A network-based approach to predict commercial activities optimal location and infer urban centrality indices

Master thesis

Spring semester 2009
June 2009

Jan Stauffer, SIE-PME, EPFL

Supervised by Dr. Stéphane Joost and Prof. François Golay, LASIG, EPFL

Abstract

The goal of this master thesis is to analyze the structure of the commercial activities in the cities of Geneva and Barcelona. This study implemented a network-based approach to predict the location of commercial activities, on the basis of which a tentative inference of urban centrality indices was carried out. This was done by using a 'grouping algorithm' and by calculating a 'location quality index', both presented and described in the '*Network-based predictions of retail store commercial categories and optimal locations'* paper published by Pablo Jensen (2006).

In the first part of the work, an existing prototype of the algorithm – implemented during a semester project as *VBscript* in the GIS software 'Manifold' – was optimized. Time gains of a factor 10 for the 'link computation' and more stable results for the 'grouping algorithm' could be achieved.

A spatial representation of the grouping-results was elaborated and permitted to recognize the dominance areas of the activity-groups for both cities. The maps created with the 'location quality index' (Q-Index) allowed to identify preference areas and location patterns for the different categories of commercial activities. It was possible to notice that the initial definition or choice of the categories is very important to get clear results. Indeed, the results on Geneva with only 18 categories performed less well than the ones with 45 and 48 categories on Geneva and Barcelona respectively.

Furthermore, tests were carried out on the basis of hectometric referenced data, like the database of the Swiss federal census of enterprises. The results obtained are very similar to the ones from the postal address-based data (precision 2-3m). This shows that the algorithm used here could be applied to whole Switzerland in a possible future study to determine the overall structure of commercial activities in the country, to identify possible central locations, or also to propose it as useful tool for cantonal or regional economic development agencies to determine optimal locations for new coming large foreign industrial commercial groups.

The results of the comparisons between the Q-Index and centrality indices showed that an approximation of the latter by the first one is not possible. But the Q-Index (describing location preference per category of commercial activity) can be used as a complementary indication to centrality (which describes accessibility in a broad sense).

As a perspective, it would be particularly interesting to work on a definition of a Q-Index for *groups* of categories or even of a global integrated one: it would then be probably possible to use this index as a surrogate for the evaluation of value landed property in particular.

Résumé

Ce projet de master a pour but d'analyser la structure des activités commerciales des villes de Genève et de Barcelone, et de tenter de prédire l'emplacement optimal des commerces avec l'aide d'une méthode de regroupement par réseau, et enfin d'essayer d'en déduire des indices de centralité urbaine. Cela a été fait sur la base d'un algorithme spécifique et d'un indice de qualité d'emplacement (indice Q) décrits dans l'article «*Network-based predictions of retail store commercial categories and optimal locations*» écrit par Pablo Jensen (2006).

Dans la première partie du projet, un prototype de cet algorithme – préalablement implémenté au cours d'un projet de semestre en *VBscript* dans le logiciel SIG Manifold – a été optimisé. Un gain de temps d'un facteur 10 pour le calcul des liens et des résultats plus stables pour l' «algorithme de regroupement» ont pu être obtenus.

Une représentation spatiale des résultats du regroupement a été produite. Elle permet de reconnaître les aires de dominance des groupes d'activités commerciales localisées à l'adresse dans les deux villes (précision 2-3 m). Les cartes créées sur la base de l'indice Q ont permis d'identifier les aires préférentielles et des caractéristiques de la localisation pour les différentes catégories d'activités commerciales. Il a également été remarqué que la définition initiale des catégories commerciales (qualité et nombre de catégories) est déterminante dans le but d'obtenir des résultats clairs. En effet, les résultats pour Genève avec seulement 18 catégories se sont avérés moins performants que ceux obtenus avec 45 et 48 catégories pour Genève et Barcelone respectivement.

En outre, des essais de calculs ont été réalisés sur des données géoréférencées à l'hectare, comme la base du recensement fédéral des entreprises. Les résultats obtenus dans ce cas sont très similaires aux premiers (précision à l'adresse). Ceci montre que l'algorithme développé dans cette recherche peut être appliqué à toute la Suisse à l'avenir, dans la perspective d'obtenir une image de la structure nationale des activités commerciales, dans celle d'identifier de potentiels lieux centraux dans le pays, ou alors de fournir aux offices cantonaux ou régionaux de promotion économique des outils en mesure de conseiller de grosses entreprises étrangères en ce qui concerne le choix d'une localisation géographique optimale.

Les résultats des comparaisons entre l'indice Q et divers indices de centralité indiquent qu'une approximation du deuxième par le premier n'est pas possible. Mais l'indice Q (décrivant les aires de préférence par catégorie d'activité) peut être utilisé comme complément à l'indice de centralité (qui décrit l'accessibilité au sens large). Il serait intéressant de travailler à l'avenir sur une définition d'un indice Q pour les *groupes* de catégories ou voire même un indice intégré global. Ceux-ci pourraient probablement être utilisés comme substitut pour l'évaluation de la valeur de la propriété foncière notamment.

Table of contents

List of figures

List of tables

1 Introduction

Everyone can observe diverse location patterns of commercial activities in town. Jeweller's shops for example are more likely to be found in the center of the city, mostly concentrated in the same area. Bakeries and kiosks on the other hand are rather spread over the territory. These particular patterns depending on the activity illustrate how location is essential for most commercial activities. But not only does the absolute location matter (characterized by accessibility for example), so does the relative location of one store to the other. In fact, the relative location patterns are the point of departure for Jensen's article '*Network-based predictions of retail store commercial categories and optimal locations'*, which appeared in 2006 in the *Physical review E 74.* He explains in the article how "[…] *location data suffices to reveal many important facts about the commercial organization of retail trade."* He defines the interactions of the different activities with attraction/repulsion links between different categories of activities and then uses the so called *'Potts algorithm'* to group the activities in commercial categories. The cited algorithm is adapted from physics and interprets the activities as magnetic spins attracting or repulsing each other. He also defines a location quality index *Q* which depends on the surrounding commercial activities of the activity in question and their interaction coefficient with that same activity.

I have already developed a prototype of the algorithm described in Jensen's article as *VBscript* in the GIS software 'Manifold' during a 4 credits semester project.
The aim of this master thesis is now to improve and adapt this code, to apply it on Geneva and Barcelona, to analyze the results and finally try to approximate centrality indices.
The last part – the approximation of centrality indices – arose because Timothée Produit is doing his master thesis on activity-density and -diversity on Barcelona in collaboration with Prof. Sergio Porta from the Human Space Lab (Milan, Italy) to compare them with centrality indices calculated by Porta (Produit, 2009).

In the following pages, the detailed objectives are stated first, and then the methodology is explained. Further the developments and the case studies are presented and discussed before finally coming to the conclusion and perspectives.

For the sake of simplicity, the word 'activity' is used like a synonym for 'commercial activity' and the word 'category' like a synonym for 'category of commercial activity' in the whole report.

2 Objectives

There are two overall goals for this master thesis:

- Form groups of commercial activities with similar location characteristics to analyze the commercial structure of the city.
- Use network-based predictions on commercial categories for location planning and to infer urban centrality indices.

Thus the main outputs of this research will be a spatial representation of the classified groups with similar location characteristics on the one hand and maps with indications of the location quality for the different categories of commercial activities on the other hand. These latter will be used to try to approximate centrality indices.

Since there already exists a prototype of the program, which does most of the calculation part, the operational objectives are the following:

- Optimize the time consuming calculation of the 'M-Index'.
- Introduce the possibility to choose a random subset of the activities (to take into account geographical peculiarities) and test it.
- Get a deeper understanding of the process of simulated annealing and use it to optimize the existing program.
- Test and compare the classification capabilities of the algorithm.
- Study and represent location characteristics for different groups of categories.
- Calculate, analyze and validate the location quality index per category.
- Compare the results of calculations with approximate hectometer-location-data and the ones obtained from exact address-location-data for Geneva.
- Try to find a centrality index and compare it with the results found by Sergio Porta on Barcelona.

3 Methodology

The first part consisted of a literature review on the one hand, and some programming to optimize the existing code on the other hand.

The choice of the articles to read was difficult, as there was and is a lot of research done on networks and their proprieties and structure. Furthermore, my knowledge on networks and simulated annealing was little, so I had to begin with papers explaining also the basics as for example Freeman (1979) for the centrality in networks.

On the programming side, the first step was to optimize the calculation of the M-Index, which was the limiting factor for the processing of large databases in the prototype of the program.
Secondly, the possibility to use a random subset of all activities was added.
And the third step was to implement and optimize the simulated annealing (SA) part, which does the classification of the categories of activities in groups (or 'communities'). Here, the optimization should not be understood the same way as for the M-Index, since it consists in the optimization of the parameter values rather than in changing the algorithm. In fact, Jensen generously provided us with his code for the SA (in the programming language *Fortran*), which was adapted in *VBscript* for the prototype. But since there were different unclarities left concerning the method and the interpretation as well as the choice of some parameters, I decided to re-implement this part with the method described in Johnson (1989). On the one hand I did so because this algorithm has an 'intelligent stopping criterion', on the other hand because there are descriptions and recommendations for each parameter in the article.
Besides these optimizations, divers small 'helping scripts' were coded, which allow the automated running of the code with different parameters and the storage of the results for each variant.
For the visualization of the results for the location quality index (Q-Index), a form has been created, which facilitates the changing and normalizing of the different Q-Indices.
Furthermore, the code for the Q-Index was adapted to calculate its value for the activities already existing. This was used afterwards for the validation of this index.

After this 'development part', a lot of different calculations and analyses were done on the activities of Geneva and Barcelona. For Geneva, two 'sets of activities' were used: First only the retail stores (as Jenson did), then all recorded activities. For Barcelona, one 'set' with all activities was used.
For the retail stores of Geneva, sensibility analyses were done for different radiuses and different subsets of all activities. The comparison was done by comparing descriptive statistics of the matrices with the link-weights and by calculating their correlation coefficient in pairs with the help of MATLAB. Furthermore, the location-data was degraded to a hectometric grid and then the results were compared to the ones calculated with the original address-precision-data. This comparison is interesting because hectometer-data exists for all of Switzerland (from the federal census of enterprises) that could be used if the results are similar.
The classification of the categories of activities was done for different (sub)sets for all three 'sets' using the radius with the most non-zero links found in the analysis mentioned before. Since a comparison of these results is difficult just by analyzing the lists of the different groups, a spatial representation of the groups was elaborated to do so.
The Q-Indices were calculated for a 50m grid on the whole territory of Geneva, for a 200m grid covering whole Barcelona and for a 10m grid covering only a portion of the center of Barcelona. For

Geneva, a validation of this Index was done by comparing the Q-Index values of newly emerged activities with those of activities, which have vanished recently. For Barcelona, a comparison with the centrality-values of Porta was done.

Finally, this report was written to present and discuss the results of this master thesis.

4 Literature review

In this chapter, I give a short overview over the reviewed literature.

As mentioned above, my work is based mainly on Jensen's article (Jensen, 2006) which describes his approach to classify retail stores and calculate a location quality index and its application on the city of Lyon. To define the link-weight in the network of store categories, Jensen uses the M-Index recently introduced by Marcon and Puech (2008) to measure the geographic concentration of economic activities. For the classification, a 'Potts Model' and simulated annealing (SA) is used, as suggested by Reichard and Bornholdt (2004). Both the 'Potts Model' and SA come from physics. In fact, each category of retail stores is interpreted as a magnetic spin, and groups are found by minimizing the system's energy.
Such physical approaches in geographical science have already been done in other contexts. An example is the method of Gastner and Newman (2004), used to create equal-density maps with a molecule diffusing process.

Since one of my goals is to approach a centrality index, there is a short review on this theme in the next section. It is followed by an overview of the classification methods of nodes in networks, which is the same as finding communities in networks. Third, the sources for the simulated annealing part are presented.

4.1 Centrality

The concept of centrality in networks and its development are described in detail in Freeman (1979). He gives an overview of the work done before him and defines three centralities: the degree-centrality, the closeness-centrality and the betweenness-centrality, for each node but also for the whole network (graph). The developments and examples are based on an intuitive understanding of centrality: the central point of a star that has the highest value, while the outer points have minimal centrality.
There are multiple variants of the three basic centrality-measures and also several algorithms calculating them. Research on the topic is still active: A new variant combining degree and closeness called delta- or information-centrality was recently introduced by Latora and Marchiori (2007). And as the centrality calculation is time consuming for large networks, one searches still for new algorithms and variants, e.g. by Brandes (2008) and Rattigan et al. (2006).
Crucitti, Latora and Porta (2006) study different centrality measures of urban streets – a spatial network. These networks are different from relational networks, as their topology "is strongly constrained by their geographical embedding" (Crucitti et al., 2006). They calculated the closeness-, betweenness-, straightness- and information-centrality for samples of 18 world cities and found different characteristics for self-organized, planned and model (non-spatial) cities. The first three centralities presented in this paper are applied by Sergio Porta on the road-network of Barcelona, which will be compared to the Q-Index calculated in this work.

Table 1: Characterization and source of some network centralities

Name	Characterization	Source
Degree-centrality	Number of links connected to the given node.	Freeman 1979
Betweenness-centrality	Number of geodetics (shortest paths) running through a given node.	Freeman 1979
Closeness-centrality	Average (network) distance to every other node.	Freeman 1979
Information-centrality (Delta-centrality)	Relative drop in the network efficiency caused by the removal of the node.	Latora and Marchiori (2007)
Straightness-centrality	Average of the Euclidian distance divided by the network distance to every other node.	Crucitti, Latora and Porta (2006)

4.2 Finding communities in networks

As stated in several articles (e.g. Barthélemy, 2005; Radichi, 2004), a network approach has been applied on different topics like the internet, social networks, ecosystems and many more. In fact, very diverse systems can be described as complex networks (Radichi, 2004).

To describe and analyze such networks, communities are defined as "a subset of nodes within the graph such that connections between the nodes are denser than connections with the rest of the network" (Radichi, 2004). The classification of network-nodes in communities is different from the clustering of multivariate data (objects characterized by a vector of attributes), for which a number of efficient algorithms exist (Reichardt and Bornholdt, 2004). This is not the case for the discovering of communities in complex networks (e.g. networks with negative link weight).

Newman and Givan (2004) give an overview for some hierarchical clustering methods of social networks, which can be divided in two classes, agglomerative and divisive techniques. The article presents also three new divisive methods, but only for unweighted and undirected networks.

Radicchi et al. (2004) also describes two algorithms for social networks. But as stated in the article itself: "…our algorithm could be fruitfully applied also to nonsocial (disassortative) networks, although future work is needed in this direction" (Radichi, 2004).

Thus these methods are not applicable on the network of activities studied in this work (or only with significant effort to adapt them), since negative link weights exist.

For this reason, the method with the 'Potts Model' described in Reichard and Bornholdt 2004 is a good alternative for the classification of the activities, because it can be applied directly to the activity-network – as Jensen has done.

At this point I would like to mention the article '*Aggregation of retail stores*' from Jensen et al. (2005), which has a promising title for my work. This paper proposes a simple model to understand why some stores tend to aggregate and others not. It is found that "aggregation is more likely for stores selling expensive products and/or stores carrying only a fraction of the business variety" (Jensen et al., 2005). As I am mainly interested in describing and analyzing the characteristics of economic activities based only on their location (regardless of the reason for it), these findings won't be useful for the main part if the work, but will provide a comparison basis on the end.

4.3 Simulated annealing

Simulated annealing (SA) is a general optimization technique first described by Kirkpatrick (1983) and still used in different domains. Examples are the use for resource allocation (Aerts et al., 2002; Sante-Riveira, 2008), location problems for school network planning (Antunes 2001) and an optimization model for regional wastewater systems planning (Cunha, 2009).

Henderson et al. (2003) gives a detailed description in the "Handbook of metaheuristics", where diverse metaheuristic optimization algorithms are presented. But, as a non-mathematician, I failed to understand the exhaustive derivations and I missed practical examples (even if there is a chapter "Practical Guidelines" in Henderson et al. (2003)).

I finally found the article '*Optimization by simulated annealing - an experimental evaluation; part 1, graph partitioning*' by Johnson et al (1989). In fact, it describes the method used by Antunes (2001) and Cunha (2009). It presents a concrete SA-approach with suggestions and remarks on each parameter and compares its performance with problem-specific algorithms. And as the graph partition problem used as example is closely related to classification in networks (Newman, 2004), there is much chance that this approach is adapted for the problem of this work.

5 Development

5.1 Basic knowledge and definitions

In this section I will summarize some concepts and definition from the articles, which should facilitate the further reading of this report (for more details refer to the cited articles):

The so-called M-Index (M) is defined to describe the interactions between the categories of commercial activities: "The definition of M_{AB} at a given distance r is straightforward: draw a disk of radius r around each store of category A, count the total number of stores n_{tot}, the number of B stores n_B, and compare the ratio n_B/n_{tot} to the average ratio N_B/N_{tot}, where N refers to the total number of stores in town. If this ratio, averaged over all A stores, is larger than 1, this means that A 'attracts' B, otherwise that there is repulsion between these two activities." (Jensen, 2006)
So M_{AB} = mean(n_B/n_{tot}) / (N_B/N_{tot}). Then the attraction/repulsion-link a_{AB} is defined as the natural logarithm of M_{AB}. Therefore a network with attraction and repulsion is defined.

Then simulated annealing (explained in the following section) is used to form groups of categories which maximize the overall site satisfaction K which is defined as follows (Jensen, 2006):

$$K \equiv \sum_{i,j=1,nCat;i\neq j} a_{ij}\pi_{\sigma_i\sigma_j}$$

where σ_i is the group of the category i and $\pi_{\sigma_i\sigma_j} \equiv 1$ if $\sigma_i = \sigma_j$ and $\pi_{\sigma_i\sigma_j} \equiv -1$ if $\sigma_i \neq \sigma_j$ and a_{ij} is the link-coefficient defined above.

The location quality index Q for an activity 'i' is defined as follows (Jensen, 2006):

$$Q_i(x,y) \equiv \sum_{j=1,nCat} a_{ij}[\, n_{ij}(x,y) - \overline{n_{ij}}\,]$$

where $n_{ij}(x,y)$ represents the number of neighbor stores around x, y and $\overline{n_{ij}}$ the average number of neighbor stores j in the neighborhood of i.

5.1.1 Johnson's simulated annealing method

Before describing Johnson's method, I would like to quote Kirkpatrick (1983) on what is required for simulated annealing (SA):
"Four ingredients are needed: a concise description of a configuration of the system; a random generator of 'moves' or rearrangements of the elements in a configuration; a quantitative objective function containing the trade-offs that have to be made; and an annealing schedule of the temperatures and length of times for which the system is to be evolved." (Kirkpatrick, 1983)
In our case, a configuration of the system is simply every possible repartition of the categories in groups; a 'move' is the change of a randomly chosen category from one to another group and the objective function is K defined above. So the only thing that is missing is the so called annealing- or cooling schedule. But an effective cooling schedule is essential for reducing the time to find an optimal solution (Henderson et al., 2003).
The cooling schedule used in Jensen's code is defined by a starting temperature, a cooling ratio and an ending temperature. But the first and the last of these three parameters are particularly difficult to choose without if one does not have any experience using this algorithm. This is one reason because I used the method of Johnson (1989), summarized on the Figure 1.

1. Get an initial solution S.
2. Get an initial temperature $T > 0$.
3. While not yet *frozen* do the following.
   3.1 Perform the following loop L times.
      3.1.1 Pick a random neighbor S' of S.
      3.1.2 Let $\Delta = cost\ (S') - cost\ (S)$.
      3.1.3 If $\Delta \leq 0$ (downhill move),
         Set $S = S'$.
      3.1.4 If $\Delta > 0$ (uphill move),
         Set $S = S'$ with probability $e^{-\Delta/T}$.
   3.2 Set $T = rT$ (reduce temperature).
4. Return S.

Figure 1: Simulated annealing, from Johnson (1989).

Johnson describes an 'intelligent' cooling schedule with 4 parameters:

"INITPROB

Used in determining an initial temperature for the current set of runs. Based on an abbreviated trial annealing run, a temperature is found at which the fraction of accepted moves is approximately INITPROB, and this is used as the starting temperature (if the parameter STARTTEMP is set, this is taken as the starting temperature, and the trial run is omitted).

TEMPFACTOR

This is a descriptive name for the cooling ratio r of Figure 1.

SIZEFACTOR

We set the temperature length L to be N*SIZEFACTOR, where N is the expected neighborhood size. We hope to be able to handle a range of instance sizes with a fixed value for SIZEFACTOR; temperature length will remain proportional to the number of neighbors no matter what the instance size.

MINPERCENT

This is used in testing whether the annealing run is frozen (and hence, should be terminated). A counter is maintained that is incremented by one each time a temperature is completed for which the percentage of accepted moves is MINPERCENT or less, and is reset to O each time a solution is found that is better than the previous champion. If the counter ever reaches 5, we declare the process to be frozen." (Johnson, 1989)

The two important things in this cooling schedule is the method to define the starting temperature and the 'intelligent' stopping criterion. In fact, the starting temperature should not be too high, because if almost every move is accepted, only a random group-distribution is generated. Johnson got good results with INITPROB = 40%. The stopping criterion is 'intelligent' in the sense that it detects itself when the system is 'frozen', i.e. when there are only less than MINPERCENT 'moves' that are accepted. Johnson suggests a value of about 2% for MINPERCENT.

The two remaining parameters TEMPFACTOR and SIZEFACTOR control together the time taken from a given starting temperature to a final one. The tests done on these two parameters in Johnson lead to the statement that "[…] increasing TEMPFACTOR seems to be the preferred method for improving the annealing results by adding running time to the schedule." (Johnson et al., 1989)

5.2 Program optimization

In the following section, the optimizations and adoptions done to the prototype are described.

