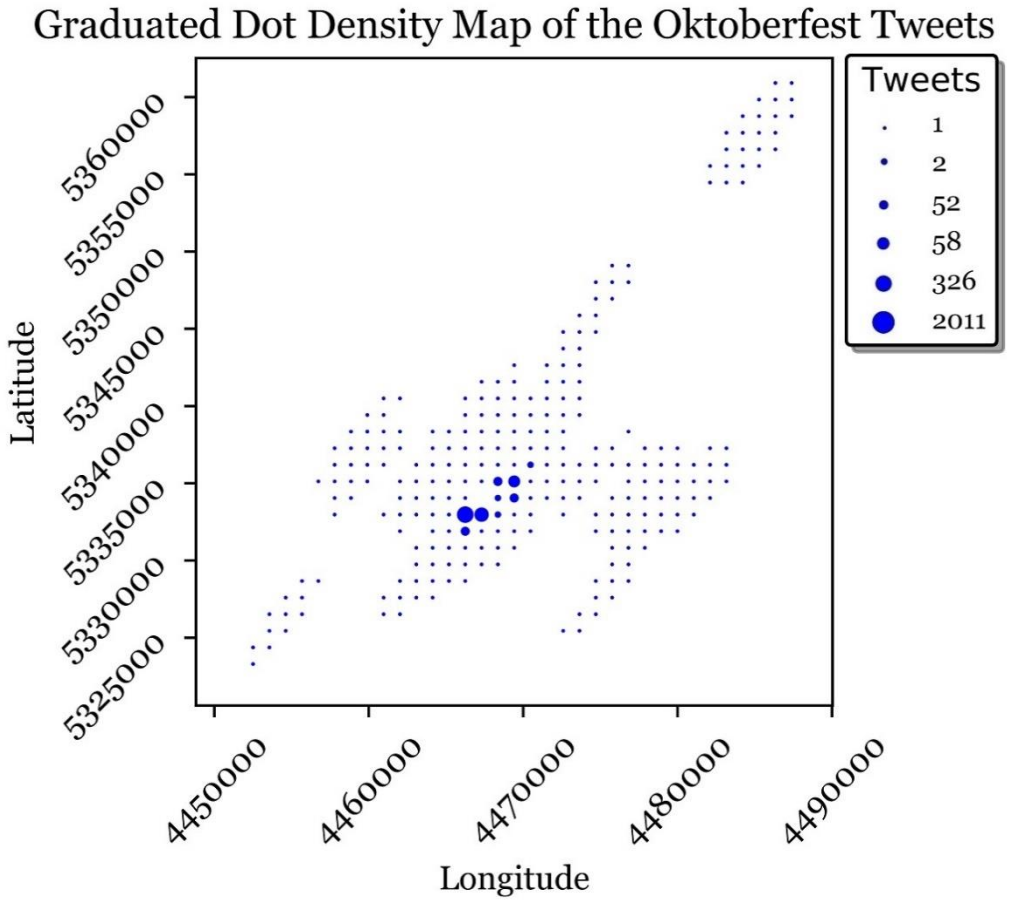
User test

Comparison of conventional and graduated dot maps



User test

Comparison of conventional and graduated dot maps



User test

Comparison of conventional and graduated dot maps

Data	Objective	Conventional Dot Map	Equal	Graduated Dot Map
POI	Density overview	8	1	5
	Quantification	6	3	5
Tweets	Density overview	7	3	4
	Quantification	7	5	2

Comparison of the preferences of 14 users

User test

Estimation of the quantity in a specific region

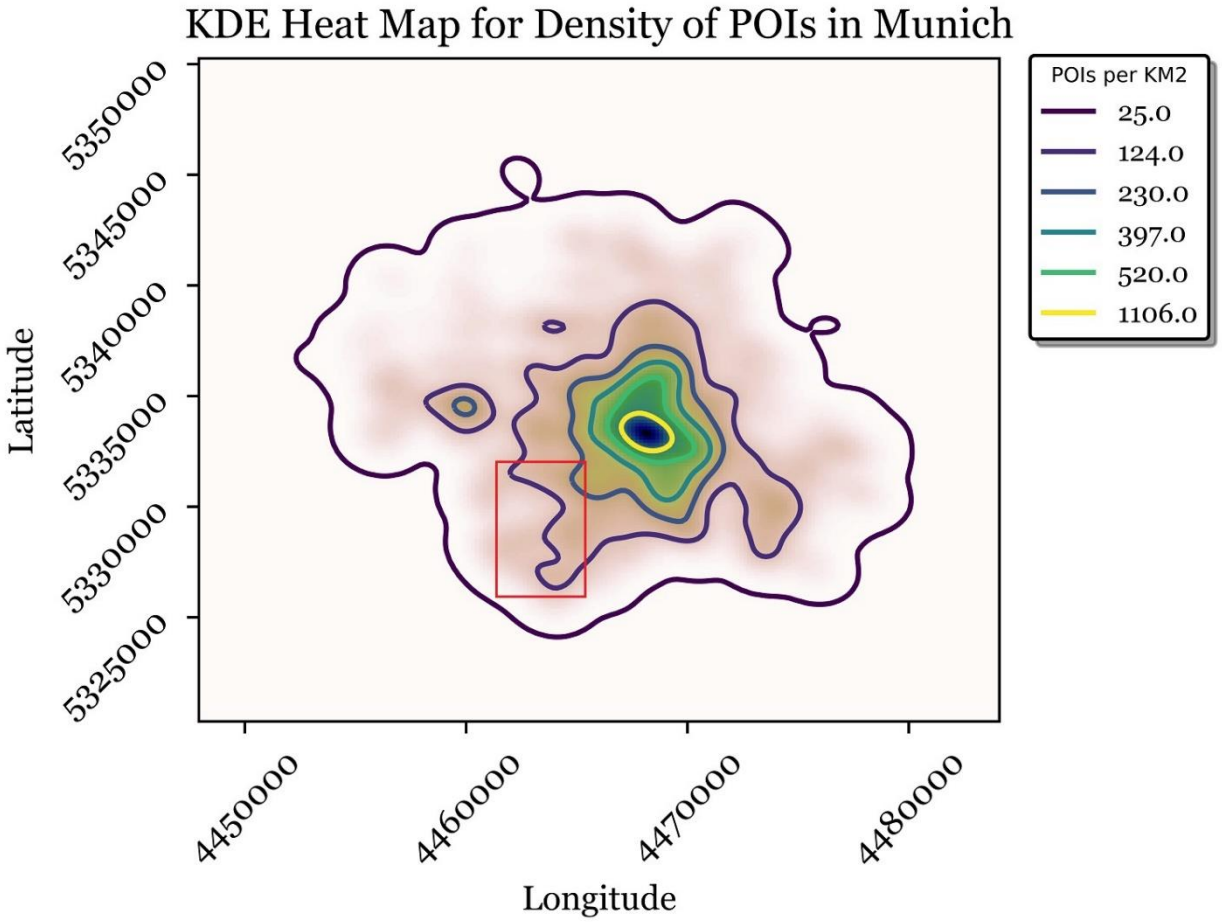
Questions:

1. How easy is the estimation of the density in the specified box?

2. What is the result of the estimation?

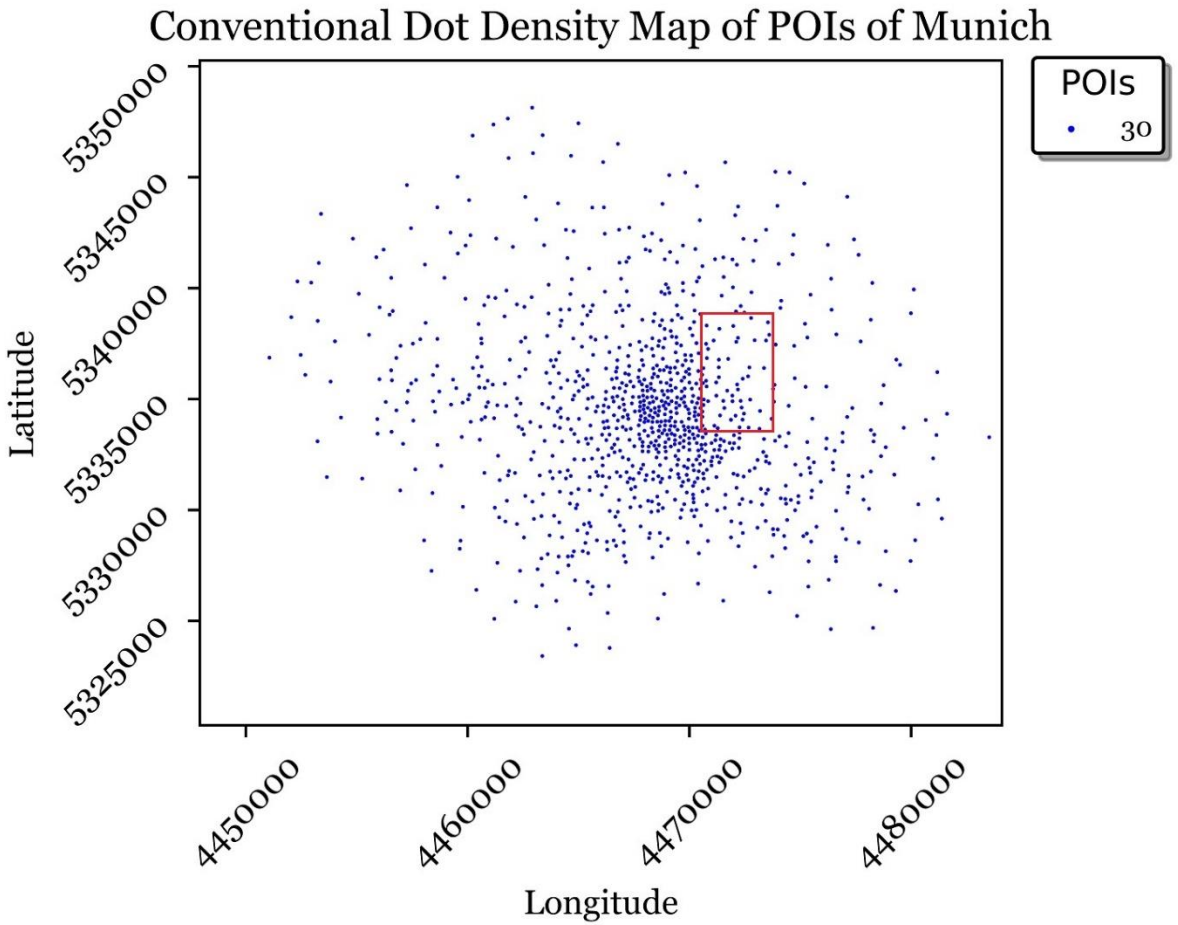
User test

Estimation of the quantity in a specific region



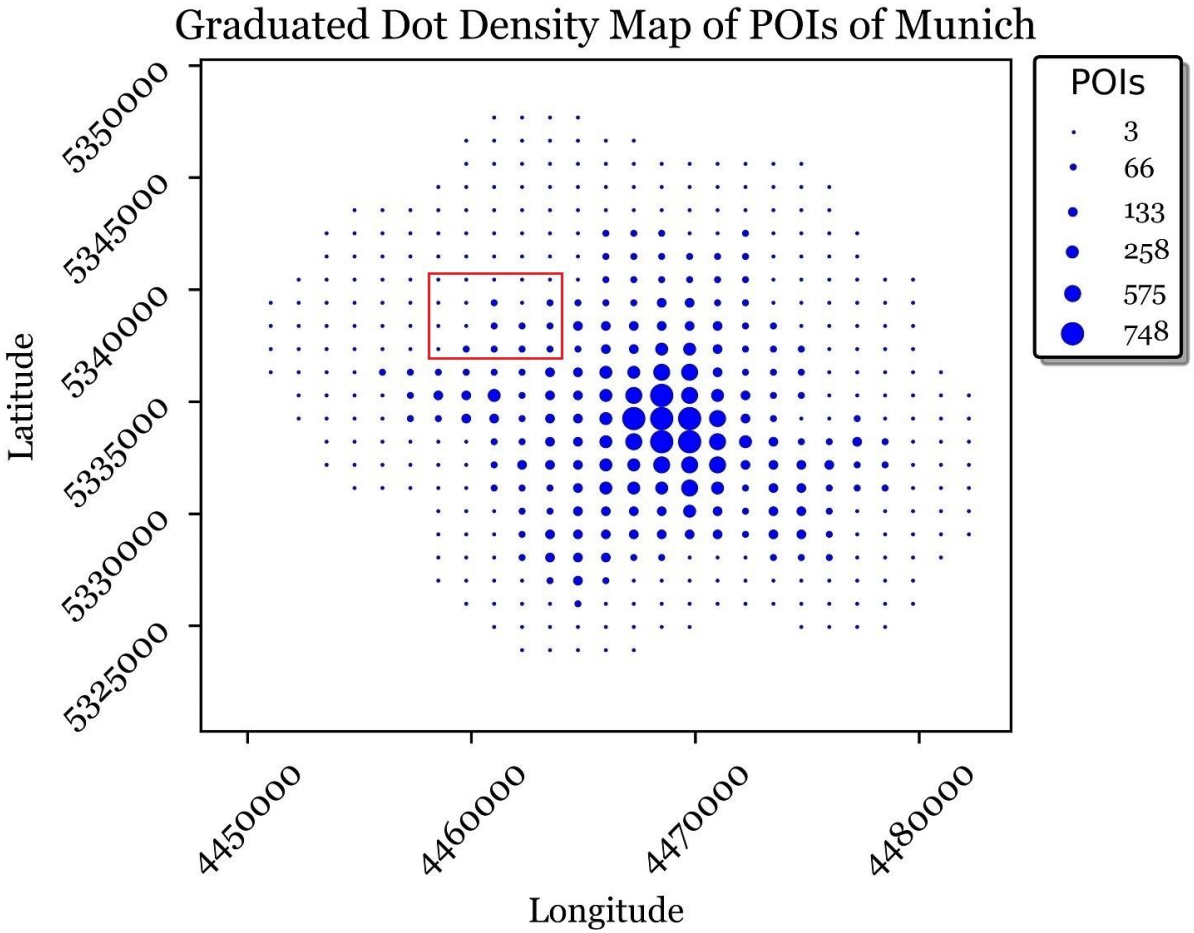
User test

Estimation of the quantity in a specific region



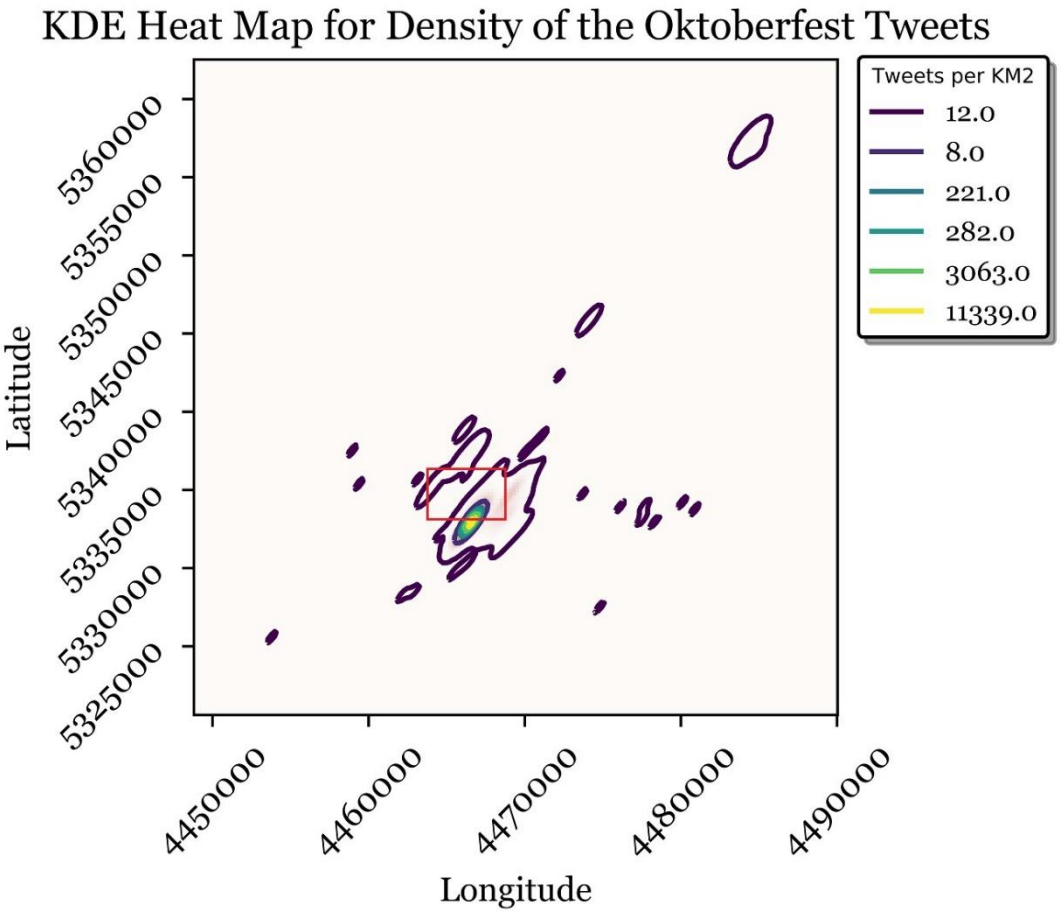
User test