The calculation of the M-Index took a lot of computation-time because there were SQL-queries on a large table inside a loop (done $N_{categories\_total}^2$ times*)*. The solution I found was to do only one complex query before the loop and storing the values in an array to use them into the loop. The result was a reduction of the computation-time by a factor of about 10 for the calculation for this part of the program on two test-sets with 66 and 3467 activities and 12 and 46 categories respectively.

Moreover, as fore-mentioned, the SA-part was re-implemented using the method of Johnson described above, with these two differences:

- The 'cooling-value' is called *timestep* and has the inverse value of TEMPFACTOR (i.e. 1/r; because it was already used that way in the first SA-implementation, the temperature reduction step is in this case: T = T/*timestep*)
- No function to find the starting temperature for a given INITPROB was implemented. I just did a trial annealing run with a large *timestep* manually to find the temperature where the acceptance ratio was 0.4 or a little bit more (to be on the safe side) and defined it as TMAX (=STARTTEMP).

Here one uses the same initial solution as for the first implementation, obtained by a classification based on the strongest links. In fact, starting with a good solution based on the structure of the system instead of a randomly generated one can be advantageous for the quality of the solution and/or the running time (Johnson, 1989).
The results found with this new method were similar or better using generally less time than the first implementation. Another positive point is the fact that the results are also more stable.

Furthermore, the code was optimized by making it handier for multiple calculations and large datasets by adding automated saving and some 'helping scripts', which will be explained in the end of the next section.

5.3 Program description

Here I explain the procedures of the program and give remarks for their handling.
The code itself can be found in Appendix 9.10.

In the following section describing the program, the word 'store' is used for 'activity', because the scripts and its variable were at first only written for stores.

An overview of the important tables and scripts and how they are related is presented graphically in the next section. This should help to understand the following explanations.

The program can be separated into three parts:

The first part calculates the M-values, a-values and mean-values of the numbers of stores of each category around a given category. It consists of two scripts: *Fill_ResTable* and *Fill_MTable.*
The script *Fill_ResTable* counts the number of neighbor stores for each single store and saves the result in the table *ResTable*. To do so, it needs a 'Manifold-Drawing', named *Stores,* composed of all stores as points with the attributes *ID_pt* (unique integer), *Category* (text(10)), *CategoryName*

(text(160)). Furthermore, the parameter 'radius' (in working units) has to be defined. The script *Fill_MTable* then calculates the *M-value* and $\overline{n_{ij}}$ (*n_mean*) defined above from the entries in *ResTable* and finally computes the *a-values* from the calculated *M-values.* These calculations are done by summing up all the corresponding values for all combinations of categories from *ResTable* (which can become very large, its size is more or less proportional to $N_{stores\_total} * N_{categories\_total}$, but depends also on the locations of the stores). This is the part that was optimized as described above.

The second part is the 'grouping algorithm' which consists of the two scripts *Make_a_sym* and *Make_Groups*. The first one makes the a-values from *MTable* symmetric. This means that the attraction/repulsion coefficients between two categories are rendered equal. This is done by taking either the mean of both values if they have both the same sign, or taking zero if they have inversed signs (this means one is attracting the other, but this latter is repulsing the former). The results are stored in *a_symTable.*

There is a parameter *shift* which corresponds to the γ in Reichardt and Bornholdt (2004). This value is subtracted to all *a-values* and so changes the interaction to overall less attraction but more repulsion. But as noted in Jensen (2006), this value can remain null in our case, as there are already negative links in the network. It was not deleted from the script because on the one hand it does not slow down the script and on the other hand it could probably be used in some future experiments.

Furthermore, the script *Make_a_sym* changes the 'Category-Identifier' from 'Text(10)' to successive integers, which simplifies some manipulations in the next script, *Make_Groups.*

The second script of this part creates first 'temporary groups' by using the highest *a-values* (from *a_symTable*), i.e. the strongest attraction links. This process determines the maximal number of groups that can be formed. It can be analyzed afterwards by looking at the entries in *LogTable*, which indicates how categories were associated to groups according to their *a-values.*

This 'temporary groups' are then used as the initial solution for the simulated annealing. If the parameter *Johnson* is set as 'true', then the SA method described above with the parameters *timestep*, *TMAX*, *MINPERCENT* and *SIZEFACTOR* is used. And for each try, a category is chosen randomly and placed in the 'best' other group. If K is better for this new grouping, then this group is used. If not, it is only used by a certain probability depending on the temperature (as in any standard SA-approach).

If the parameter *Johnson* is set as 'false', the 'old' method (implemented after Jensen's code) is used, where the only parameter is the *timestep*, by which the 'artificial temperature' is divided after each try. The theoretical maximum of the overall satisfaction K (*Kmax*) is taken as starting value for the 'temperature'. After *ln(Kmax)/ln(timestep)* tries, the process reaches the temperature 1 and ends.

For both methods, the script saves the actual K-value and the category that changed (or not) in the *LogTable* for each try. Therefore the user can analyze afterwards what happened during the SA.

At the end, the best result is saved in the table *Groups.* This result includes the 'satisfaction' (*quotK*) of each category and the theoretic maximum of K for this category (*CatKmax*), so that the actual K-value for the category can be computed: *CatK = quotK*CatKmax.*

The parameter *rand* (Boolean) specifies if the random number generator is randomized before the SA or not. If it is not randomized, the same pseudo-random-sequence is used all the time (this can be useful for tests).

Furthermore, there is a little script called *AddFill_GroupID* with two parameters *Column* and *Add.* If the latter is 'true', a new column with the name *Column* is created in the table *Stores* and the

corresponding group numbers are inserted for all the stores. If the parameter *Add* is 'false', the group numbers are re-inserted without the creation a new column.

The third and last part concerns the calculation of the Q-Index. There are two variants: *Calc_QIndex* and *Calc_QIndex_perStore*. The first variant is to calculate Q for points contained in a 'Manifold-Drawing' *QIndex* for all categories in the table *Categories.* The second variant is a modified form of the first to calculate the value of Q for all existing Stores (in the table *Stores*).
Calc_QIndex needs points in *Qindex* with an attribute *ID_pt* containing a unique integer and the table *Categories* with the categories, for which the *QIndex* has to be calculated. If the calculation has to be done on all existing categories, the query *Fill_Categories* can be used.
First *Calc_QIndex* adds a column in *QIndex* for each category in *Categories*. Then it processes point after point in order to get the number of neighborhood stores $n_{ij}(x, y)$. Finally it computes the Q-value after the definition given above. The parameter *radius* specifies the size of the neighborhood to be taken into account.
Calc_QIndex_perStore works the same way, but without *QIndex* and *Categories*. It takes the points in *Stores* and their category for the calculation of Q. The store itself is not taken into account when doing the calculation of his Q-Index.
Both variants have the parameter *Nsave* which specifies the number of points to be processed before saving automatically the file (if *autosave* is set, which is explained later).

The scripts explained above are called by a script *Potts*, which also transmits the parameters to the scripts and saves the starting and ending time of each step in the *LogTable.*
In addition, there are the little scripts *MultipleCalc, ChangeR* and *Copy.* They are used to do automated calculations on different 'sets' and/or with different parameters. *MultipleCalc* changes the name(s) of the corresponding 'Manifold-Components' to have the desired 'set' and calls *ChangeR* which in turn calls *Potts* repeatedly, each time with different parameter-values. *Copy* creates a new folder for each execution of *Potts* and saves the necessary results in it. But these scripts need to be adapted for the data and the calculations that should be done each time they are used.

Furthermore, there is the global variable *autosave* which can be set as 'true' if the program is called from outside of the active file (i.e. as simple *VBscript* plug-in with the code in a text-file), so that the file is saved automatically after each script (or after each *Nsave* processed points for the third part).

The four queries *M_value*, *a_value, n_mean* and *a_sym* show the content of *MTable*/*a_symTable* as a matrix. With their help, these matrixes can be exported and analyzed.

Separately from these scripts, a little form *QIndex_viewer* has been developed to facilitate the visualization of the Q-values in *QIndex*. The form and the visualizing 'Map' are shown on Figure 2.
In the first field, the category can be chosen. The corresponding column-name appears in the *Formula*-field, in which the formula can be changed if desired (combine categories for example).
By clicking on *Calculate*, the result of the formula is saved in the specified column (normally the one used as 'Theme' on the 'Map') and normalized between -50 and 50 if the corresponding option is chosen. If *Zero rests Zero* is checked, the scaling is only done on the larger of the two extremes, so that the sign of all values is conserved. The option *View Stores* sets the values of the column *Visible*

to ‘true’, so that the corresponding stores are highlighted on the ‘Map’ (larger, blue dots on Figure 2). The ‘Select by year of creation’ enables to view the stores which match the criterion in red.

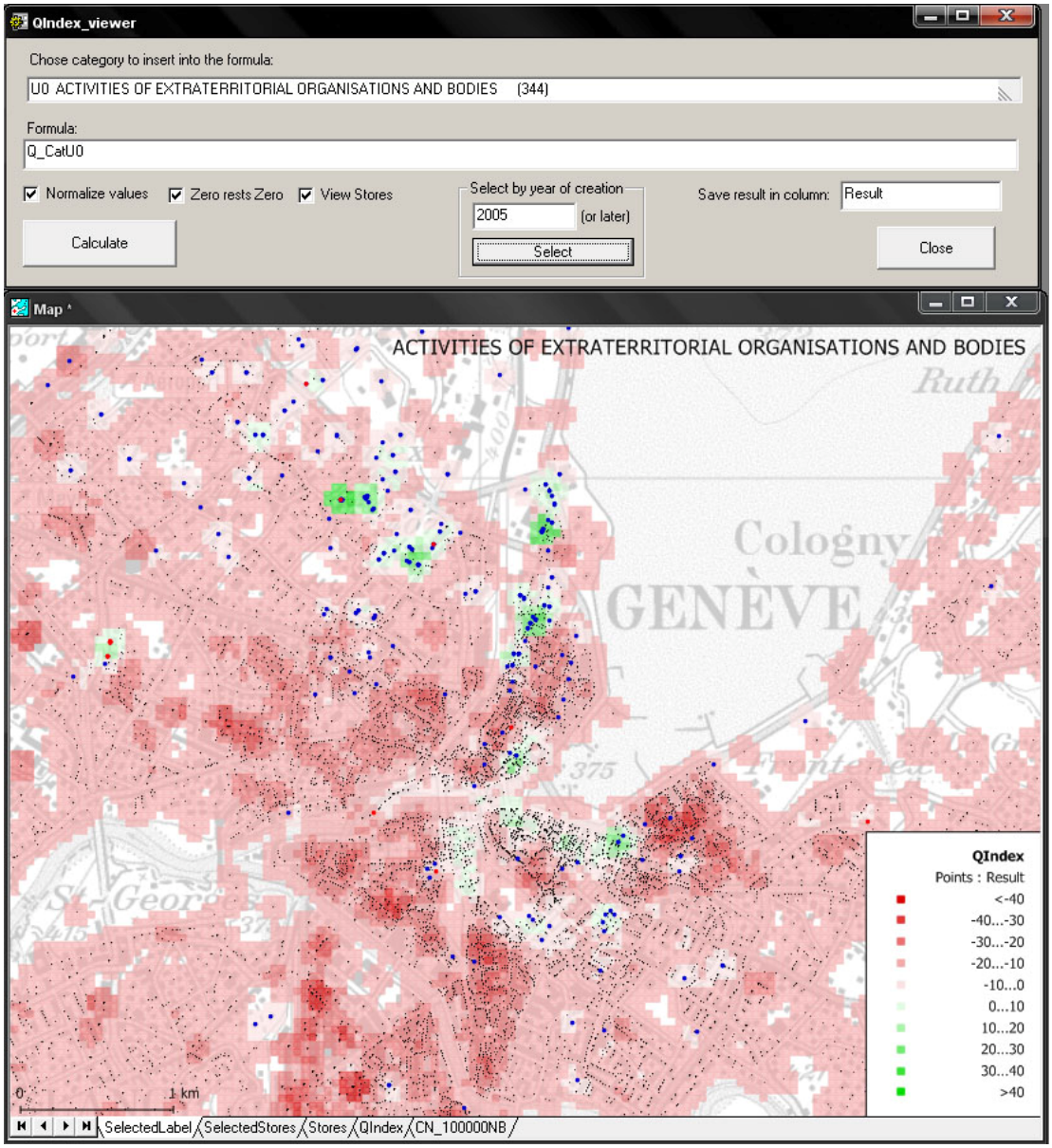


Figure 2: Form 'QIndex_viewer' and resulting 'Map'

5.4 Scheme of the program

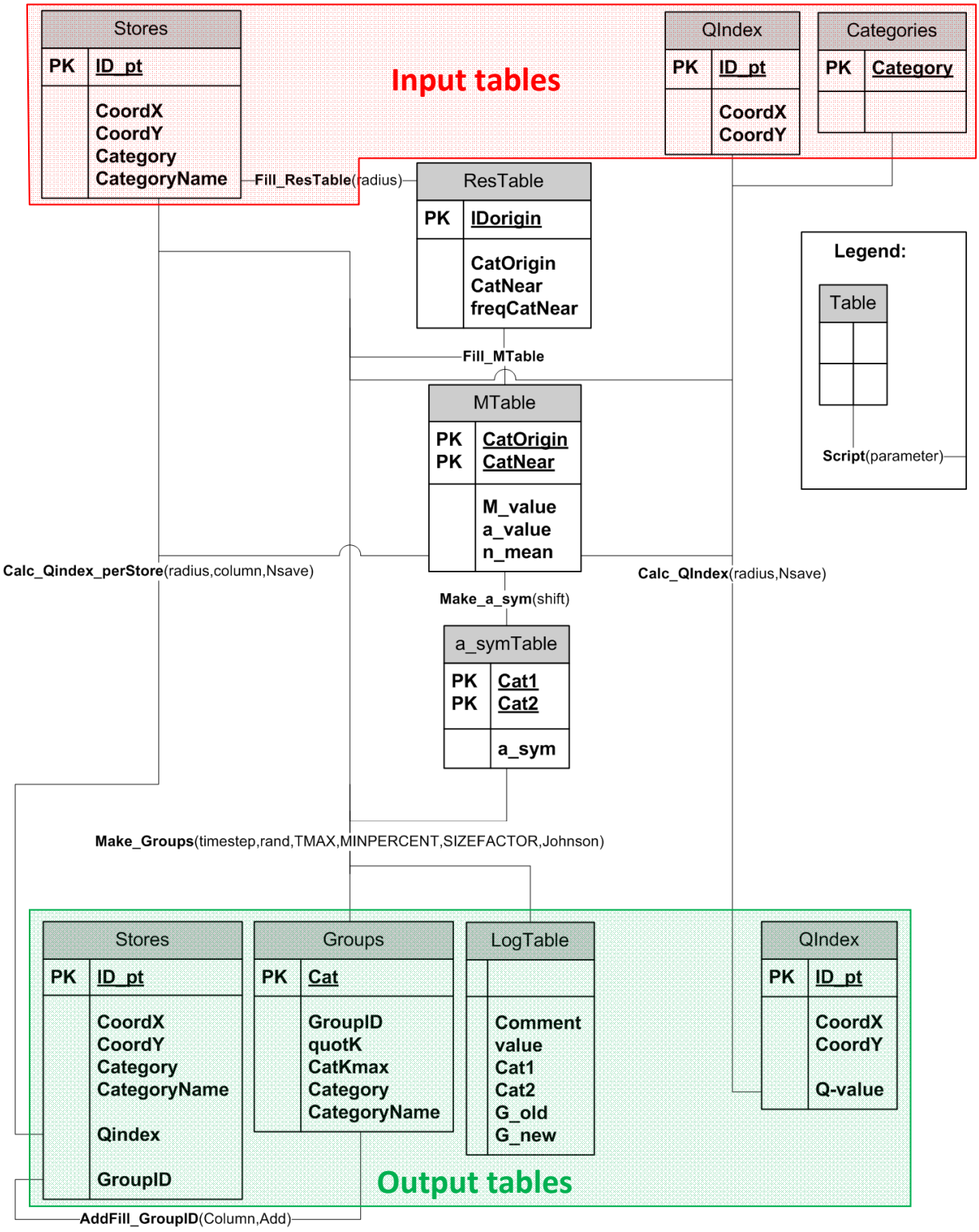


Figure 3: Scheme of the program

Remarks:

- For simplicity's and clarity's sake, some of the less important tables are not represented here.
- The tables *Stores* and *QIndex* are represented two times: in the input and in the output section (as column(s) added and filled by scripts).

6 Case studies

In this chapter, the results of the case studies on the two cities Geneva and Barcelona are presented.

6.1 Geneva

Geneva is a Swiss city (188'068 inhabitants) and also a Swiss canton (453'439 inhabitants in December 2008)[1]. More than 99% of the cantons population is living in the agglomeration of the city according to the 2000 federal population survey[2]. The canton has a special geographical situation, since it shares more than 90% of its boundary with France. Furthermore, the Lake Geneva separates the eastern part of the territory.

The activity database used for the calculations comprises all activities registered in December 2008 for the whole canton. Their location is given with address precision (2-3m) and their category in the Swiss 'General Classification of Economic Activities' (NOGA). There are many other data attached to each activity (name, employees …), but I used only the 'year of creation' for the validation.
Two sets of activities were used:

- The retails stores of Geneva using the most detailed NOGA-level ('NOGA-6') as categories. But some of the categories where regrouped together so that each category comprises more than 15 activities. This set contains 4506 activities in 45 categories.
- All activities of Geneva using the topmost level of the NOGA (the sections). This set contains 40330 activities in 18 categories.

A detailed list of the categories of these sets can be found in Appendix 9.1.

The calculations are always done for the whole canton, but the analyses focus on the city of Geneva.

6.1.1 Retail stores of Geneva

6.1.1.1 Sensibility analysis

With the set 'Geneva retail-stores' and for 4 subsets of it (33% randomly chosen activities of the whole set), I calculated the a-values and a_sym-values for 6 different radiuses (25, 50, 100, 150, 200, 500m).
Copies of these five sets containing retail stores were modified to a hectometer-grid-model ('hecto') by rounding the coordinates off. The calculations were repeated with radiuses of 50, 101, and 150 meters. The radius 101m may seem special, but it was chosen instead of 100m so that the next grid points (which are exactly 100m distant) are clearly included.
The resulting 45 a-matrixes and a_sym-matrixes (with each 2025 link-values) were analyzed in two ways:

- Whit descriptive statistics and counting of zero-, positive- and negative-links resumed in a table as illustrated in Table 2 and Table 3. The tables for all activity-sets can be found in Appendix 9.2.
- By calculating the (linear) correlation-coefficient for all combinations of matrixes and representing the results as color-coded correlation-matrix like on Figure 4 for the a-matrixes. The same figure for a_sym is in Appendix 9.2.3.
  Note that these figures are symmetric, which facilitates their reading.

[1] http://www.ge.ch/statistique/tel/domaines/01/01_02_1/T_01_02_1_1_2008.xls [11.6.09]
[2] http://www.are.admin.ch/themen/agglomeration/00667/00673/index.html?lang=fr [11.6.09]

The results of these analyses are presented in the three following sub-sections, one for each parameter. The first is the radius, followed by the random subsets and the comparison of the 'hecto'- and address-location-model.