Estimation of the quantity in a specific region



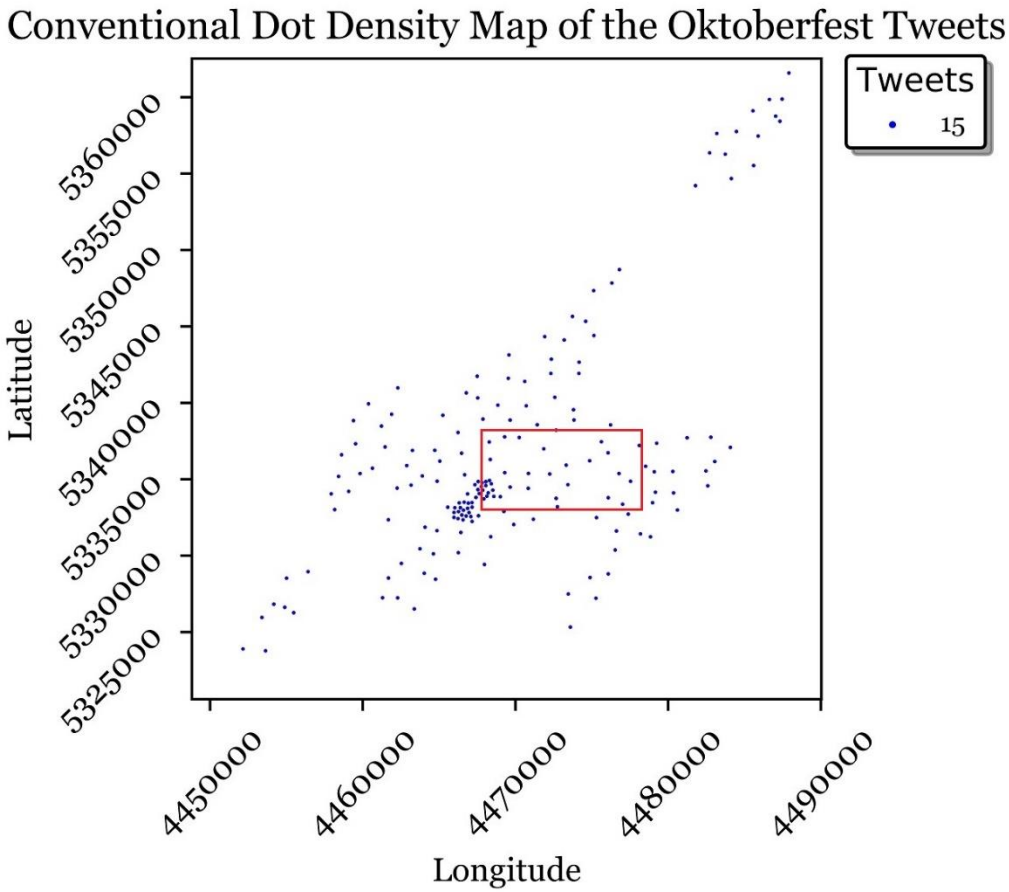
User test

Estimation of the quantity in a specific region



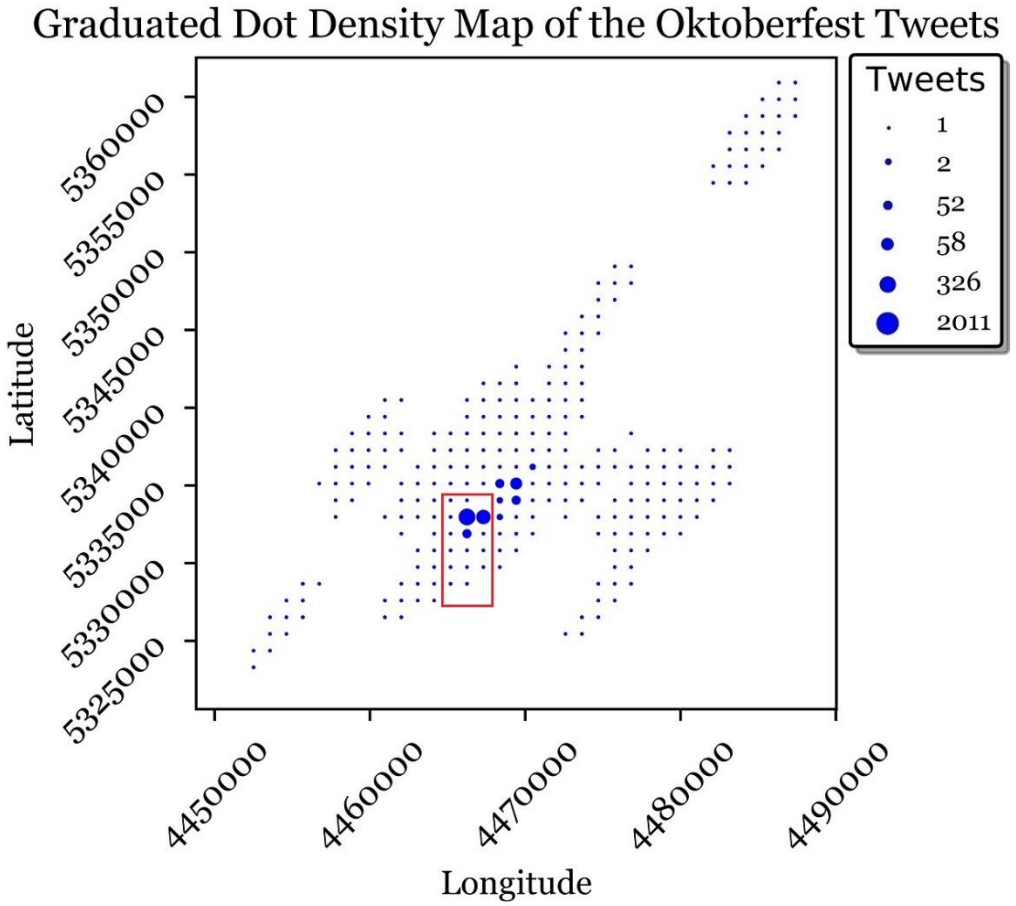
User test

Estimation of the quantity in a specific region



User test

Estimation of the quantity in a specific region



User test

Estimation of the quantity in a specific region

How easy is the estimation of the density in the specified box?

Map	Data	Easy	Medium	Hard
Conventional dot map	POI	5	5	7
	Tweets	4	4	6
Graduated dot map	POI	5	3	6
	Tweets	8	0	6
Heatmap	POI	0	0	14
	Tweets	0	0	14

Comparison of the preferences of 14 users

User test

Estimation of the quantity in a specific region

What is the result of the estimation?