Table 2: Descriptive statistics of the a-matrixes for the activity-subset 'Retail stores 33% A'

activity set	Retail stores 33% A					
radius (m)	25	50	100	150	200	500
# zero links	1147	759	302	152	92	6
# non zero links	878	1266	1723	1873	1933	2019
# positive links	643	765	903	929	946	1039
# negative links	235	501	820	944	987	980
# strong links (> 0.5)	450	464	430	365	342	226
# strong anti-links (< 0.5)	100	228	429	460	467	370
# strong anti- and links	550	692	859	825	809	596
Maximum link	3.98	2.83	1.81	1.67	1.54	1.08
Minimum link	-2.73	-3.32	-3.07	-2.79	-3.48	-2.81
Overall mean	0.22	0.13	-0.02	-0.07	-0.10	-0.09
Sum of positive links	575.1	549.6	493.5	437.5	403.4	330.3
Sum of negative links	-133.8	-295.4	-537.0	-588.1	-612.5	-511.9
Mean of postive links	0.89	0.72	0.55	0.47	0.43	0.32
Mean of negative links	-0.57	-0.59	-0.65	-0.62	-0.62	-0.52

Table 3: Descriptive statistics of the a_sym-matrixes for the activity-subset 'Retail stores 33% A'

activity set	Retail stores 33% A					
radius (m)	25	50	100	150	200	500
# zero links	1435	1247	923	873	919	953
# non zero links	590	778	1102	1152	1106	1072
# positive links	494	512	584	562	524	562
# negative links	96	266	518	590	582	510
# strong links (> 0.5)	400	376	344	280	228	146
# strong anti-links (< 0.5)	60	160	356	350	352	252
# strong anti- and links	460	536	700	630	580	398
Maximum link	2.88	2.03	1.54	1.48	1.34	0.97
Minimum link	-1.88	-2.33	-2.32	-2.08	-2.66	-2.06
Overall mean	0.20	0.11	-0.02	-0.05	-0.07	-0.04
Sum of positive links	471.1	413.4	358.8	304.1	266.4	216.8
Sum of negative links	-61.7	-189.3	-390.6	-401.2	-414.1	-293.6
Mean of postive links	0.95	0.81	0.61	0.54	0.51	0.39
Mean of negative links	-0.64	-0.71	-0.75	-0.68	-0.71	-0.58

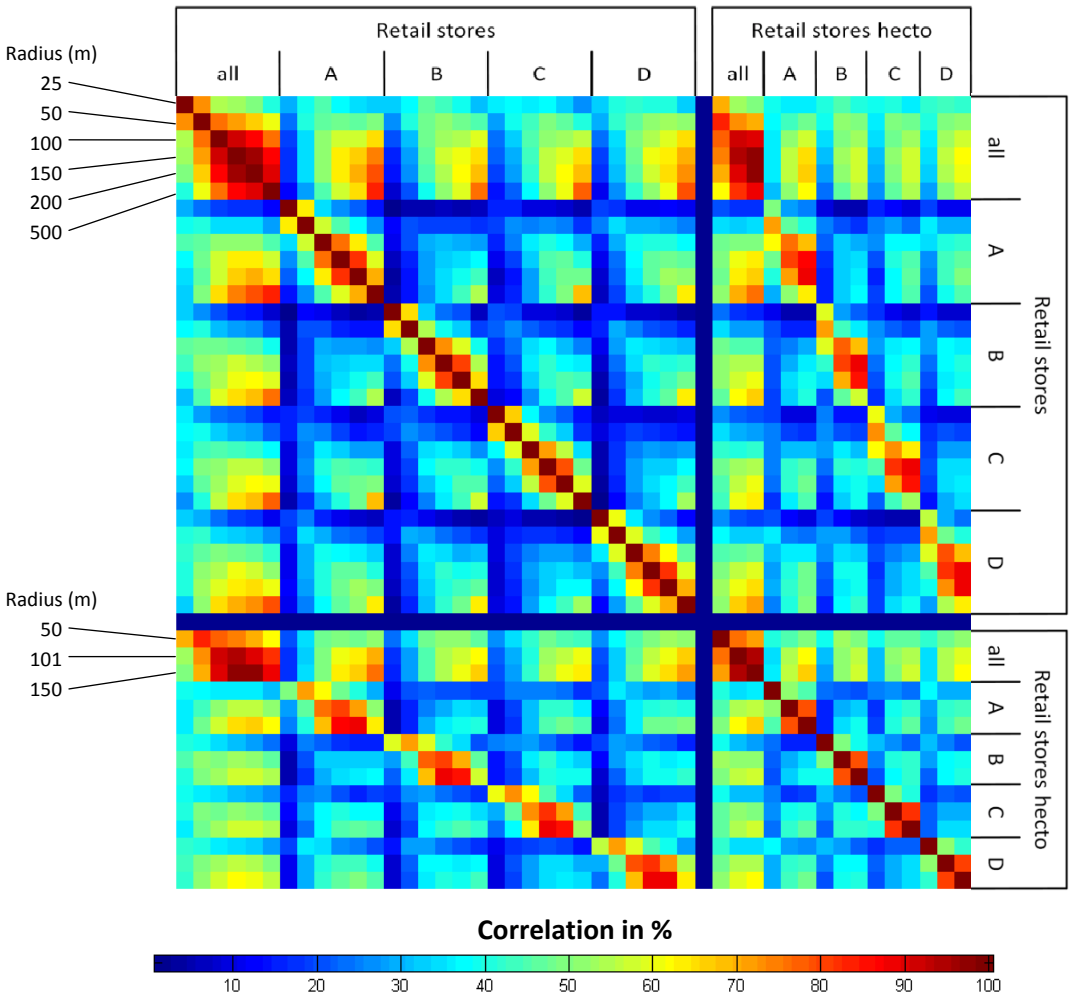


Figure 4: Correlation of the a-matrixes for 'Geneva retail-stores'

6.1.1.1.1 Radius

In Table 2, we can see that the number of positive and negative links increases with growing radius (while zero-links are decreasing). This is easily explained since there are always more neighbors included, therefore more and more categories 'reach' each others. On the other hand, the value of the maximum link and the sum and the mean of the positive links decrease. The sum of negative links increases while the mean of the negative values remains rather stable until 200m. Strong links have a maximum with a radius of 50m and decrease afterwards while strong anti-links increase until they reach the maximum at a radius of 200m. So one can conclude that there are more and stronger positive links for small radiuses compared to the negative links. Moreover there are more but generally weaker links for large radiuses.

The tendencies outlined here for the subset 'A' on Table 2 are mostly true for all four subsets, as can be verified in Appendix 9.2.1. But it is not the case for the original set ('all'). This comes from a difference in the activity-density, as there are three-times more activities in this set than in its subsets.

If we look at the statistics of the symmetric link values (a_sym-values) in Table 3, we see that there are much more zero-links than for the original link values (a-values). This comes from the fact that opposite link values are set to zero when the a_sym-matrix is calculated. The positive and negative

sums as well as the maximum and minimum links are lower (in absolute values). On the other hand, the links and anti-links are in general stronger, and the mean of the positive and negative links are both higher (also in absolute values). That is because a lot of values near zero were set to zero and that the values left are the mean of two links.
Since these symmetric values will be used in the classification part, I chose the 'best radius' on their basis: The number of zero links decreases at first, and then increased with growing radiuses. For the subset 'A' represented here, the minimum zero links is reached with radius=150m. It is also 150m for 'B', but 200m for 'C' and 'D' and only 50m for the whole set ('all'). I chose a 'best radius' of 150m for all subsets. I took 100m for the whole set, since there are slightly more zero-links than for 50m but it is less 'local' and therefore one can choose a larger spacing for the Q-index-grid afterwards (and so less points have to be calculated).

On Figure 4, we can see that the a-matrixes of similar radius values are correlated more than distant ones, which is an obvious result. Yet this effect is stronger for large radiuses, e.g. the 500m- and 200m-a-values are generally more correlated than the 50m- and 100m-a-values, even if the radius-difference of the first pair is much larger. This can be explained by the fact that for larger radiuses, the link values approaches zero (exactly zero for a very large radius covering all stores), so that the differences are smaller.

6.1.1.1.2 Random subsets

It is difficult to draw conclusions for the subsets from the descriptive statistics (Appendix 9.2). One can only say that all the values clearly vary among the different subsets. Yet one can not at all state whether this is significant or not.
On the other hand, Figure 4 – with the correlations between all the variants – lets us discover interesting facts: the correlation between the subsets is rather poor, generally only 50% or less, with the exception of the 500m-radius-values, which reach 60% to 70%. This good correlation of large radiuses can be explained by the same fact as above (general trend towards zero, thus towards little link values). But the bad correlation of e.g. only about 40% for 150m means that the random choice of subgroups introduces non negligible differences.
The correlation between the whole set and its subsets is higher, yet still not that strong.
These results do not encourage the use of the subgroups for further calculations, even if this particular whole set ('all') has the disadvantage of the geographical peculiarities. However we will see how the differences influence the results of the classification.

6.1.1.1.3 Hectometric and address location

As for the random subsets, it is not easy to compare the hecto- and address-location-models on the basis of the descriptive statistics in Appendix 9.2. But it can be said that the variations from hecto to address are smaller than the variations between the different subsets. Furthermore, for the a-matrixes, there are less no-zero links for the hecto-model and the mean of the positive links is generally lower, while the mean of negative links remains about the same. Also, the sum of the positive and negative links is mainly lower (so more repulsion in general). The same trends as observed here were noticed for a growing radius. One can thus assume that the hecto-values represent somehow a larger radius. This can be explained by the fact that the stores in the hector-model were concentrated on one point, so that the 101m radius includes also points that were originally further away than this distance.

Yet all this does not explain if the results from the hecto- and address-model are really similar. Thus we have to look at the correlations on Figure 4: for the 50m-hecto, the correlations lie around 75%, for the 101m around 80% and for the 150m around 90%. All these values are remarkably higher than between the different subsets. It is also interesting to note that the hecto-values are correlated similar or even higher with the address-values of the next higher radius (of the corresponding set, of course). So the 150-adress-values are all correlated around 85% with the 101m-hecto-values. This confirms the guess made before that the hector-values represent a larger radius in the address-model. Moreover, the correlations of the set 'all' are even higher than the ones mentioned for its subsets. In this set, the 150m-adress- and the 101m-hecto-values have a correlation of 95%. And the 150m- and 200m-address- and the 150m-hecto-values have even a correlation of 98%.
On the basis of this comparison, the use of hectometer-precision-data seems to be adapted to this program. We will see afterwards how it is well adapted to the classification and to the location quality index.

6.1.1.2 Classification

The classification was done for all five 'sets' with the simulated annealing method described above. Different parameter-sets were used. TMAX was always fixed to 5 (value obtained by a trial run for about 50% acceptance rate) and SIZEFACTOR to 1 while MINPERCENT was either 1% or 0.5% and 'timestep' was set to 2, 1.5, 1.4, 1.2, 1.1 and 1.05. There were at least 12 runs per set (sometimes more, so some parameter combinations were used twice). The grouping of the best result was conserved. These groups can be found in Appendix 9.3. Since the analysis and the comparison of these tables is difficult, a graphical representation of them was elaborated as follow:

A grid of 40x40m was created over the whole area. Then a group was assigned to each grid cell, giving the priority to the smaller groups (so if more than one store was in a cell, the group with the smallest number of stores (in total) was chosen). This method is not the reality-nearest one (this would be the use of the mode), but it has been chosen so that small groups do not disappear.

Since this colorful image is not that easy to read (Figure 5), it has been simplified with the function 'Image-Simplify' in Manifold (with parameters: level=0, neighbors=1, iterations=1). But since the result (Figure 6) was not satisfying (too much information was lost), a special mode-filter was coded and applied to the image in MATLAB. This 3x3 moving-window-filter takes the mode of the values (group-numbers), without taking into account the background (zero-values). Thus the region of colored pixels grows. As the filtered image was still overloaded (Figure 7), another simplifying (with parameters: level=0, neighbors=3, iterations=9) was applied to get the final result (Figure 8).

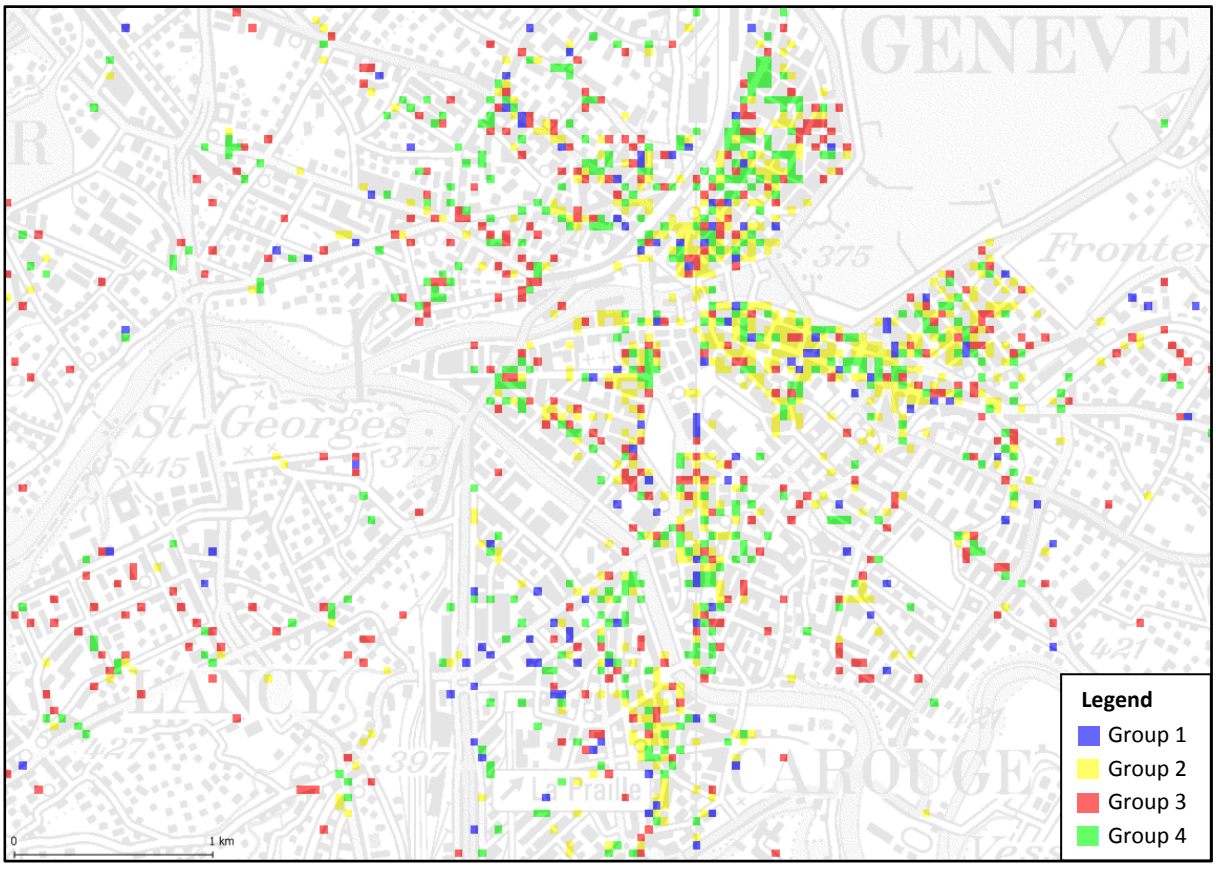


Figure 5: Groups of the set 'all ' represented on 40m-grid

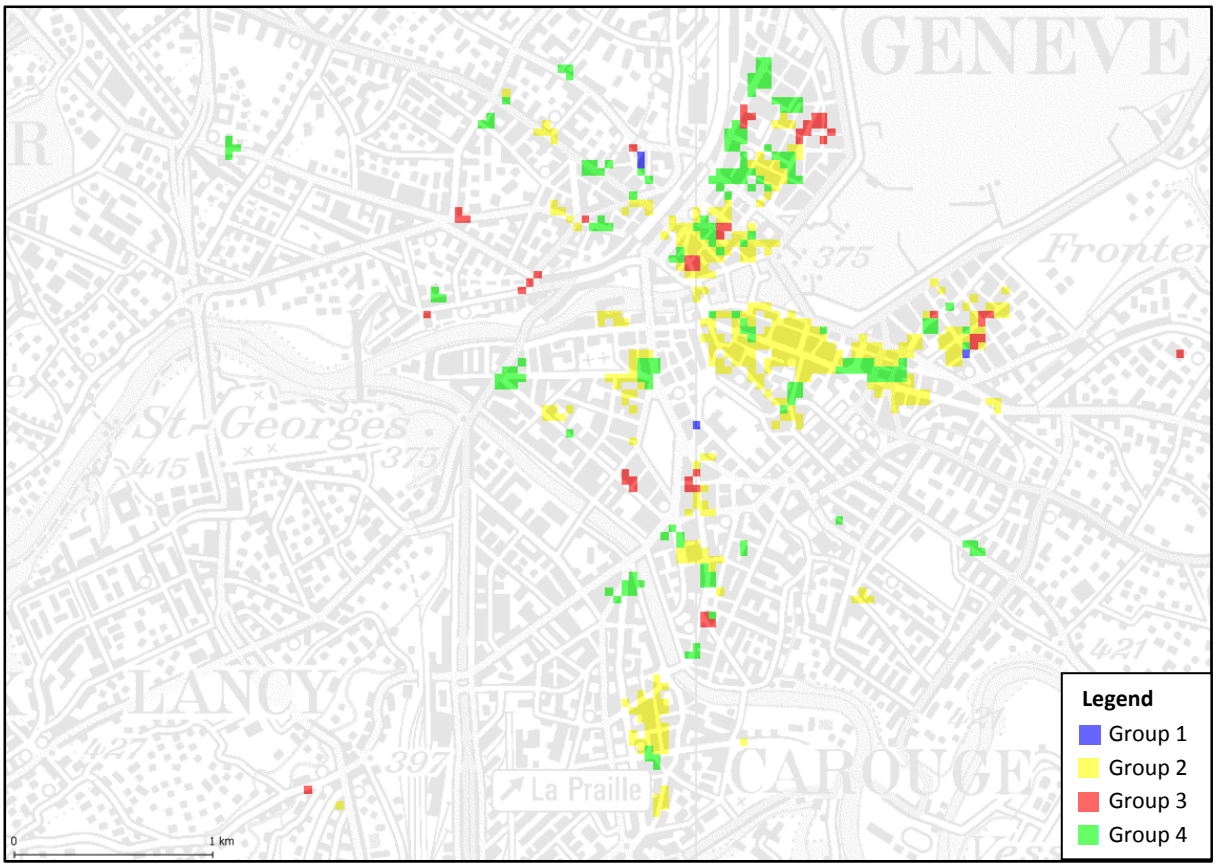


Figure 6: Groups of the set 'all ' represented on 40m-grid – simplified

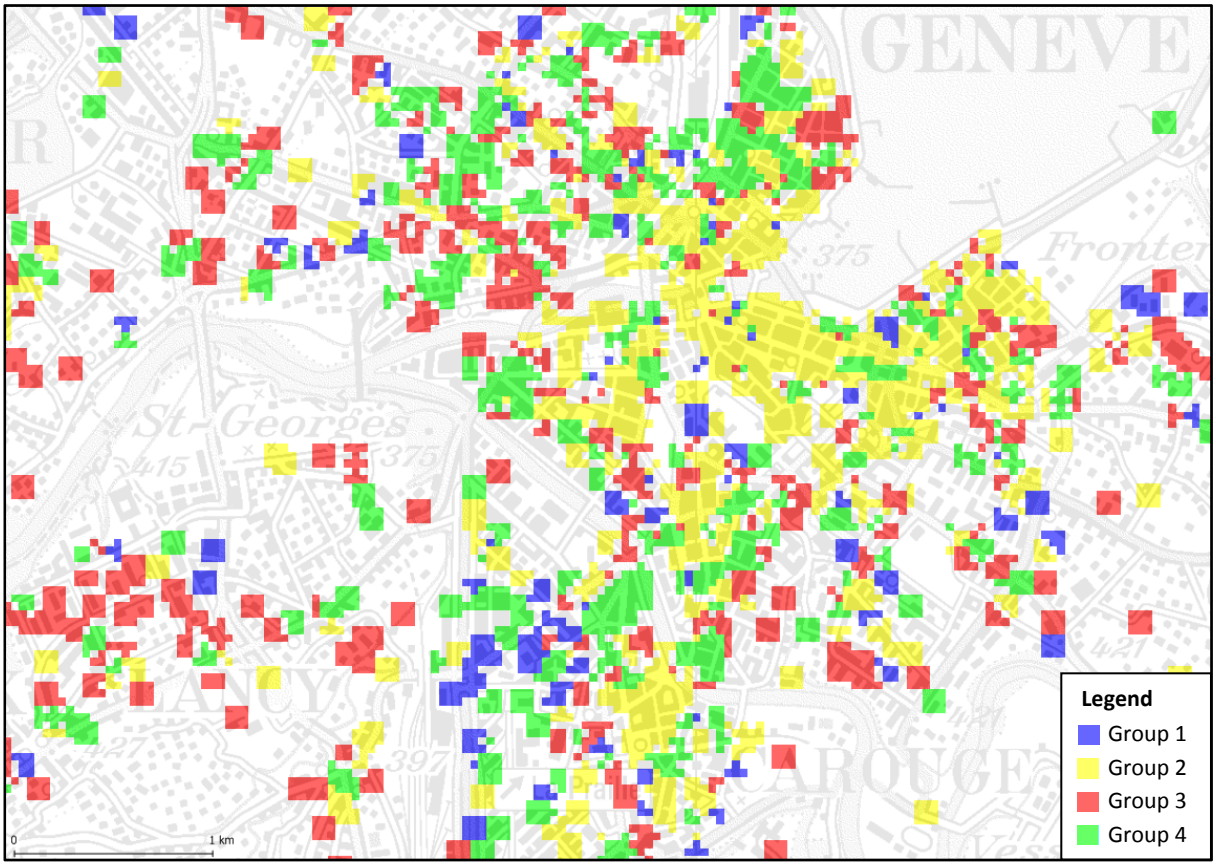


Figure 7: Groups of the set 'all ' represented on 40m-grid – filtered

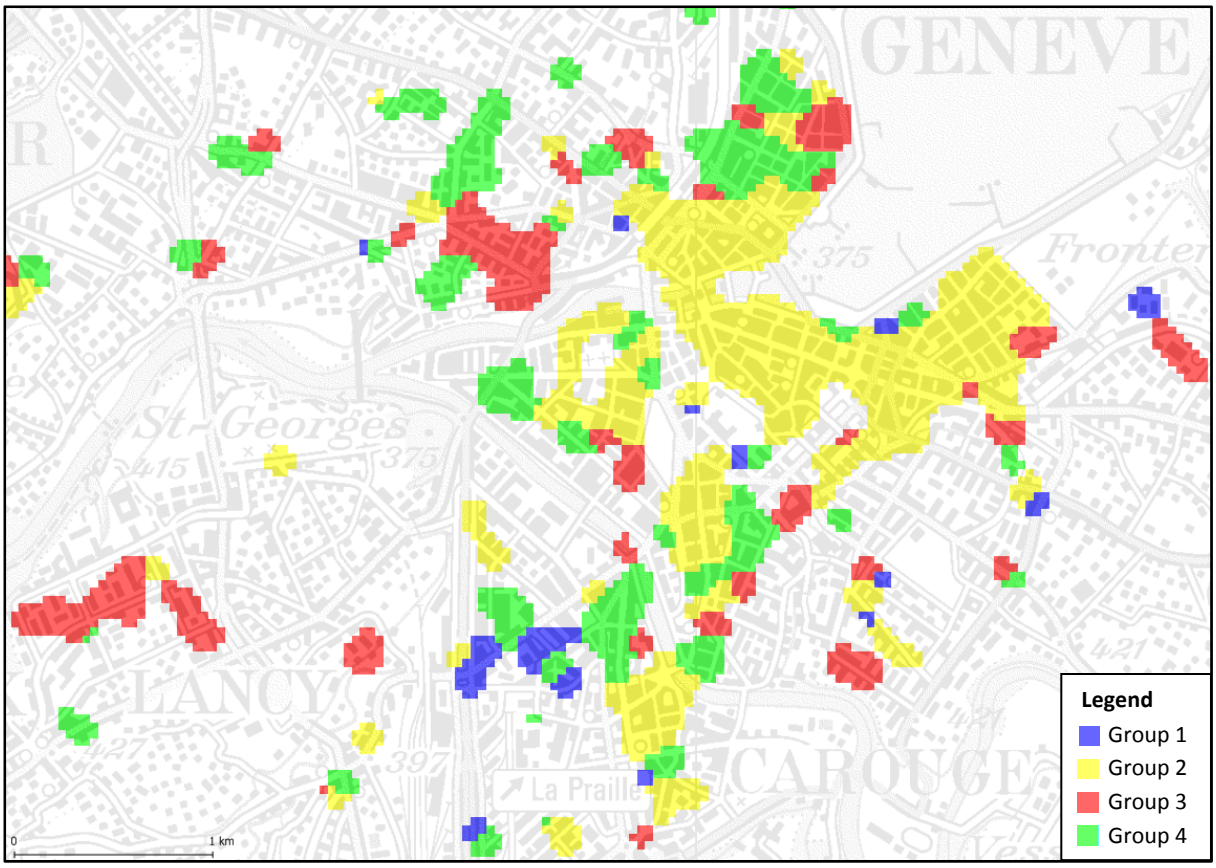


Figure 8: Groups of the set 'all ' represented on 40m-grid – filtered and simplified

Group 2 clearly dominates and is occupies 'the center'. It is evidently the largest group as it is composed of 27 out of 45 categories and of 2451 out of 4506 stores (as can be counted in Appendix 9.3.1). It comprises the stores selling clothes, footwear, watches and jewellery, textiles, video equipment and leather, to name only the ones with the highest 'satisfaction' to belong to this group. The furniture-stores, second-hand-good stores and retail sale of bread and confectionery are also in this group, but with less satisfaction (*quotK*). Most of the stores in this group sell expensive goods and/or have only a little part of business-variety. According to Jensen et al. (2005), these stores tend to cluster, which is the characteristic of the group 2.