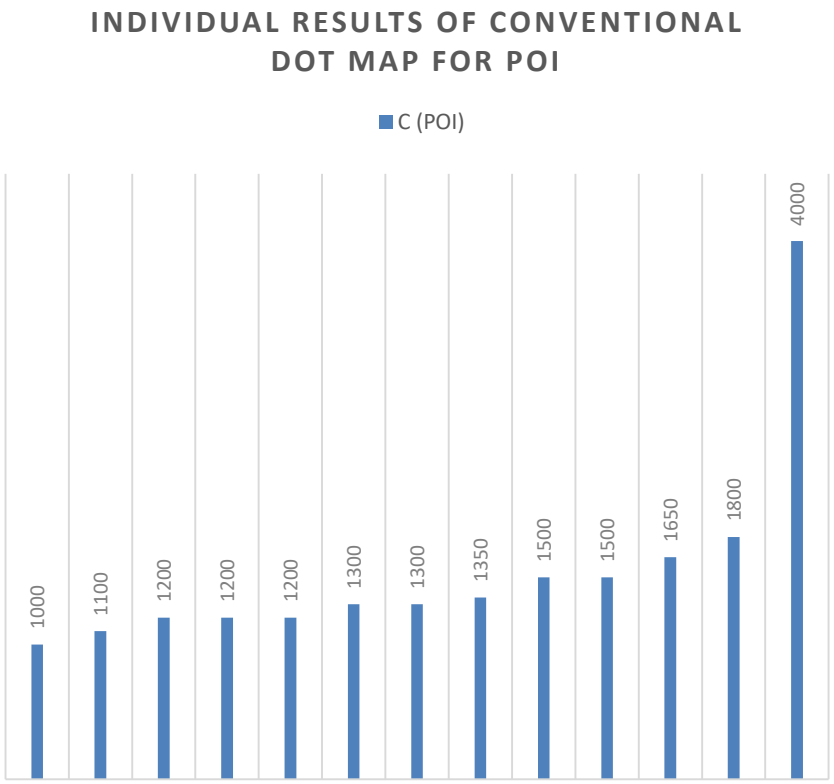
Maps	Data	Successful counters	True Number	Mean	Std. Dev.	Error Percentage	Min.	Max.
Conventional dot map	POI	13	1,740	1,546	765	11%	1,000	4,000
	TW	13	594	499	319	16%	400	630
Graduated dot map	POI	14	525	730	108	39%	539	1,400
	TW	14	2,402	1,544	1,429	36%	350	4,050
Heatmap	POI	---	---	---	---	---	---	---
	TW	---	---	---	---	---	---	---

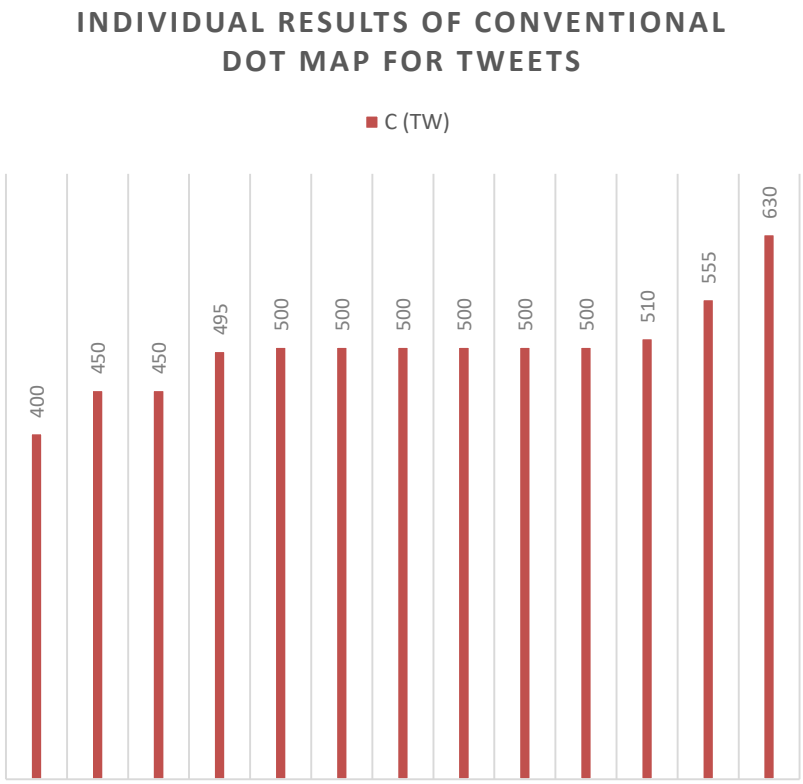
Comparison of the preferences of 14 users

User test

Estimation of the quantity in a specific region

Individual estimations

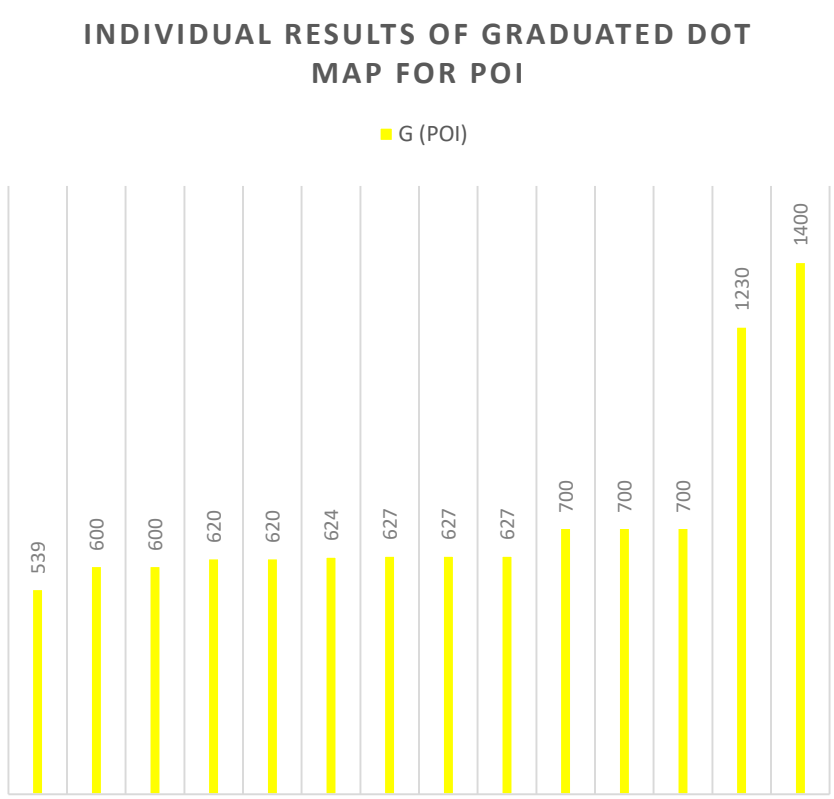


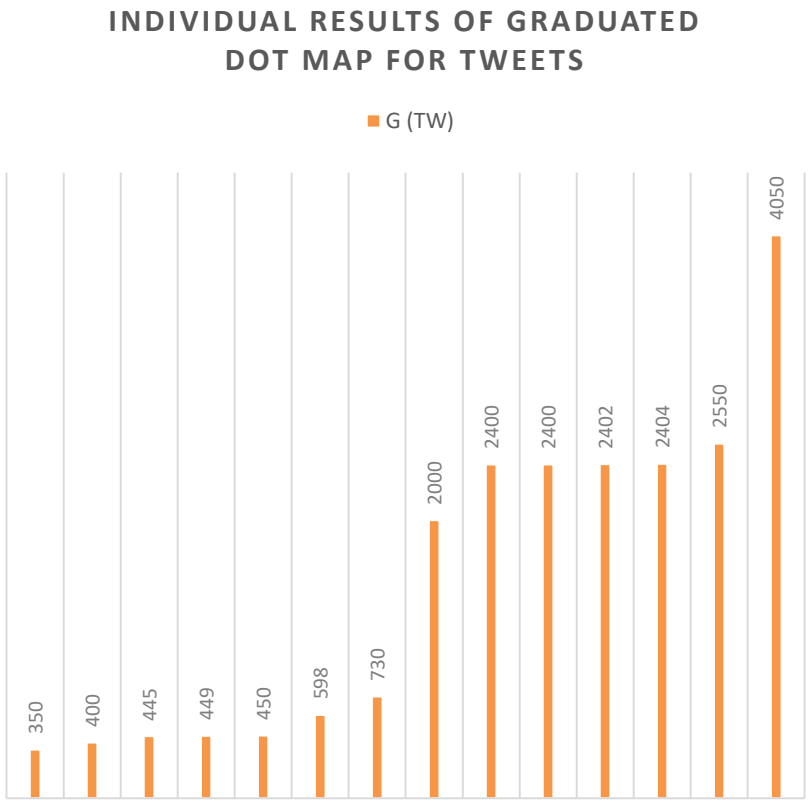


User test

Estimation of the quantity in a specific region

Individual estimations

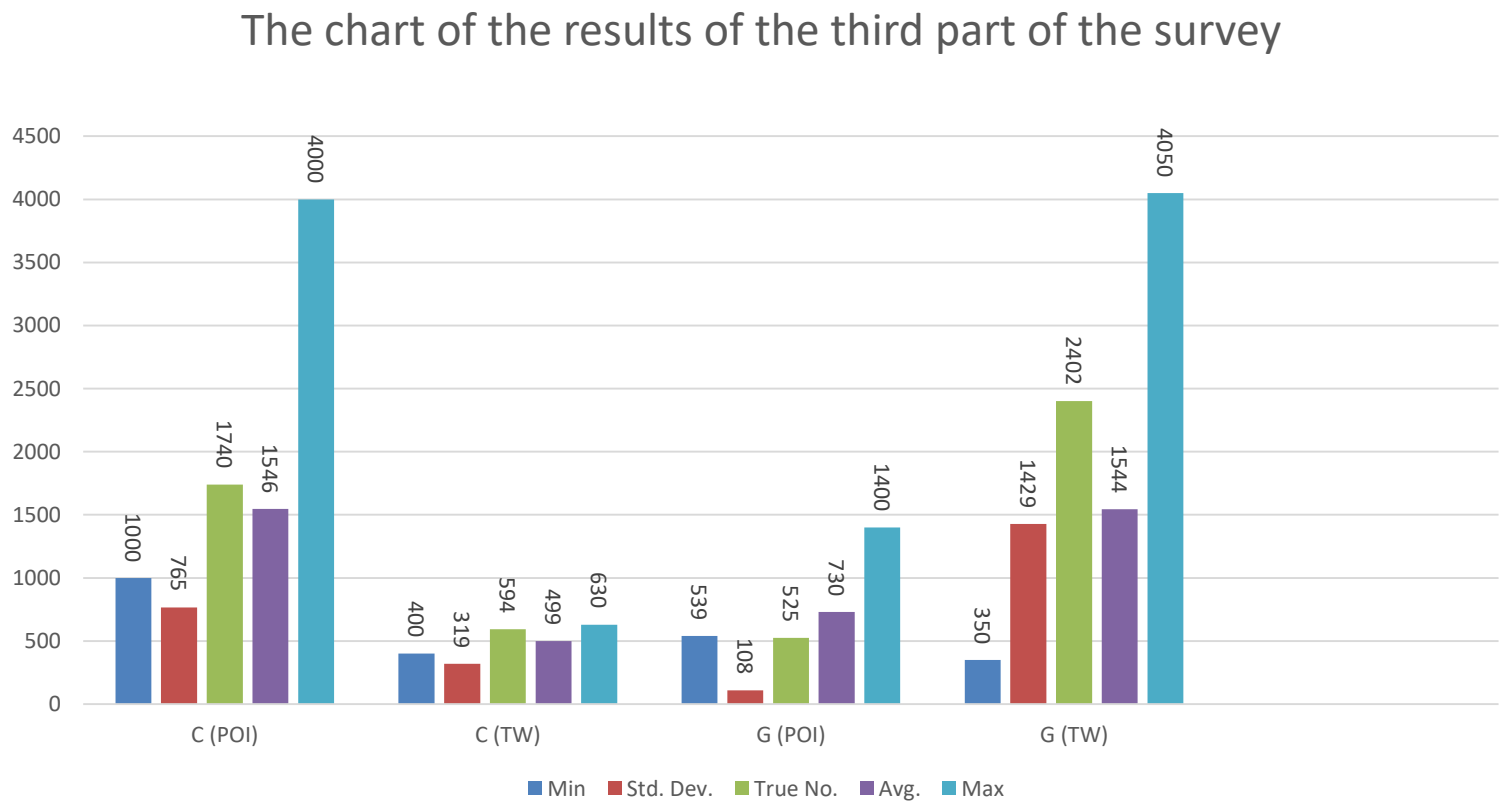




User test

Estimation of the quantity in a specific region

Main indexes



User test

Choosing the best type of map

Map type	Conventional dot map	Graduated dot map	Heatmap	None of them
Single selections	2	9	0	0
Multiple selections	3		0	

Answers to research questions

What are the most important criteria to make a dot map?

- Resolution of the raster
- Dot size
- Dot value
- Minimum of dot coverage
- Randomness of dot placement
- Avoiding dot overlap and coalescence

Answers to research questions

What are the difference between KDE heatmaps and dot maps?

- Dot maps do not use colors and cryptic density values.
- They present a quantitative estimation for any desired specific region on the map.
- They give a better and more understandable density overview of a phenomenon.

Answers to research questions

Do final dot maps provide quantitative information to the user?

- The participants were able to estimate numbers for specific areas using the dot maps.
- None of the users could get a number from the heatmaps.