Secondly, group 4 is found on a lot of large and small spots. It comprises stores like filling stations, paint- and hardware-stores, beverage-stores and bakeries (the ones with the most satisfaction). Newspaper-sale, butcher's shops and supermarkets (='non-specialised stores with food…') are also in this group. If we look on Figure 5 or Figure 7, we can see that this group is rather dispersed. This is a characteristic that matches the categories of this group. In fact, these shops sell rather inexpensive goods and cover most of the business variety, thus have the opposite characteristics of the ones who tend to cluster of the group 2.

Group 3 has two large spots outside the center and diverse little spots dispersed everywhere. This group is composed of the retail sale via mail- or internet-order, computer-shops, the sale of musical instruments and the sale of medical goods. The last two categories have only 24 and 22 stores respectively, thus the first two dominate. I think that, for all these shops, the location does not matter much (since they are very specialized or send their goods to the customer).

Finally, the group 1 is only present on some little and very little spots. It is composed of the retail sale via stalls and markets, 'other retail sale not in stores, stalls or markets' and the pharmacies. The

first are satisfied to 90% or more, while the pharmacies are only satisfied to 58%. With the exception of this last category, the stores mentioned do not have the means to rent a premise, so they are located where place is given to them.
Overall, it can be said that four groups with rather different characteristics could be found (only on the basis of their location) and the places where they are present by a majority are represented in Figure 8. Even if this figure has clearly lost a lot of information compared to the original 40m-grid (Figure 5), it gives valuable and fast readable indications. More detailed location patterns for each category can be analyzed with the location quality index in the next section, after the comparison of the classification between the random subsets and the hecto- and address-model.

6.1.1.2.1 Random subsets

The tables of the groups, the original- and the filtered+simplified-representations of the groups for every subset are in Appendix 9.3. When comparing them, one dominant group comprising the clothing-, footwear-, jewellery-stores and others from the group 2 described above is always found. Yet beside this, it is difficult to see clear parallels. Thus the random choice of the subsets induces much variability between the subsets and their usefulness is therefore questionable.

6.1.1.2.2 Hectometric and address location

The classification has also been done for the sets 'hecto-all' and 'hecto-A' (with several runs as described above). The best groups are presented in tables in Appendix 9.3.6 and 9.3.7 respectively. By comparing them to the corresponding address-location-results, a surprising similarity can be found: for the whole set ('all'), only one category is in another group. The category of furniture stores is in the 'clustering group' (group 2) for the address-model, but it is in the 'dispersive-group' (group 3) in the hecto-model. In fact, this category is the least satisfied in both models, so this little change is easily possible.
For the subset A, there are a few more differences, there is one group less in the 'hecto-model', but a larger one composed of two little ones of the address-model. Moreover, two categories are in another group.
This strong similarity confirms the fore-mentioned statement that the hecto-location-data is also suitable for these calculations. But it has to be noticed that the graphical representation with this data would be less informative, as a 100x100m grid has to be chosen.

6.1.1.3 Location quality index

The location quality index has been applied on the whole set of retail stores with a radius of 100m. The values were computed for each category at each point on the 50m-grid over the whole area (grid-points more than 100m distant from all activities were deleted, as they do not 'catch' any activity).
To visualize the results, the 'QIndex-Viewer' has been used and the values have all been normalized to -50 or 50 (yet keeping the sign of all values, so only the larger extreme was set to -50 respectively 50).
A selection of the 45 maps is presented and analyzed in the next sub-section.

6.1.1.3.1 Analysis

On nearly every map, the center (where most of the stores are located) stands out. For most of the categories it appears as green, which indicates that the location quality index is positive there.
One example for a very 'center-liking'-category are the jewellers' shops represented on Figure 9.

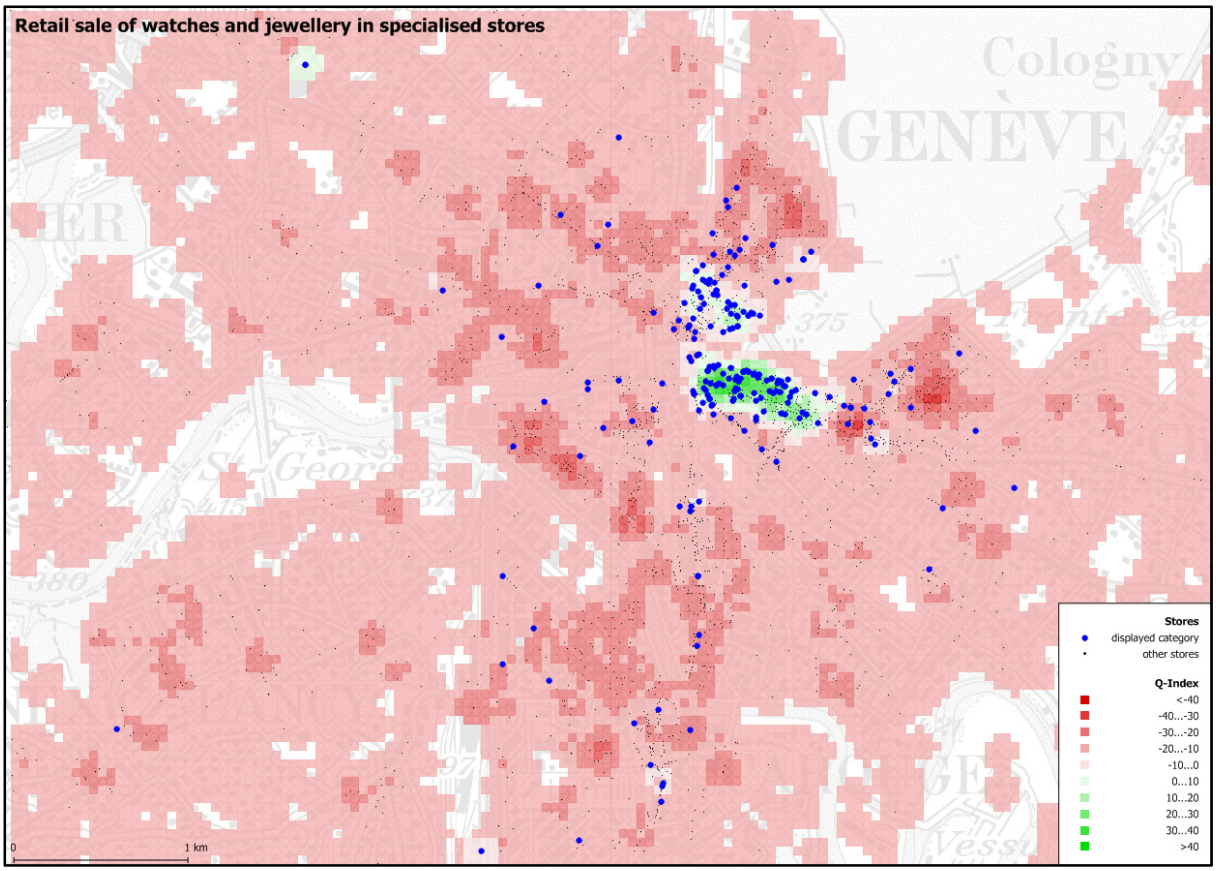


Figure 9: Q-Index-map for jeweller's shops

One can note that most of the jewellers' shops are located in this center (composed of a little area in the north of the Rhone River and a larger one in its south). Moreover, the Q-Index-values are far below zero if this center is left. Just below the word 'jewellery' of the title, there is a little green spot which seems to have very similar characteristics than the center. I guess it is a big commercial center with a wide variety of stores.

There are 7 other categories that have very similar patterns like the jewellers'-shops on Figure 9: stores selling cosmetic, four categories selling clothes, footwear stores and stores selling leather and travel goods. All these categories belong to the dominating clustering-group found for the classification, mostly with a very high satisfaction. Thus the two results confirm each other.

There are a lot of categories liking the center, which do not dislike the other regions as the 8 categories cited above. There are different levels of 'disliking' the surroundings. The jewellers' shops are the extreme example, as they completely dislike the surroundings. They are followed by the textile- and book-shops (maps in Appendix 9.4), and finally by the bakeries (showed on Figure 10). The bakeries belong to the 'dispersed group' from the classification. This is verified by looking at their distribution (blue dots). But according to the Q-Index they prefer the center. This tendency is not as marked as it is on the first sight with the red-green contrast, since nearly all the negative values are in the -10 to zero range. Nevertheless the Q-Index does not really reflect the location pattern of this category. I think this can be explained by the very different activity-densities between the center and its surroundings. The fact that some of the stores are near or in the center increases the n_mean-values. It is therefore difficult for the locations in the surroundings to reach this number of neighbors.

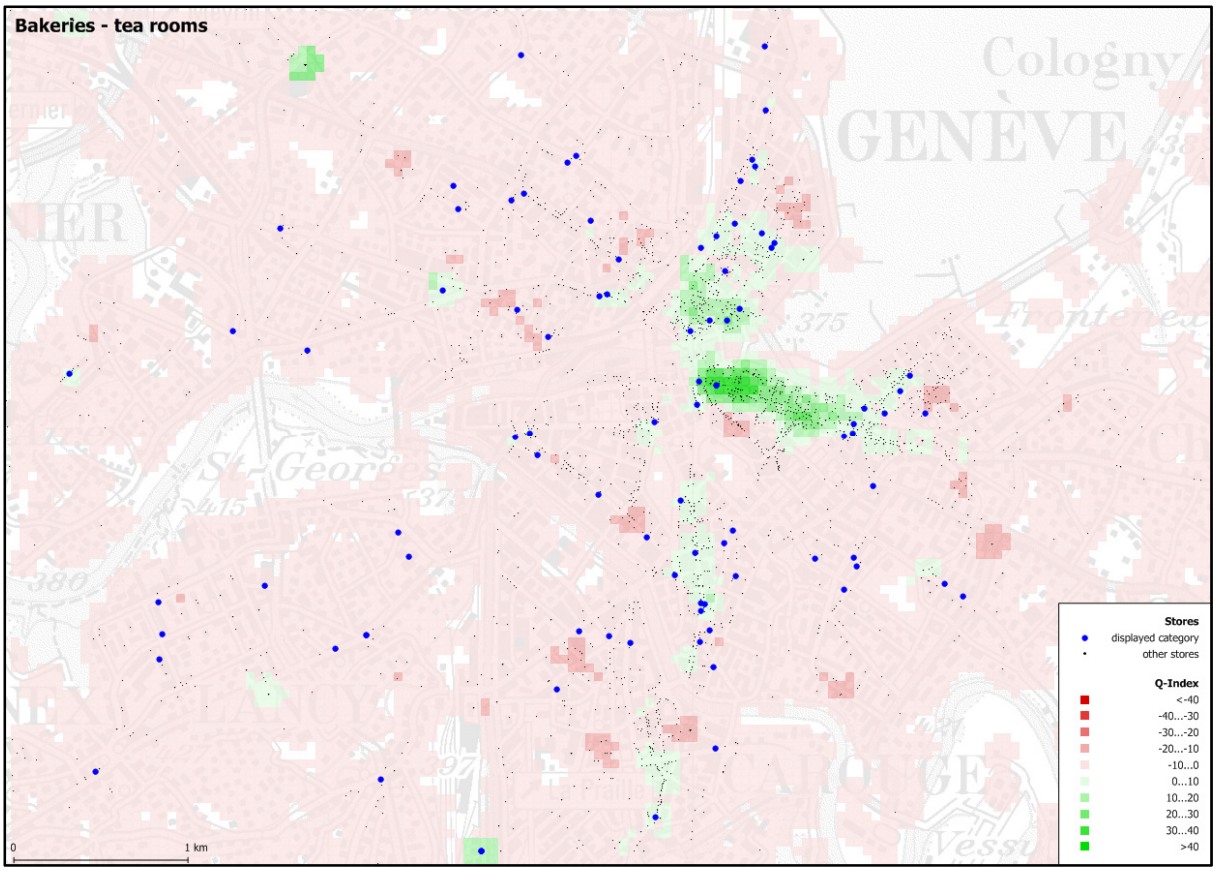


Figure 10: Q-Index-map for bakeries – tea rooms

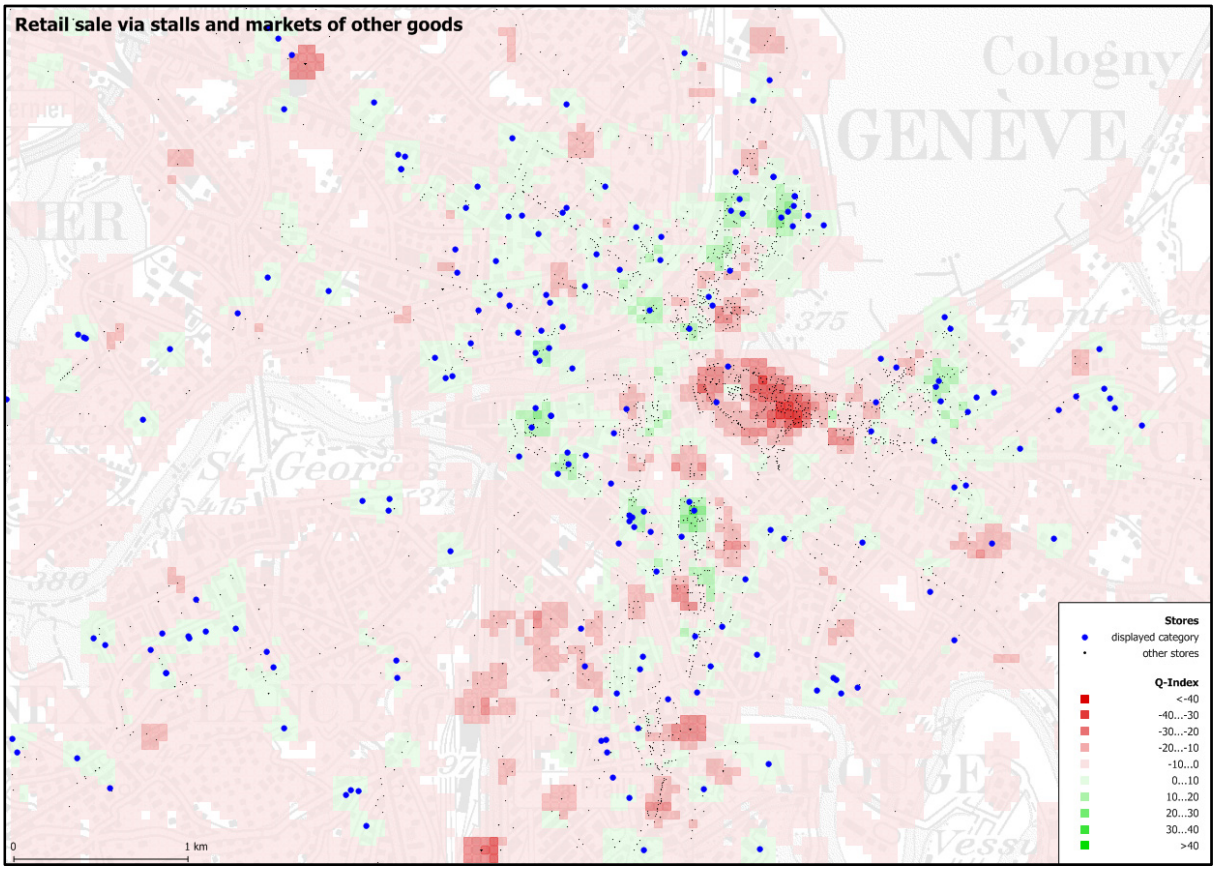


Figure 11: Q-Index-map for retail sale via stalls and markets of other goods (than food)

A way to take these different densities into account would be an adaptive radius or the taking of the median instead of the mean. Yet to adapt the code and recalculate all the steps would go beyond the limits of this work.
Some categories dislike the center, such as 'retail sale via stalls and markets of other goods' (than food) showed on Figure 11. The same is true for shops selling hardware, paint and glass (map in Appendix 9.4).
As illustrated in Figure 12, art trade is also very 'center-fixed', but not exactly on the same center as most of the other categories. The same is true for the stores selling antiques, which have a very similar pattern, as can be verified in Appendix 9.4. There seems to be an ideal quarter for such stores.

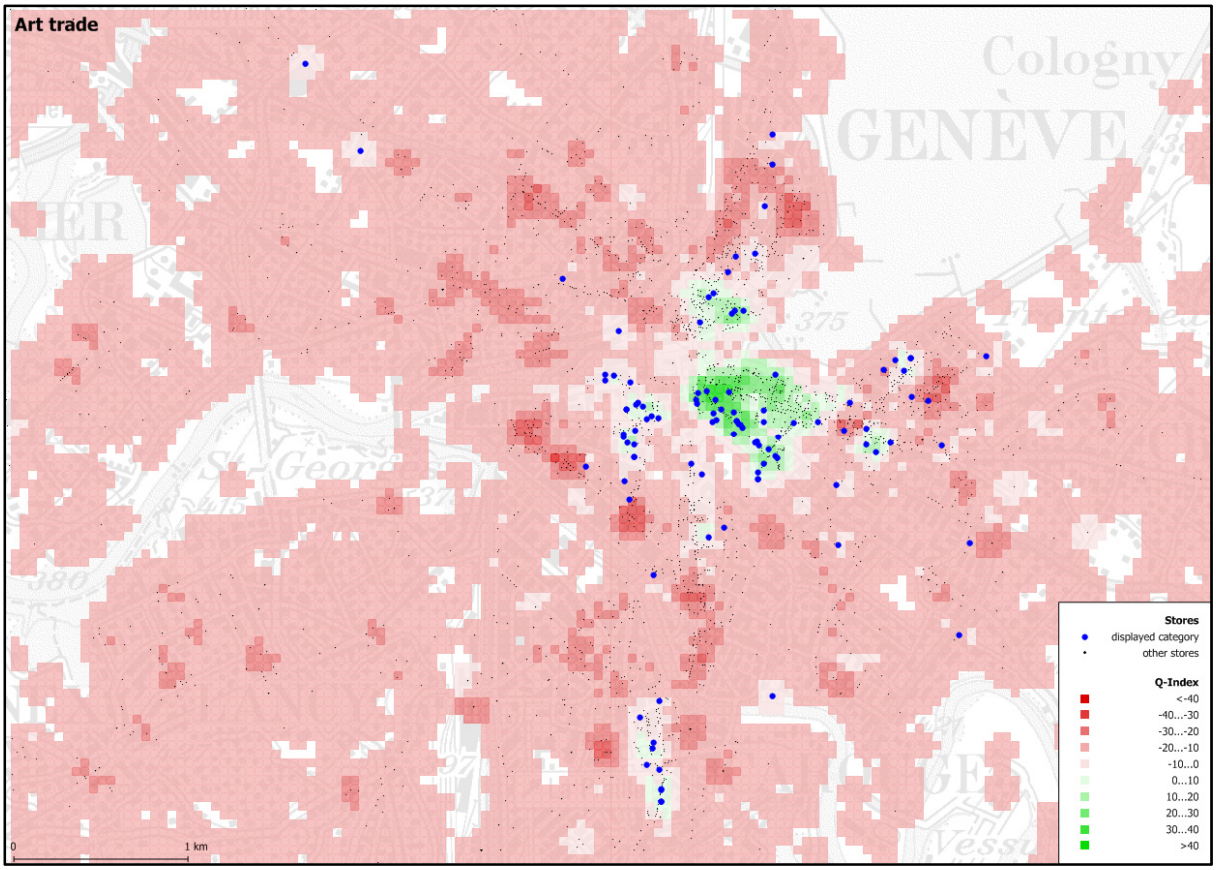


Figure 12: Q-Index-map for art trade

Thus the different location-patterns and preferences can be shown with the location quality index. Yet can the Q-Index - which is (only) a reflection of the actual activity-distribution - really be used to evaluate the location quality? This question is treated in the following sub-section, the validation.

6.1.1.3.2 Validation

To validate the Q-Index, the same method as Jensen (2006) has been used: The Q-Index-values of stores who closed since 2001 are compared to the one of newly opened stores during the same period. To do so, data of closed stores is needed, but this is lacking in the database. A complete database, including closed stores, is available online[3], but it only allows the search for keywords. The closed stores have to be searched manually. I used their postal address to find their coordinates

[3] 'Répertoire des Entreprises du canton de Genève' (http://reg.ge.ch)

manually[4] (since no coordinates are published online). I have chosen the butcher's-stores to do all this work on their category. When the data-collection was done, it was imported in Manifold and the Q-Index was computed for all the points. For the newly opened stores, the database and the calculated 'Q-Index per store' could be used. The results are presented as simple statistics of the obtained Q-values in Table 4.

Table 4: Q-Index statistics for closed and opened butcher's stores since 2001

	closed	opened
# stores	22	30
mean	-2.6	0.2
median	-2.3	-2.2
min	-6.1	-4.8
max	0.8	17.2

All the values but the mean are clearly higher for the opened stores. The median is only little higher. These positive differences are an indication that the Q-Index is a good measure for location-quality. Yet this is neither confirmation nor a proof. There should be much more validation-data to actually assert this.

As the 'Q-Index per store' was calculated for all stores, I created a map with all of them colored after their Q-Index-value (normalized). The result is shown on Figure 13.

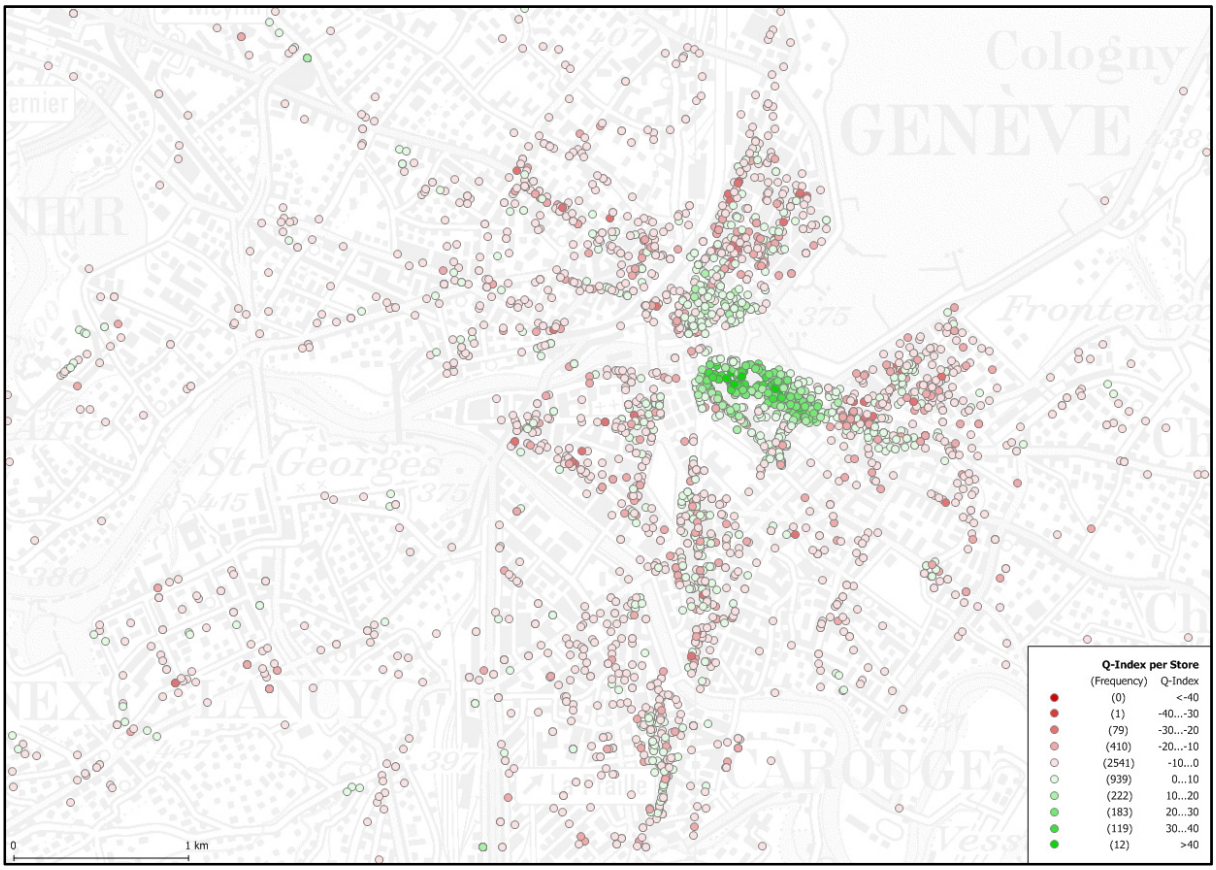


Figure 13: Map of colored stores after 'Q-Index per store'

[4] www.mapplus.ch

Once more, the center is clearly singled out with positive values. Around the center, there is a mix of positive and negative values, most values being near zero.

6.1.1.3.3 Hectometric and address location

The 'Q-Index per store'-values were also calculated for the 'hecto-all' set and compared to the normal ones: the correlation between these values is 97.3%.
Thus we have one more indication, that the hecto-model can also be used without much difference in the results.

6.1.2 All activities of Geneva

For the set of all activities of the canton Geneva, three subsets with 20% randomly chosen activities of the original set were created. The a-link-values and a_sym-values were calculated for all of them with a radius of 100m. The correlations between the values of the different sets are shown in Table 5. As for the retails stores, the correlations of a_sym are weaker than the ones of the a-values. Moreover, they are significantly higher. This is due to the fact that there are more activities and fewer categories, so a random subset reflects better the original set.

Table 5 a, b: Correlations between the set 'Geneva all' and its subsets

a-value	*all*	*A*	*B*	*C*
all	1	0.84	0.84	0.88
A	0.84	1	0.81	0.74
B	0.84	0.81	1	0.77
C	0.88	0.74	0.77	1

a_sym	*all*	*A*	*B*	*C*
all	1	0.37	0.46	0.60
A	0.37	1	0.43	0.25
B	0.46	0.43	1	0.39
C	0.60	0.25	0.39	1

6.1.2.1 Classification

The classification was done on all sets, with SIZEFACTOR fixed to 1, TMAX to 5 (the test runs with these sets gave a TMAX of about 2.5 for an acceptance rate of about 50%, but as there are only a few categories, SA is fast even when TMAX is a little bit too high). MINPERCENT was 1% or 0.5% and *timestep* 2, 1.5, 1.2 and 1.05. Most of the combinations were used twice, so that each set had at least 12 runs. The group of the best classification-result can be found in Appendix 9.5.
It is very difficult to find similarities among all of the groups, because the a_sym-values are rather badly correlated.

To analyze the groups' spatial characteristics, they have been represented on a 20m-grid (since the density is much higher for all activities than for only the retail stores). This time, the mode has been used to assign a group-number to each pixel. Then, exactly the same filter and simplification as for the retail stores have been applied. The filtered and simplified map for the set 'all' is shown on Figure 14. The original- and filtered-map of the set 'all' and also the results for the subsets can be found in Appendix 9.5. Yet the filtered and simplified maps of the subset-groups do not contain much information, because there is always one group who dominates and covers nearly all the area. For the subsets, the mode assignment of the pixels was probably not well adapted. But it was certainly adapted for the whole set. As for this set, each group has similar numbers of activities and the result allows to identify the characteristics of each group, as can be seen on the next page.

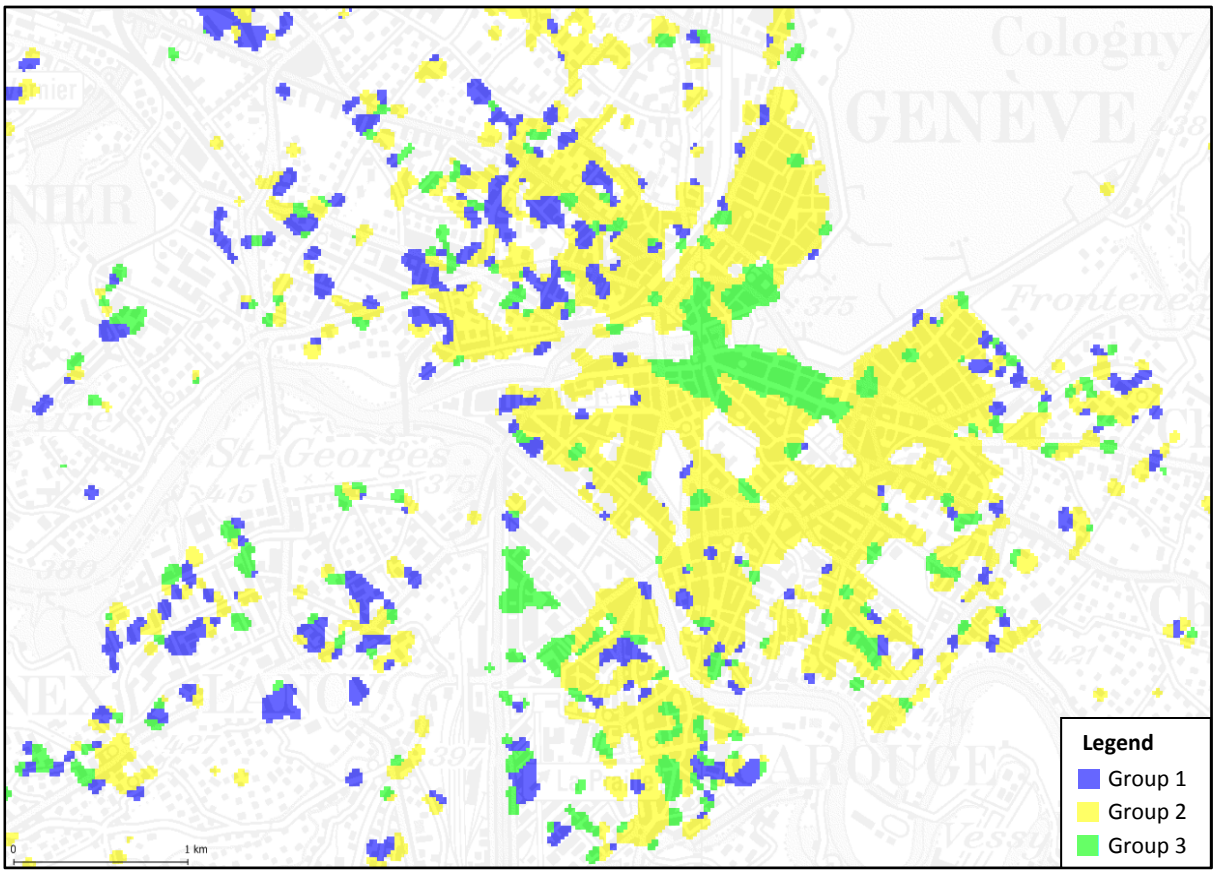


Figure 14: Groups of the set 'all ' represented on 20m-grid – filtred and simplified

The group 1 is composed of the categories ‘construction’, ‘administrative and support services’, ‘transportation and storage’ and ‘information and communication’. They are located mainly outside the city (center), sometimes clustered on a medium spot. For them, location does not matter much.
The group 2 (17903 activities) comprises, among others, the large of category ‘professional, scientific and technical activities’ (6105 activities), ‘human health and social work’, ‘accommodation and food services’, ‘education’ and ‘public administration’. This group dominates all the central part of the city except the core of the center.
Group 3 (14622 activities) has the largest category ‘wholesale and retail trade’ (8371 activities). There are furthermore ‘financial and insurance activities’ and ‘real estate activities’. They dominate the core of the center and are also present on a lot o of little and medium spots on the whole area. I cannot explain why ‘agriculture’, ‘water supply’ and ‘manufacturing’ are also in the same group (moreover even with a high satisfaction). In fact, the number of groups was limited to three by the ‘classification by strongest links’ in the beginning and thus I think that a compromise had to be found (by the algorithm).
It also has to be noted that wholesale does not belong to the center, yet retail trade does. They are nevertheless in the same category. It is therefore difficult to find common patterns. In fact, the little number of categories used here has the advantages that it makes it faster to calculate and easier to keep the overview. However on the other hand, the categories may be heterogeneous and thus neither can strong characteristics be found nor owes one subcategory dominate the others.
The findings of above allows nevertheless the analysis of the city structure, which is one of my principal objectives.

6.1.2.2 Location quality index

The Q-Index has been calculated on the same grid as for the retail-stores (50m spaced, point only calculated if not farther away than 100m from its nearest activity). The radius used was also 100m. In fact, a lesser radius could be used, since this set with all activities is denser than the one with the retail stores. Yet this should have been done from the beginning and would also necessitate a finer grid (to represent the local characteristics), which rises the calculation time considerably. It took already about 1 day for the Q-Index on this dataset, which has only few categories (18) to calculate.

As before, the values are always normalized and displayed with the help of the 'QIndex-viewer'. A selection of the 18 maps are shown and analyzed below in the following pages.

There is no surprise while analyzing the Q-Index-maps: Most of the categories like the center but do not really dislike the surroundings as the 'information and communication'-activities (Figure 15).
An interesting observation is the oppositeness of construction-activities (Figure 16) and 'public administration and defence' (Figure 17). The only category which is really center-fixed (dislikes the surroundings) is 'financial and insurance activities'. Furthermore, the embassies ('extraterritorial organizations') have their own pattern, since they mainly congregate in diverse spots (the corresponding maps are in Appendix 9.6).

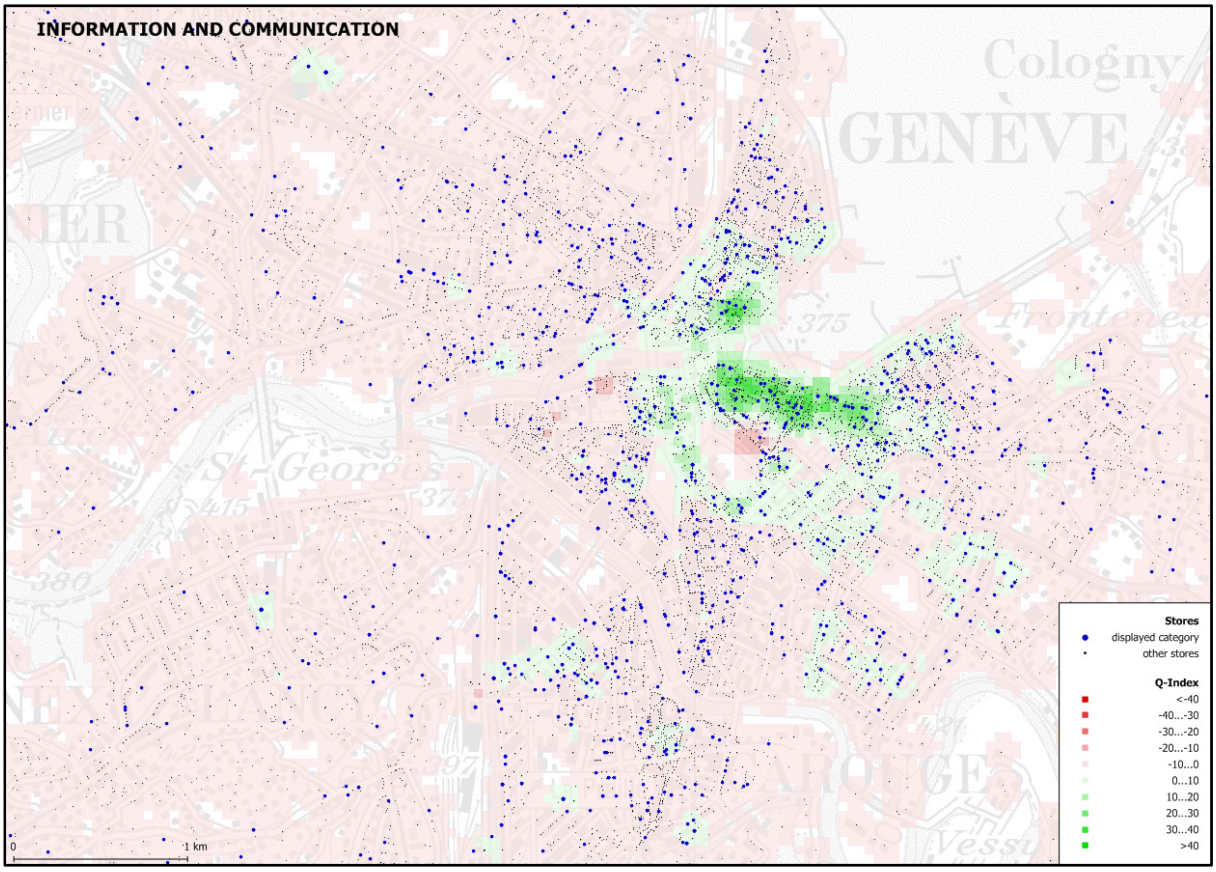


Figure 15: Q-Index-map for 'information and communication'-activities

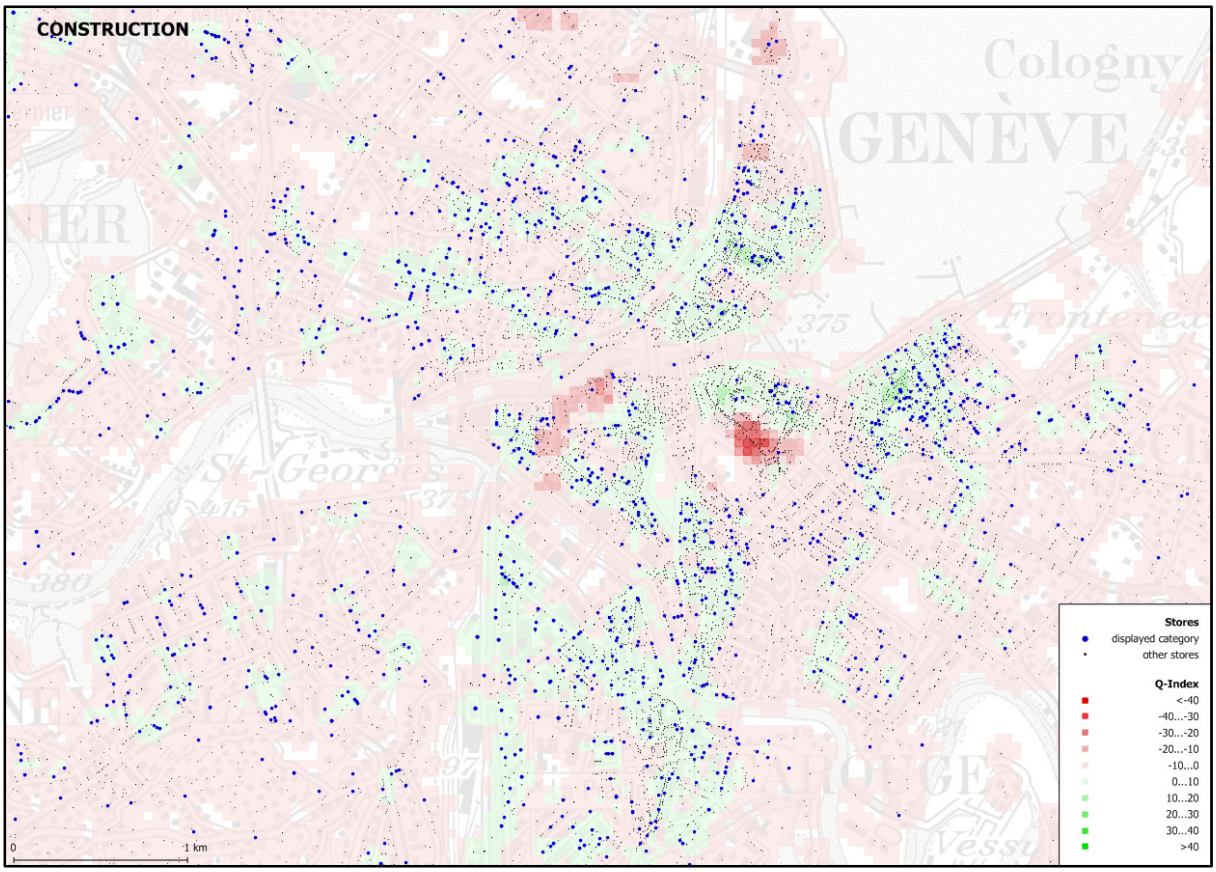


Figure 16: Q-Index-map for construction activities

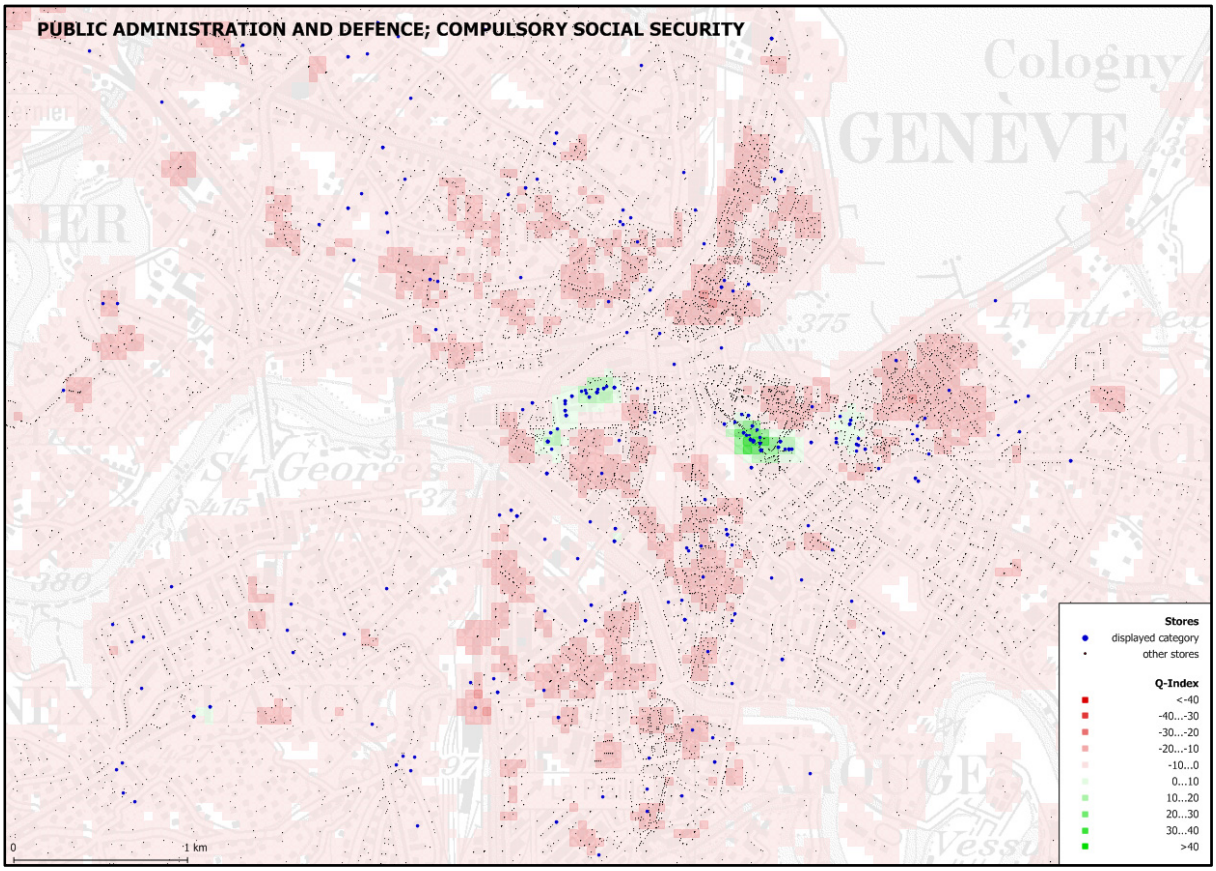


Figure 17: Q-Index-map for public administration and defence

6.1.3 Discussion Geneva

The results from both activity-sets analyzed in this chapter allowed to describe and represent diverse location patterns of categories and groups of commercial activities. A city-center was clearly identified with the Q-Index value for most of the categories, but also with the spatial representation of the classification results.