- A minority of users who preferred a heatmap in the first part of the survey could not give any estimation based on the heatmap in the third part of the survey.
- At the end, no users chose heatmaps anymore.

Answers to research questions

What are the advantages and disadvantages of graduated and conventional dot maps?

Conventional Dot Map		Graduated Dot Map	
Pros	**Cons**	**Pros**	**Cons**
Fast calculation	Greater number of dots	Fewer dots	More calculation
More detailed estimation	Randomness of dots	Regularity of dots	Less detailed estimation
Only one dot value	More parameters in the algorithm	Fewer parameters in the algorithm	Various dot values
No need to classification	More iterative processes for settings	Fewer iterative processes for settings	Requirement to a classification method

Conclusions

- The proposed dot maps are clearly superior to heatmaps
- Dot maps help users to quantify the number of points of a phenomenon in any favorable region on the map.
- Dot maps do not use any cryptic approach to show the density.
- There is no definite advantage between graduated and conventional dot maps.
- Choosing between the two dot map types highly depends on data and users' objectives.

Thank you for your attention

Questions?


ITC
UNIVERSITY OF TWENTE.
TECHNISCHE
UNIVERSITÄT
DRESDEN
Technical
University
of Munich
TU
WIEN
TECHNISCHE
UNIVERSITÄT
WIEN
Vienna University of Technology