This domination of the center with a high density of activity raises the question if the algorithm should be adapted to this. A possibility would be the use of the median instead of the mean or an adaptive radius.

The use of subsets is questionable as the differences found between them are significant, particularly for the little set of the retail stores. In fact, subsets should only be used for large datasets, and even in this case with care.

The fact that the results from the hectometer-precision-data and the address-location-data are similar allows to state that these data can also be used as the basis for the calculations. Only the disadvantage that the spatial representation of the groups is limited to a 100m grid remains.

The definition of the categories in the beginning is critical. Indeed too many categories are difficult to analyze, yet if one does not chose enough of them, some characteristics may not be taken into account. From this point of view, the first set with the retail stores in 45 categories was more adapted for the analysis than all activities in only 18 categories. In fact, it was sometimes difficult to identify patterns, as some categories were too broad. For this reason, more categories were used for the case study of Barcelona.

In the same context, there arises also the problem of categories with only very little activities: is it better to integrate them into others, group them in higher level categories or just delete them. The first possibility is not recommended as the other category, in which the small category would be integrated, would be influenced. The second brings the risk of heaving too broad categories and the third is questionable because of its potential influence on the final result (although I think this influence remains modest for small categories). For the second case study presented in the following chapter, a mix of the second (when reasonable) and the third possibility has been applied (if the second was not reasonable).

6.2 Barcelona

Barcelona is the second largest city in Spain, with a population of 1'615'908 in 2008. It is located on the Mediterranean coast between the mouths of two rivers and is bounded to the west by the Serra de Collserola ridge.[5]

The database of Barcelona dates from 2002 and has 166311 activities and they were grouped in 48 categories based on the second-level of the 'Statistical classification of economic activities in the European Community' (NACE Rev.1.1). Some little categories were regrouped if possible or otherwise deleted, to have at least 50 activities in each category. So finally, the calculation-database consisted of 166246 activities. A table with the categories used can be found in Appendix 9.1.3.

6.2.1.1 Classification

The link-values (a, a_sym) for Barcelona were computed for the whole set of Barcelona with a radius of 100m. Then, the classification has been made with the same parameters than for Geneva (SIZEFACTOR = 1; MINPERCENT 1% or 0.5%; timestep 2, 1.5, 1.2, 1.05) with the exception of TMAX, which was changed to 10, so that one has about 50% of accepted tries in the beginning of the SA.

The groups for the best classification are presented in a table in Appendix 9.7.1. On the first sight, all the 'manufacture'-activities in group 3 stand out. In fact, all but two categories of the NACE-section D 'Manufacturing' are in this group. Furthermore, the 'Mining and quarrying'-category and the category of the activities for motor vehicles are part of the group 3.
Then, there is a group (number 4) with only one category: 'Hotels and restaurants'.
Another small group (in terms of number of categories) – group number 1 – comprises the largest category 'Retail trade', 'Other service activities', 'Land transport' and 'Manufacture of office machinery and computers'. All but the last category have high satisfactions (more than 90%).
Finally, there is the large group number 2 (in terms of categories and activities) with the 23 remaining categories such as financial-activities, public administration, real estate, 'Health and social work', education and also wholesale trade (but with a satisfaction of only 59%).
I think that this classification result can be qualified as satisfactory, as it represents meaningful groups of categories belonging together (with exception of the group 1).

Looking at the spatial distribution of the groups, one remarks that they are also represented on a 20m-grid. As for 'Geneva activities', the mode of the groups has been assigned to the cells and the filter and simplification were applied to get the result showed on Figure 18 (the original-grid and intermediate result are in Appendix 9.7.2 and 9.7.3). The road-network has been superposed to facilitate orientation.
On the first sight, only two groups are visible: Group 2 occupying a sort of center and group 1 dominating in the northern and southern parts. Group 3 is found on some spots, mainly in the eastern part. Only some little spots show the one-category group 'hotels and restaurants', mainly in the south-east.
It is thus possible to define a large center-area of public-, financial-, real-estate- (and many more) activities with the group 2. Then some manufacturing areas in the east can be found with the group 3. Furthermore, hotels and restaurants seems to be gathered mainly in the south-east and all the

[5] http://en.wikipedia.org/wiki/Barcelona [16.6.2009]

large remaining space is dominated by the retail trade, land transport (there is a lot of taxi-transport) or ‘other service activities’.

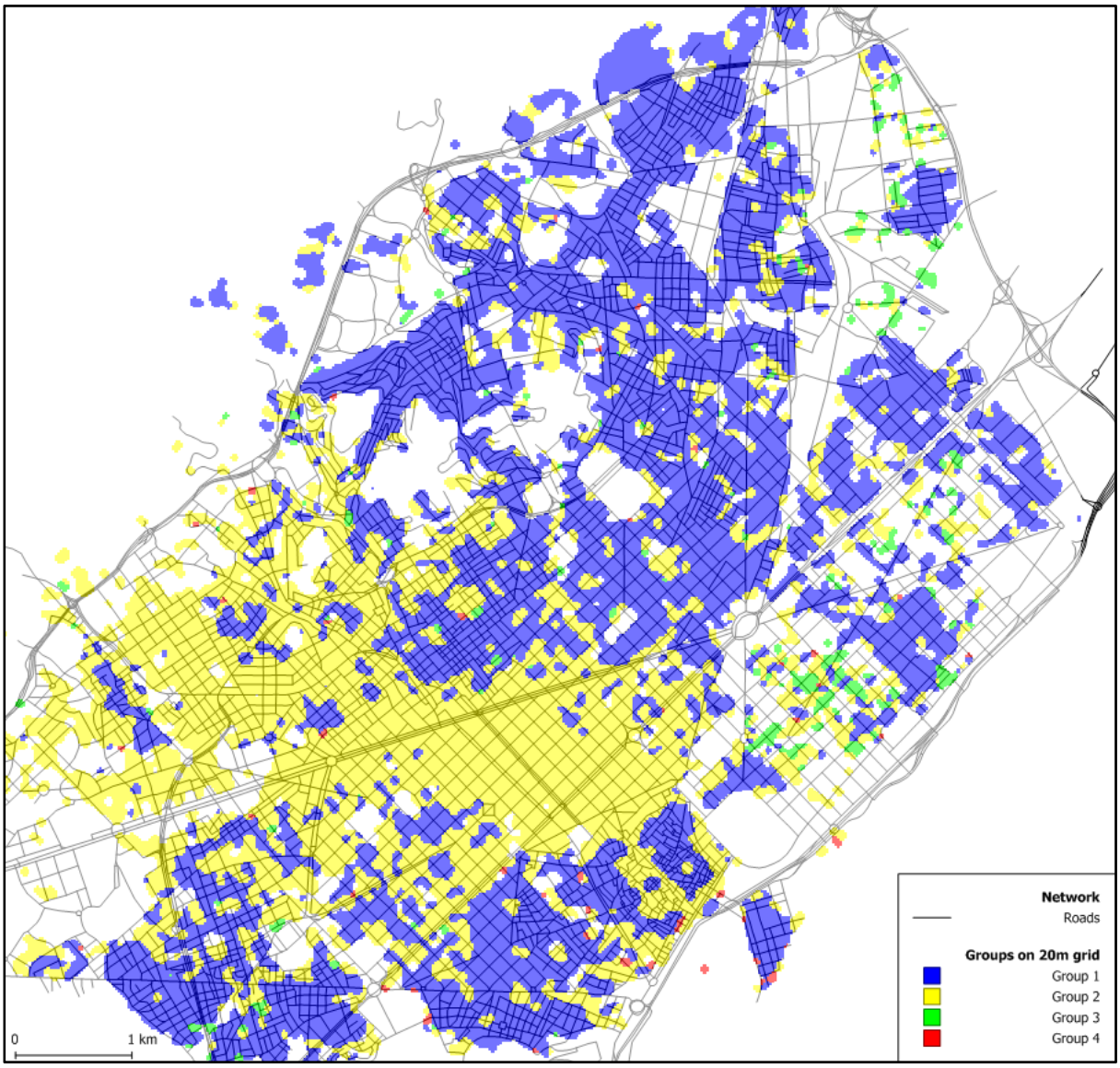


Figure 18: Groups of the set 'Barcelona-all' – filtered and simplified

6.2.1.2 Location quality index

The location quality index for all categories has been calculated on a 200m-grid and with a radius of 100m for the whole area of Barcelona. As the grid spacing is larger than the radius, the result is only a sampling and local particularities are not captured by this grid. In the center, a 10m-grid has been used to represent such local effects. The maps showed in the following analysis show both grids. This can be seen as there is a detailed structure in the center but large 'pixels' around.
As for the group-representation, the superposed road-network is used to orientate.

6.2.1.2.1 Analysis

A lot of categories like the center-area defined above form the classification. One example is the category of real-estate-activities, for which the Q-Index-map is shown on Figure 19. Here we can also see the center-core, which seems to be just above the main road. Furthermore, there are some roads going south-east from the main roads, which have also very high Q-Index-values.

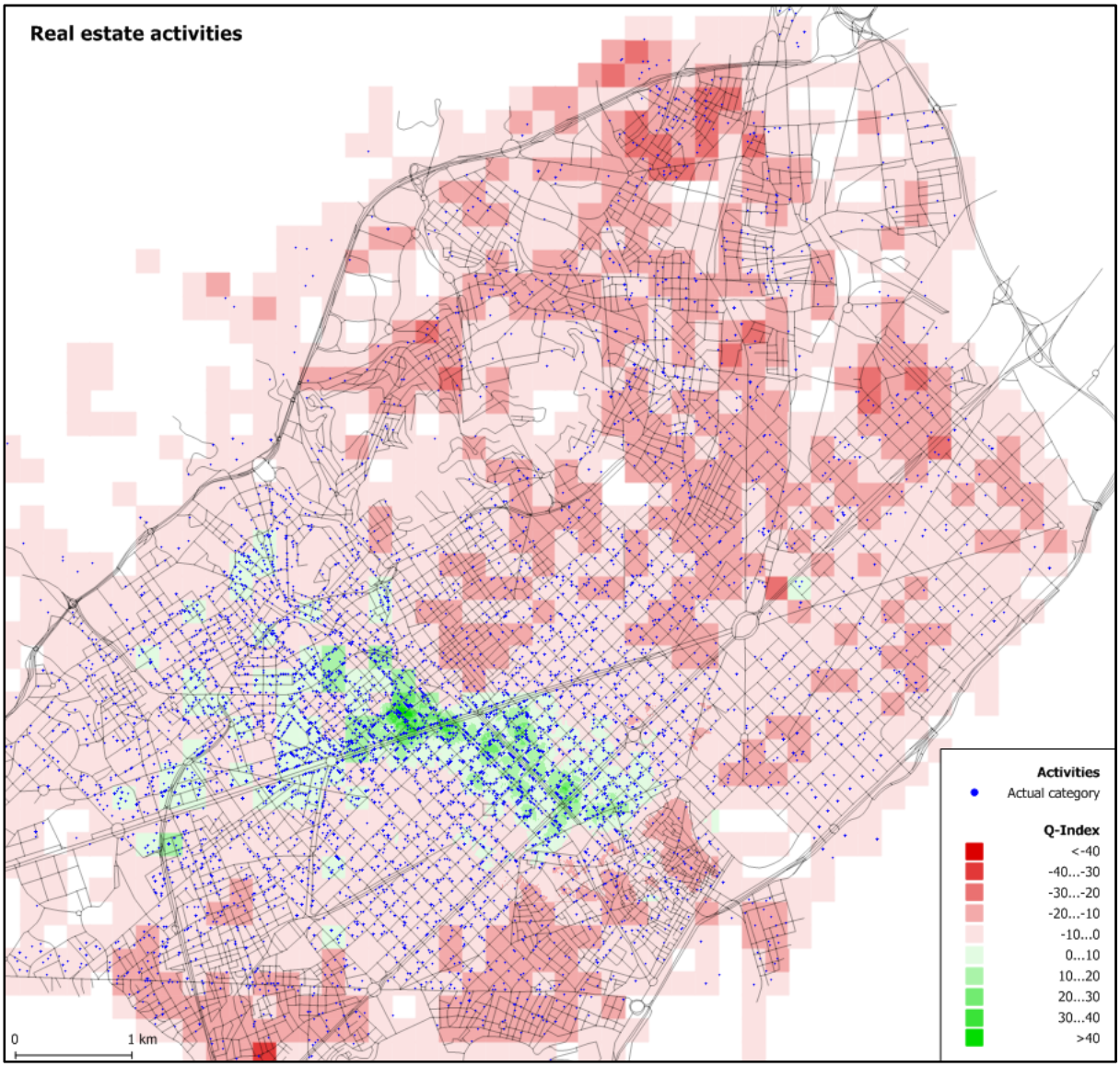


Figure 19: Q-Index-map for real estate activities

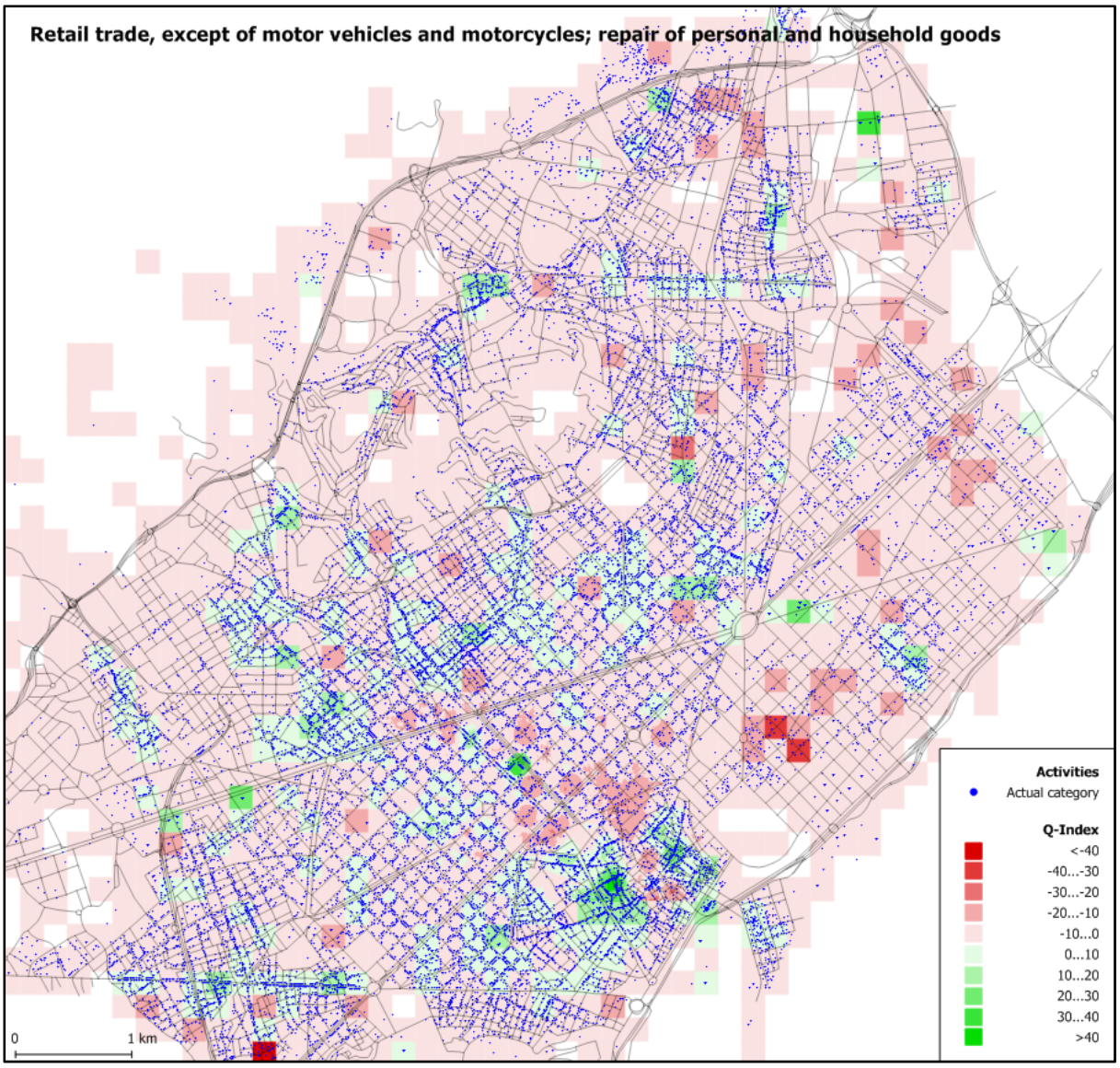


Figure 20: Q-Index-map for the retail trade

On Figure 20, we can see that the retail trade is spread nearly everywhere in the city, but mostly concentrated in the south-east (where I suspect the old-town is located). There are also a lot of other little areas where the Q-Index is good for the retail stores. One of these areas is located just in the west of the center-core defined above. There are two spots with very low values. They seem to be a manufacturing center, as they are the preferred locations for some manufacturing activities as 'Manufacture of rubber and plastic products' (map in Appendix 9.8).

It is not surprising that the hotels and restaurants prefer the center-area, but even more the old-town (of which the location was confirmed by the look on a satellite image[6]). There are also some other spots with high Q-Index-values for this category, for which the map is also in Appendix 9.8.

A lot of the manufacturing categories dislike the center-core and/or the 'old-town', but are found spread all over the rest of the area as for example furniture-manufacturing (map in Appendix 9.8).

[6]http://maps.google.com/maps?hl=en&ll=41.379707,2.180958&t=h&z=14 [16.6.09]

6.2.1.2.2 Comparison with centrality indices

Comparisons of the Q-Index-values with the betweenness-centrality of the road-network have been done. There are also global straightness- and closeness- and some local closeness-centralities available, but they were not used because they show rather 'basic' patterns (e.g. areas where the network is dense or the most straight). The only interesting finding with these centralities is that the center of the global closeness-centrality is very close to the center-core found above (a map with this comparison is in Appendix 9.9.1).

But the approximation of the betweenness-centrality is not easy, since it has very special patterns. The Q-Index of all categories and also some combinations have been tried out, without success. The category with the most similarities is represented on Figure 21. It is a logical result, as activities for motor vehicles have to be accessible by roads, and the betweenness is high mostly for important roads.

Nevertheless, this category also does not satisfactorily approach the betweenness centrality, so this aim of the work could not be reached.

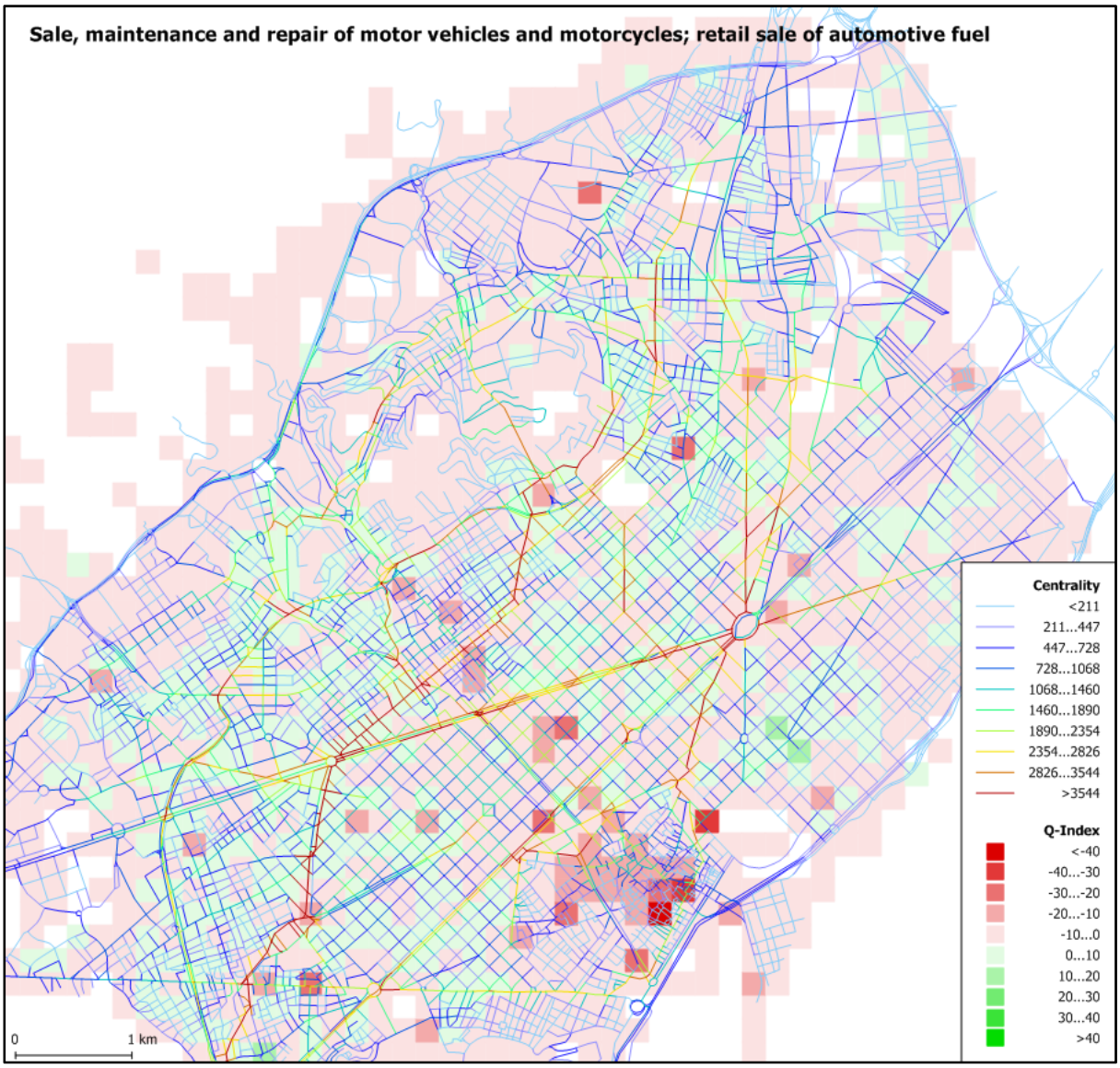


Figure 21: Map with betweenness-centrality and Q-Index of motor vehicle-activities

6.2.2 Discussion Barcelona

For this city, a good classification could be found. In fact, the algorithm grouped 18 of 20 categories of the NACE-section 'Manufacturing' together in one group. The graphical representation of the groups allowed to identify areas with different activity-characteristics. Furthermore, diverse centers and preference areas could be found with the help of the Q-Index-maps.

Besides the coincidence of the closeness-centrality-center and center-core identified with the Q-Index, the comparison with the centrality-indices gave no valuable result.

But as the Q-Index values represent also a sort of 'centrality' (defined by the location of the activities), the question arises if an 'activity-centrality' based on the Q-Index-values is not a better means to understand and analyze the city than by approximating road-network-centralities. I think when both information are used complementary, it is the best solution: the first indicating the location quality, the second the accessibility (in a broad sense). At the moment, the Q-Index is used per category, but I think an extension to groups (of categories) could be found or possibly even a global integrated one.

6.3 Comparison Geneva - Barcelona

In this section, a comparison of the results found for both case studies is done.

Yet unfortunately, no deep comparison is possible, for two reasons: first, because the older NACE Rev. 1.1 classification of the Barcelona-dataset does not match the NOGA-2008 classification (which is compatible with the newer NACE Rev. 2) of the Geneva-dataset. The second reason is that different levels of the categories were used. In fact, I chose the second-level categories for Barcelona because the results with the top-level categories for Geneva raised some some problems with too broad categories, as explained before.

Nevertheless, some similarities and differences can be presented: in both cities, centers were identified, one in Geneva and two in Barcelona. The center in the old town of Barcelona (where most of the retail stores and hotels and restaurants like to be located) can be compared to the one in Geneva (which the two categories noted above prefer also). The center-area in Barcelona can be compared to the surrounding of the center in Geneva. In fact, there are several parallels between the groups 2 of the classifications for both cities, as they comprise both categories like 'health and social work', 'public administration' and 'education'. The financial- and real estate-activities are more likely to be located in the 'old-town-center' in Geneva, but in the center-area or even more in the center-core (so the other center) of Barcelona. In both cities, hotels and restaurants (called 'accommodation and food services' in Geneva) really like the 'old-town-center' but are also likely to be located in the surroundings in Geneva and the center-area in Barcelona. In this last town, a manufacturing area was found, but not in Geneva (probably because it is outside the analyzed area). And finally, the retail stores like clearly the old-town-center in both cities, but in Barcelona different other preference-areas were found, which has certainly to do with the different size of the cities.

7 Conclusion and perspectives

In the first part of the work, the program was optimized. The adoptions for the limiting M-Index-calculation resulted in a time gain of a factor 10 for this part of the code.
The new simulated annealing implementation after Johnson et al. (1989) allowed mostly to obtain better and certainly more stable results in the same amount of time or even less than before.

In the case studies, the program was applied and the results analyzed.
The spatial representation of the classified groups allowed to recognize the dominance areas of the different groups of economic activities. But the initial definition of the categories seems to be very important to get clear results. In fact, a group matching nearly the official NACE-classification was found for the set of Barcelona with 48 categories, but the ones for the set 'Geneva activities' with only 18 (to broad) categories were difficult to interpret.
The location quality index allowed to identify diverse preference areas of the categories. Construction activities for example have nearly opposite preference-pattern than public administration in Geneva. The result of the validation indicates that this index is meaningful.
It can therefore be concluded that the applied methods are adapted to analyze the commercial structure of cities and the location quality index seems to be a useful indicator for location planning.

Tests with different randomly selected subsets resulted in considerable differences between their link values and the classification results. Thus the use of such subsets should only be done with care.
The link values, the classification results and the location quality indices obtained with hectometric data were very similar to the ones from the original address-based data. This shows that the algorithm used here could be applied to whole Switzerland with the data of the Swiss federal census of enterprises. Possible future studies could be done to determine the overall structure of commercial activities in the country and to identify possible central locations. It is also possible to propose the program as a useful tool for cantonal or regional economic development agencies to determine optimal locations for new coming large foreign industrial commercial groups.

The comparisons of the location quality indices with the betweenness-centrality gave no valuable result. But I think that the Q-Index should be used as an independent index complementary to the centrality. The Q-Index, being based on activities, takes their spatial distribution into account, while centrality indices based on the road network characterize accessibility (in a broad sense).

As a perspective, it would be particularly interesting to work on a definition of a Q-Index for groups of categories or even of a global integrated one: it would then be probably possible to use this index as a surrogate for the evaluation of value landed property in particular.

8 Bibliography

Aerts, J.C.J.H. and Heuvelink, G.B.M. 2002. Using simulated annealing for resource allocation. International journal of geographical information science 16 (6): 571-587.

Antunes, A. and Peeters, D. 2001. On solving complex multi-period location models using simulated annealing. European journal of operational research 130 (1): 190-201.

Barthélemy, M. et al. 2005. Characterization and modeling of weighted networks. Physica A 346 (1-2): 34-43.

Brandes, U. 2008. On variants of shortest-path betweenness centrality and their generic computation. Social Networks 30 (2): 136-145.

Crucitti, P., Latora, V. and Porta, S. 2006. Centrality measures in spatial networks of urban streets. Physical review. E, Statistical, nonlinear, and soft matter physics 73 (3): 036125.

Cunha, M.C. et al. 2009. Optimization Model for Integrated Regional Wastewater Systems Planning. Journal of Water Resources Planning and Management Volume 135, Issue 1: 23-33

Freeman, L. 1979. CENTRALITY IN SOCIAL NETWORKS CONCEPTUAL CLARIFICATION. Social Networks. 1 (3): 215-239.

Gastner, M. and Newman, M. 2004. Diffusion-based method for producing density-equalizing maps. Proceedings of the National Academy of Sciences of the United States of America 101 (20): 7499-7504.

Henderson, D. et al. 2003. The theory and practice of simulated annealing. Handbook of metaheuristics. 57 : 287-319.

Jensen, P. 2006. Network-based predictions of retail store commercial categories and optimal locations. Physical review. E, Statistical, nonlinear, and soft matter physics 74 (3): 035101.

Jensen, P. et al. 2005. Aggregation of retail stores. Physica A - Statistical mechanics and its applications. 351 (2-4): 551-570.

Johnson, D. et al. 1989. Optimization by simulated annealing - an experimental evaluation; part 1, graph partitioning. Operations research 37 (6): 865-892.

Kirkpatrick, S. 1983. Optimization by simulated annealing. Science. 220 (4598): 671-680.

Latora, V. and Marchiori, M. 2007. A measure of centrality based on network efficiency. New journal of physics 9 (6): 188.

Manifold® System Release 8 User Manual, Manifold Net Ltd., Carson City, USA. www.manifold.net/doc

Marcon, E. and Puech, F. (2008) Measures of the geographic concentration of industries: improving distance-based methods.
http://halshs.archives-ouvertes.fr/docs/00/37/26/17/PDF/HAL_Marcon_Puech_-_MeasuresOfTheGeographicConcentrationOfIndustriesImprovingDistanceBasedMethods.pdf [4.6.09]

Newman, M.E.J. and Girvan, M. 2004. Finding and evaluating community structure in networks. Physical review. E, Statistical, nonlinear, and soft matter physics 69 (2): 026113.

Produit, Timothée. 2009. A novel GIS method to determine an urban centrality index applied to the Barcelona metropolitan area. Master thesis, LASIG, EPFL.

Radicchi, F. et al. 2004. Defining and identifying communities in networks. Proceedings of the National Academy of Sciences of the United States of America. 101 (9): 2658-2663.

Rattigan, M. J. et al. 2006. Using structure indices for efficient approximation of network properties. In Proceedings of the 12th ACM SIGKDD international Conference on Knowledge Discovery and Data Mining (Philadelphia, PA, USA, August 20 - 23, 2006). KDD '06. ACM, New York, NY, 357-366.

Reichardt, J. and Bornholdt, S. 2004. Detecting fuzzy community structures in complex networks with a Potts model. Physical review letters. 93 (21): 218701.

Sante-Riveira, I. et al. 2008. Algorithm based on simulated annealing for land-use allocation. Computers & geosciences 34 (3): 259-268.

9 Appendix

9.1 Categories of the 'activity-sets'

9.1.1 Categories of the set 'Geneva retail stores'

Category	NOGA Categories	Store Count	Category Name
4711	4711	191	Retail sale in non-specialised stores with food, beverages or tobacco predominating
4719	4719	18	Other retail sale in non-specialised stores
472200	472200	68	Retail sale of meat and meat products in specialised stores
472401	472401	156	Retail sale of bread, cakes, flour confectionery and sugar confectionery in specialised stores
472402	472402	131	Bakeries - tea rooms
472500	472500	48	Retail sale of beverages in specialised stores
4729	4729, 472100, 472300, 472600	210	Other retail sale of food in specialised stores
473000	473000	79	Retail sale of automotive fuel in specialised stores
474100	474100	105	Retail sale of computers, peripheral units and software in specialised stores
474200	474200	35	Retail sale of telecommunications equipment in specialised stores
474300	474300	76	Retail sale of audio and video equipment in specialised stores
475100	475100	49	Retail sale of textiles in specialised stores
4752	4752	60	Retail sale of hardware, paints and glass in specialised stores
475300	475300	17	Retail sale of carpets, rugs, wall and floor coverings in specialised stores
475400	475400	62	Retail sale of electrical household appliances in specialised stores
475901	475901	24	Retail sale of musical instruments
475902	475902	79	Retail sale of furniture
475903	475903	64	Retail sale of household utensils and fittings n.e.c.
476100	476100	76	Retail sale of books in specialised stores
476201	476201	267	Retail sale of newspapers and magazines, newspaper stands
476202	476202	26	Retail sale of writing materials and stationery
476300	476300	21	Retail sale of music and video recordings in specialised stores
4764	4764	82	Retail sale of sporting equipment in specialised stores
476500	476500	34	Retail sale of games and toys in specialised stores
477101	477101	239	Retail sale of women's clothing
477102	477102	50	Retail sale of men's clothing
477103	477103, 477104	45	Retail sale of babies' and children's clothing
477105	477105	268	Retail sale of clothing accessories and clothing equally combined
477201	477201	103	Retail sale of footwear
477202	477202	25	Retail sale of leather and travel goods
477300	477300	176	Dispensing chemist in specialised stores
477400	477400	22	Retail sale of medical and orthopaedic goods in specialised stores
4775	4775	49	Retail sale of cosmetic and toilet articles in specialised stores
4776	4776	99	Retail sale of flowers, plants, seeds, fertilisers, pet animals and pet food in specialised stores
477700	477700	245	Retail sale of watches and jewellery in specialised stores
477802	477802	90	Retail sale of glasses and other optical goods
477804	477804	60	Retail sale of gifts and souvenirs
477805	477805	118	Art trade
477806	477806, 477801, 477803	170	Other retail sale of goods in specialised stores n.e.c.
477901	477901	66	Retail sale of antiques
477902	477902	81	Retail sale of second-hand goods in stores
478100	478100	105	Retail sale via stalls and markets of food, beverages and tobacco products
478900	478900, 478200	272	Retail sale via stalls and markets of other goods
479100	479100	78	Retail sale via mail order houses or via Internet
479900	479900	167	Other retail sale not in stores, stalls or markets

9.1.2 Categories of the set 'Geneva activities'

Category	NOGA section	StoreCount	CategoryName
A0	A	427	AGRICULTURE, FORESTRY AND FISHING
C0	C	1919	MANUFACTURING
E0	E	134	WATER SUPPLY; SEWERAGE, WASTE MANAGEMENT AND REMEDIATION ACTIVITIES
F0	F	2614	CONSTRUCTION
G0	G	8371	WHOLESALE AND RETAIL TRADE; REPAIR OF MOTOR VEHICLES AND MOTORCYCLES
H0	H	1640	TRANSPORTATION AND STORAGE
I0	I	2545	ACCOMMODATION AND FOOD SERVICE ACTIVITIES
J0	J	1800	INFORMATION AND COMMUNICATION
K0	K	2745	FINANCIAL AND INSURANCE ACTIVITIES
L0	L	1026	REAL ESTATE ACTIVITIES
M0	M	6105	PROFESSIONAL, SCIENTIFIC AND TECHNICAL ACTIVITIES
N0	N	1751	ADMINISTRATIVE AND SUPPORT SERVICE ACTIVITIES
O0	O	536	PUBLIC ADMINISTRATION AND DEFENCE; COMPULSORY SOCIAL SECURITY
P0	P	1273	EDUCATION
Q0	Q	3108	HUMAN HEALTH AND SOCIAL WORK ACTIVITIES
R0	R	853	ARTS, ENTERTAINMENT AND RECREATION
S0	S	3139	OTHER SERVICE ACTIVITIES
U0	U	344	ACTIVITIES OF EXTRATERRITORIAL ORGANISATIONS AND BODIES

9.1.3 Categories of the set 'Barcelona'

Category	NACE	StoreCount	CategoryName
CC	C	86	Mining and quarrying
D15	15	479	Manufacture of food products and beverages
D17	17	275	Manufacture of textiles
D18	18	1640	Manufacture of wearing apparel; dressing and dyeing of fur
D19	19	103	Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear
D20	20	451	Manufacture of wood and of products of wood and cork, except furniture
D21	21	163	Manufacture of pulp, paper and paper products
D22	22	3802	Publishing, printing and reproduction of recorded media
D24	24	344	Manufacture of chemicals and chemical products
D25	25	308	Manufacture of rubber and plastic products
D26	26	160	Manufacture of other non-metallic mineral products
D27	27	58	Manufacture of basic metals
D28	28	1390	Manufacture of fabricated metal products, except machinery and equipment
D29	29	310	Manufacture of machinery and equipment n.e.c.
D30	30	161	Manufacture of office machinery and computers
D31	31	246	Manufacture of electrical machinery and apparatus n.e.c.
D32	32	138	Manufacture of radio, television and communication equipment and apparatus
D33	33	483	Manufacture of medical, precision and optical instruments, watches and clocks
D34	34	80	Manufacture of motor vehicles, trailers and semi-trailers
D35	35	79	Manufacture of other transport equipment
D36	36	969	Manufacture of furniture; manufacturing n.e.c.
EE	E	58	Electricity, gas and water supply
G50	50	3375	Sale, maintenance and repair of motor vehicles and motorcycles; retail sale of automotive fuel
G51	51	12723	Wholesale trade and commission trade, except of motor vehicles and motorcycles
G52	52	36310	Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods
H55	55	12758	Hotels and restaurants
I60	60	16097	Land transport; transport via pipelines
I61	61	61	Water transport
I62	62	72	Air transport
I63	63	2961	Supporting and auxiliary transport activities; activities of travel agencies
I64	64	866	Post and telecommunications
J65	65	4598	Financial intermediation, except insurance and pension funding
J66	66	657	Insurance and pension funding, except compulsory social security
J67	67	563	Activities auxiliary to financial intermediation
K70	70	10343	Real estate activities
K71	71	1110	Renting of machinery and equipment without operator and of personal and household goods
K72	72	138	Computer and related activities
K73	73	202	Research and development
K74	74	17189	Other business activities
L75	75	2966	Public administration and defence; compulsory social security
M80	80	4655	Education
N85	85	4727	Health and social work
O90	90	515	Sewage and refuse disposal, sanitation and similar activities
O91	91	5721	Activities of membership organizations n.e.c.
O92	92	9162	Recreational, cultural and sporting activities
O93	93	6492	Other service activities
P95	95	115	Activities of households as employers of domestic staff
Q99	99	87	Extra-territorial organizations and bodies
Deleted:			
X	A	5	Agriculture, hunting and forestry
X	16	3	Manufacture of tobacco products
X	23	12	Manufacture of coke, refined petroleum products and nuclear fuel
X	37	45	Recycling

9.2 Tables and graphics of the sensibility analysis 'Geneva retail stores'

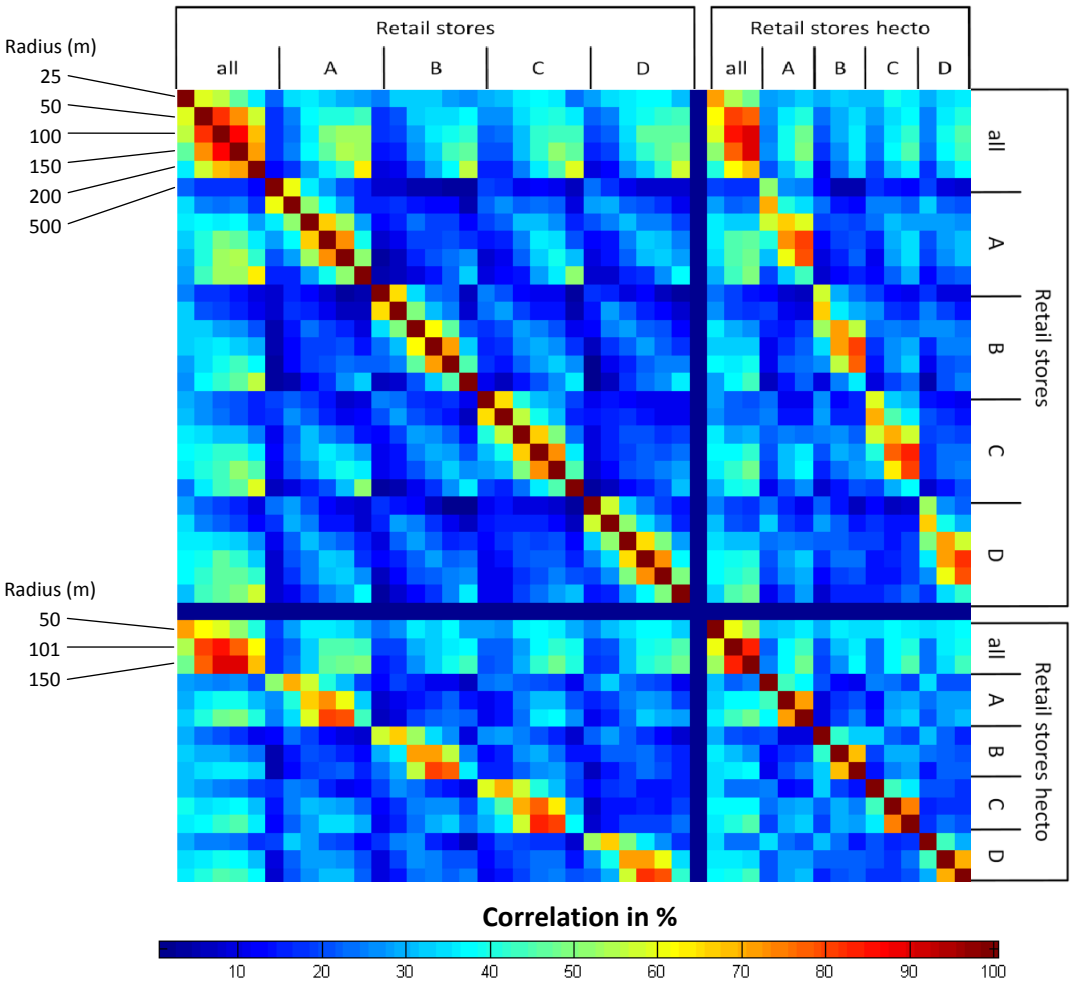
9.2.1 Descriptive statistics of the a-matrixes

activity set	Retail stores all						Retail stores 33% A						Retail stores 33% B						Retail stores 33% C						Retail stores 33% D					
radius (m)	25	50	100	150	200	500	25	50	100	150	200	500	25	50	100	150	200	500	25	50	100	150	200	500	25	50	100	150	200	500
# zero links	265	100	3	3	2	0	1147	759	302	152	92	6	1126	815	384	197	106	13	1140	850	492	275	176	20	1114	733	322	168	102	10
# non zero links	1760	1925	2022	2022	2023	2025	878	1266	1723	1873	1933	2019	899	1210	1641	1828	1919	2012	885	1175	1533	1750	1849	2005	911	1292	1703	1857	1923	2015
# positive links	848	887	959	951	958	1010	643	765	903	929	946	1039	634	732	833	857	878	884	592	716	812	865	890	979	675	815	911	915	917	969
# negative links	912	1038	1063	1071	1065	1015	235	501	820	944	987	980	265	478	808	971	1041	1128	293	459	721	885	959	1026	236	477	792	942	1006	1046
# strong links (> 0.5)	445	370	258	228	190	117	450	464	430	365	342	226	436	451	389	349	308	166	412	433	418	375	346	167	459	480	399	340	302	192
# strong anti-links (< 0.5)	485	521	495	437	400	307	100	228	429	460	467	370	125	234	416	516	513	435	140	233	369	435	454	399	99	251	384	443	457	365
# strong anti- and links	930	891	753	665	590	424	550	692	859	825	809	596	561	685	805	865	821	601	552	666	787	810	800	566	558	731	783	783	759	557
Maximum link	3.24	2.68	1.66	1.53	1.10	0.98	3.98	2.83	1.81	1.67	1.54	1.08	3.69	3.42	2.38	2.01	2.05	1.02	3.82	3.93	3.53	2.43	2.04	1.40	5.32	4.31	3.09	2.53	2.17	1.08
Minimum link	-3.62	-3.25	-2.77	-2.71	-2.67	-2.29	-2.73	-3.32	-3.07	-2.79	-3.48	-2.81	-2.18	-3.57	-2.93	-2.98	-3.41	-3.17	-2.52	-2.49	-3.14	-3.75	-4.25	-3.15	-3.24	-2.72	-3.16	-2.75	-3.07	-2.41
Overall mean	-0.09	-0.14	-0.14	-0.12	-0.11	-0.08	0.22	0.13	-0.02	-0.07	-0.10	-0.09	0.19	0.12	-0.04	-0.13	-0.15	-0.15	0.17	0.12	0.01	-0.07	-0.10	-0.11	0.22	0.12	-0.01	-0.08	-0.11	-0.10
Sum of positive links	488.3	415.8	349.4	317.4	296.3	248.3	575.1	549.6	493.5	437.5	403.4	330.3	547.0	541.2	444.0	395.7	357.4	259.4	528.7	526.1	481.2	433.1	405.1	291.2	581.2	559.8	468.6	408.4	370.4	305.1
Sum of negative links	-680.3	-693.4	-626.6	-558.5	-526.2	-413.3	-133.8	-295.4	-537.0	-588.1	-612.5	-511.9	-157.1	-305.4	-530.5	-652.5	-670.6	-565.9	-174.8	-284.7	-455.9	-570.6	-600.8	-514.7	-128.4	-311.6	-498.6	-572.5	-587.8	-512.9
Mean of postive links	0.58	0.47	0.36	0.33	0.31	0.25	0.89	0.72	0.55	0.47	0.43	0.32	0.86	0.74	0.53	0.46	0.41	0.29	0.89	0.73	0.59	0.50	0.46	0.30	0.86	0.69	0.51	0.45	0.40	0.31
Mean of negative links	-0.75	-0.67	-0.59	-0.52	-0.49	-0.41	-0.57	-0.59	-0.65	-0.62	-0.62	-0.52	-0.59	-0.64	-0.66	-0.67	-0.64	-0.50	-0.60	-0.62	-0.63	-0.64	-0.63	-0.50	-0.54	-0.65	-0.63	-0.61	-0.58	-0.49
activity set	Retail stores hecto all						Retail stores hecto 33% A						Retail stores hecto 33% B						Retail stores hecto 33% C						Retail stores hecto 33% D					
radius (m)		50	101	150				50	101	150				50	101	150				50	101	150				50	101	150		
# zero links		68	5	2				674	238	131				789	284	161				805	346	229				696	238	131		
# non zero links		1957	2020	2023				1351	1787	1894				1236	1741	1864				1220	1679	1796				1329	1787	1894		
# positive links		915	932	947				845	920	944				747	862	853				739	849	877				810	904	911		
# negative links		1042	1088	1076				506	867	950				489	879	1011				481	830	919				519	883	983		
# strong links (> 0.5)		356	241	214				522	401	357				447	365	326				440	398	350				469	368	318		
# strong anti-links (< 0.5)		525	458	422				268	452	483				236	480	520				249	390	431				264	432	468		
# strong anti- and links		881	699	636				790	853	840				683	845	846				689	788	781				733	800	786		
Maximum link		2.42	1.51	1.31				2.31	1.78	1.70				2.38	2.03	2.00				2.56	2.61	2.10				4.22	2.83	2.47		
Minimum link		-3.13	-3.28	-2.88				-3.44	-2.58	-3.03				-2.88	-2.94	-3.21				-2.45	-3.44	-3.85				-2.93	-2.82	-3.16		
Overall mean		-0.15	-0.13	-0.12				0.12	-0.05	-0.09				0.11	-0.09	-0.14				0.11	-0.03	-0.08				0.10	-0.06	-0.10		
Sum of positive links		407.0	329.1	305.2				582.3	460.2	420.0				529.9	423.4	379.8				517.7	444.6	412.8				538.6	435.5	389.0		
Sum of negative links		-710.5	-598.3	-546.8				-329.8	-561.8	-601.0				-314.6	-610.9	-660.9				-301.1	-510.0	-568.6				-336.5	-558.8	-590.6		
Mean of postive links		0.44	0.35	0.32				0.69	0.50	0.44				0.71	0.49	0.45				0.70	0.52	0.47				0.66	0.48	0.43		
Mean of negative links		-0.68	-0.55	-0.51				-0.65	-0.65	-0.63				-0.64	-0.70	-0.65				-0.63	-0.61	-0.62				-0.65	-0.63	-0.60		

9.2.2 Descriptive statistics of the a_sym-matrixes

activity set	Retail stores all						Retail stores 33% A						Retail stores 33% B						Retail stores 33% C						Retail stores 33% D					
radius (m)	25	50	100	150	200	500	25	50	100	150	200	500	25	50	100	150	200	500	25	50	100	150	200	500	25	50	100	150	200	500
# zero links	993	897	907	1011	1011	1127	1435	1247	923	873	919	953	1457	1279	1015	855	875	961	1493	1251	1019	941	897	961	1399	1175	955	921	911	991
# non zero links	1032	1128	1118	1014	1014	898	590	778	1102	1152	1106	1072	568	746	1010	1170	1150	1064	532	774	1006	1084	1128	1064	626	850	1070	1104	1114	1034
# positive links	472	476	496	438	446	442	494	512	584	562	524	562	460	490	510	520	486	408	408	508	540	524	522	504	524	584	586	532	508	478
# negative links	560	652	622	576	568	456	96	266	518	590	582	510	108	256	500	650	664	656	124	266	466	560	606	560	102	266	484	572	606	556
# strong links (> 0.5)	288	228	164	126	94	76	400	376	344	280	228	146	378	346	296	246	196	82	340	368	344	266	248	120	420	392	310	238	198	130
# strong anti-links (< 0.5)	390	420	348	284	238	112	60	160	356	350	352	252	62	174	330	438	446	282	66	166	292	360	364	270	50	166	314	340	328	238
# strong anti- and links	678	648	512	410	332	188	460	536	700	630	580	398	440	520	626	684	642	364	406	534	636	626	612	390	470	558	624	578	526	368
Maximum link	1.59	1.44	1.28	1.18	1.05	0.80	2.88	2.03	1.54	1.48	1.34	0.97	3.06	2.70	1.79	1.52	1.34	0.90	3.02	2.85	2.34	2.05	1.81	1.01	2.73	2.58	1.94	1.79	1.66	0.86
Minimum link	-2.25	-2.15	-2.13	-1.82	-1.93	-1.28	-1.88	-2.33	-2.32	-2.08	-2.66	-2.06	-1.46	-2.45	-2.36	-1.96	-2.26	-2.56	-1.61	-1.80	-2.38	-2.50	-3.06	-2.34	-1.74	-1.77	-2.16	-2.00	-2.31	-2.04
Overall mean	-0.08	-0.11	-0.09	-0.07	-0.05	-0.01	0.20	0.11	-0.02	-0.05	-0.07	-0.04	0.18	0.11	-0.03	-0.11	-0.13	-0.10	0.16	0.12	0.01	-0.06	-0.08	-0.06	0.20	0.12	0.00	-0.05	-0.08	-0.05
Sum of positive links	296.9	253.9	213.2	180.4	167.8	142.8	471.1	413.4	358.8	304.1	266.4	216.8	429.7	402.6	314.4	269.7	226.2	144.2	409.0	415.2	361.3	297.5	272.6	184.9	471.9	433.8	341.5	273.7	234.0	183.5
Sum of negative links	-466.1	-472.8	-397.1	-314.5	-279.2	-164.3	-61.7	-189.3	-390.6	-401.2	-414.1	-293.6	-66.9	-185.9	-375.8	-486.8	-482.4	-338.1	-78.5	-181.9	-331.7	-410.8	-433.6	-315.6	-57.8	-197.0	-348.5	-379.7	-388.1	-282.2
Mean of postive links	0.63	0.53	0.43	0.41	0.38	0.32	0.95	0.81	0.61	0.54	0.51	0.39	0.93	0.82	0.62	0.52	0.47	0.35	1.00	0.82	0.67	0.57	0.52	0.37	0.90	0.74	0.58	0.51	0.46	0.38
Mean of negative links	-0.83	-0.73	-0.64	-0.55	-0.49	-0.36	-0.64	-0.71	-0.75	-0.68	-0.71	-0.58	-0.62	-0.73	-0.75	-0.75	-0.73	-0.52	-0.63	-0.68	-0.71	-0.73	-0.72	-0.56	-0.57	-0.74	-0.72	-0.66	-0.64	-0.51
activity set	Retail stores hecto all						Retail stores hecto 33% A						Retail stores hecto 33% B						Retail stores hecto 33% C						Retail stores hecto 33% D					
radius (m)		50	101	150				50	101	150				50	101	150				50	101	150				50	101	150		
# zero links		861	951	1021				1159	895	871				1253	985	913				1223	979	927				1121	935	909		
# non zero links		1164	1074	1004				866	1130	1154				772	1040	1112				802	1046	1098				904	1090	1116		
# positive links		506	448	430				592	582	564				504	500	468				520	524	522				588	548	516		
# negative links		658	626	574				274	548	590				268	540	644				282	522	576				316	542	600		
# strong links (> 0.5)		218	136	120				424	312	244				340	274	210				376	298	262				400	270	228		
# strong anti-links (< 0.5)		442	332	262				168	368	360				192	372	438				180	302	334				202	350	352		
# strong anti- and links		660	468	382				592	680	604				532	646	648				556	600	596				602	620	580		
Maximum link		1.29	1.09	1.09				2.08	1.62	1.51				2.33	1.87	1.54				2.31	2.06	1.76				2.56	2.10	1.61		
Minimum link		-2.37	-2.42	-1.99				-2.41	-1.96	-2.38				-2.07	-2.12	-2.03				-1.83	-2.26	-2.68				-1.97	-2.16	-2.44		
Overall mean		-0.12	-0.09	-0.06				0.12	-0.03	-0.06				0.09	-0.07	-0.12				0.10	-0.02	-0.06				0.09	-0.04	-0.07		
Sum of positive links		250.2	187.9	169.8				449.9	330.5	287.6				388.4	280.5	239.5				408.0	312.1	282.1				421.4	303.7	252.2		
Sum of negative links		-492.5	-365.3	-300.1				-198.6	-400.3	-415.8				-207.1	-427.5	-482.5				-198.1	-361.0	-400.4				-230.6	-385.9	-396.3		
Mean of postive links		0.49	0.42	0.39				0.76	0.57	0.51				0.77	0.56	0.51				0.78	0.60	0.54				0.72	0.55	0.49		
Mean of negative links		-0.75	-0.58	-0.52				-0.72	-0.73	-0.70				-0.77	-0.79	-0.75				-0.70	-0.69	-0.70				-0.73	-0.71	-0.66		

9.2.3 Correlation of the a_sym-matrixes for 'Geneva retail stores'



9.3 Classification results for 'Geneva retail stores'

9.3.1 Groups for the set 'all'

Gr	quotK	CatKmax	#	Category	CategoryName
1	0.97	17.21	272	478900	Retail sale via stalls and markets of other goods
1	0.90	14.04	105	478100	Retail sale via stalls and markets of food, beverages and tobacco products
1	0.90	18.77	167	479900	Other retail sale not in stores, stalls or markets
1	0.58	4.83	176	477300	Dispensing chemist in specialised stores
2	1.00	10.24	239	477101	Retail sale of women's clothing
2	1.00	14.80	45	477103	Retail sale of babies' and children's clothing
2	1.00	12.09	268	477105	Retail sale of clothing accessories and clothing equally combined
2	1.00	9.82	103	477201	Retail sale of footwear
2	1.00	13.75	245	477700	Retail sale of watches and jewellery in specialised stores
2	0.96	11.07	49	475100	Retail sale of textiles in specialised stores
2	0.95	4.73	76	474300	Retail sale of audio and video equipment in specialised stores
2	0.95	13.84	25	477202	Retail sale of leather and travel goods
2	0.91	12.59	50	477102	Retail sale of men's clothing
2	0.90	6.60	170	477806	Other retail sale of goods in specialised stores n.e.c.
2	0.86	10.30	76	476100	Retail sale of books in specialised stores
2	0.81	12.48	26	476202	Retail sale of writing materials and stationery
2	0.81	8.99	60	477804	Retail sale of gifts and souvenirs
2	0.81	13.30	66	477901	Retail sale of antiques
2	0.80	8.96	64	475903	Retail sale of household utensils and fittings n.e.c.
2	0.79	13.11	118	477805	Art trade
2	0.79	7.79	90	477802	Retail sale of glasses and other optical goods
2	0.76	15.10	34	476500	Retail sale of games and toys in specialised stores
2	0.75	14.00	49	4775	Retail sale of cosmetic and toilet articles in specialised stores
2	0.72	16.92	21	476300	Retail sale of music and video recordings in specialised stores
2	0.64	7.80	99	4776	Retail sale of flowers, plants, seeds, fertilisers, pet animals and pet food in specialised stores
2	0.61	5.33	82	4764	Retail sale of sporting equipment in specialised stores
2	0.58	16.32	18	4719	Other retail sale in non-specialised stores
2	0.57	5.81	156	472401	Retail sale of bread, cakes, flour confectionery and sugar confectionery in specialised stores
2	0.56	9.10	81	477902	Retail sale of second-hand goods in stores
2	0.53	9.77	62	475400	Retail sale of electrical household appliances in specialised stores
2	0.41	9.70	79	475902	Retail sale of furniture
3	0.97	14.42	78	479100	Retail sale via mail order houses or via Internet
3	0.89	8.07	105	474100	Retail sale of computers, peripheral units and software in specialised stores
3	0.87	10.74	24	475901	Retail sale of musical instruments
3	0.78	8.77	22	477400	Retail sale of medical and orthopaedic goods in specialised stores
4	1.00	21.49	79	473000	Retail sale of automotive fuel in specialised stores
4	0.98	16.09	60	4752	Retail sale of hardware, paints and glass in specialised stores
4	0.97	7.35	48	472500	Retail sale of beverages in specialised stores
4	0.95	8.52	131	472402	Bakeries - tea rooms
4	0.89	4.52	210	4729	Other retail sale of food in specialised stores
4	0.83	16.62	35	474200	Retail sale of telecommunications equipment in specialised stores
4	0.82	6.03	191	4711	Retail sale in non-specialised stores with food, beverages or tobacco predominating
4	0.71	5.49	267	476201	Retail sale of newspapers and magazines, newspaper stands
4	0.67	8.00	68	472200	Retail sale of meat and meat products in specialised stores
4	0.48	9.62	17	475300	Retail sale of carpets, rugs, wall and floor coverings in specialised stores

9.3.2 Groups for the subset A

Gr	quotK	CatKmax	#	Category	CategoryName
1	0.88	13.57	210	4729	Other retail sale of food in specialised stores
1	0.75	18.09	131	472402	Bakeries - tea rooms
1	0.70	10.83	82	4764	Retail sale of sporting equipment in specialised stores
1	0.61	19.60	99	4776	Retail sale of flowers, plants, seeds, fertilisers, pet animals and pet food in specialised stores
1	0.40	13.15	81	477902	Retail sale of second-hand goods in stores
1	0.31	18.45	68	472200	Retail sale of meat and meat products in specialised stores
1	0.26	14.75	17	475300	Retail sale of carpets, rugs, wall and floor coverings in specialised stores
2	0.88	6.30	35	474200	Retail sale of telecommunications equipment in specialised stores
2	0.70	19.46	60	4752	Retail sale of hardware, paints and glass in specialised stores
2	0.51	13.98	64	475903	Retail sale of household utensils and fittings n.e.c.
2	0.42	16.23	34	476500	Retail sale of games and toys in specialised stores
2	0.39	12.39	24	475901	Retail sale of musical instruments
3	0.83	21.32	105	474100	Retail sale of computers, peripheral units and software in specialised stores
3	0.74	14.50	78	479100	Retail sale via mail order houses or via Internet
4	0.96	10.36	176	477300	Dispensing chemist in specialised stores
4	0.92	17.48	79	473000	Retail sale of automotive fuel in specialised stores
4	0.81	12.52	191	4711	Retail sale in non-specialised stores with food, beverages or tobacco predominating
4	0.66	8.50	267	476201	Retail sale of newspapers and magazines, newspaper stands
5	1.00	19.47	245	477700	Retail sale of watches and jewellery in specialised stores
5	0.96	18.49	49	4775	Retail sale of cosmetic and toilet articles in specialised stores
5	0.92	20.24	50	477102	Retail sale of men's clothing
5	0.91	10.44	156	472401	Retail sale of bread, cakes, flour confectionery and sugar confectionery in specialised stores
5	0.83	14.69	103	477201	Retail sale of footwear
5	0.80	15.51	268	477105	Retail sale of clothing accessories and clothing equally combined
5	0.79	20.42	49	475100	Retail sale of textiles in specialised stores
5	0.75	21.45	45	477103	Retail sale of babies' and children's clothing
5	0.73	16.38	25	477202	Retail sale of leather and travel goods
5	0.73	11.61	170	477806	Other retail sale of goods in specialised stores n.e.c.
5	0.68	21.04	60	477804	Retail sale of gifts and souvenirs
5	0.66	16.69	76	476100	Retail sale of books in specialised stores
5	0.65	18.18	66	477901	Retail sale of antiques
5	0.62	14.71	90	477802	Retail sale of glasses and other optical goods
5	0.61	18.07	118	477805	Art trade
5	0.59	14.69	79	475902	Retail sale of furniture
5	0.54	12.32	239	477101	Retail sale of women's clothing
5	0.53	13.07	26	476202	Retail sale of writing materials and stationery
5	0.45	10.45	76	474300	Retail sale of audio and video equipment in specialised stores
5	0.44	13.23	22	477400	Retail sale of medical and orthopaedic goods in specialised stores
5	0.44	15.84	48	472500	Retail sale of beverages in specialised stores
5	0.44	15.95	62	475400	Retail sale of electrical household appliances in specialised stores
5	0.39	19.70	18	4719	Other retail sale in non-specialised stores
5	0.17	16.98	21	476300	Retail sale of music and video recordings in specialised stores
6	1.00	19.53	105	478100	Retail sale via stalls and markets of food, beverages and tobacco products
6	1.00	20.20	167	479900	Other retail sale not in stores, stalls or markets
6	1.00	14.50	272	478900	Retail sale via stalls and markets of other goods

9.3.3 Groups for the subset B

Gr	quotK	CatKmax	#	Category	CategoryName
1	0.89	18.29	99	4776	Retail sale of flowers, plants, seeds, fertilisers, pet animals and pet food in specialised stores
1	0.82	20.23	105	478100	Retail sale via stalls and markets of food, beverages and tobacco products
1	0.61	18.78	66	477901	Retail sale of antiques
1	0.58	21.19	49	475100	Retail sale of textiles in specialised stores
1	0.52	11.45	176	477300	Dispensing chemist in specialised stores
1	0.42	18.16	18	4719	Other retail sale in non-specialised stores
1	0.10	14.57	22	477400	Retail sale of medical and orthopaedic goods in specialised stores
2	0.73	14.48	131	472402	Bakeries - tea rooms
2	0.60	16.32	35	474200	Retail sale of telecommunications equipment in specialised stores
2	0.34	14.86	170	477806	Other retail sale of goods in specialised stores n.e.c.
3	0.97	24.43	45	477103	Retail sale of babies' and children's clothing
3	0.91	20.23	245	477700	Retail sale of watches and jewellery in specialised stores
3	0.90	14.87	50	477102	Retail sale of men's clothing
3	0.88	15.17	239	477101	Retail sale of women's clothing
3	0.83	17.49	268	477105	Retail sale of clothing accessories and clothing equally combined
3	0.79	19.23	103	477201	Retail sale of footwear
3	0.78	13.35	156	472401	Retail sale of bread, cakes, flour confectionery and sugar confectionery in specialised stores
3	0.74	7.66	210	4729	Other retail sale of food in specialised stores
3	0.71	17.54	34	476500	Retail sale of games and toys in specialised stores
3	0.70	19.54	82	4764	Retail sale of sporting equipment in specialised stores
3	0.68	18.42	49	4775	Retail sale of cosmetic and toilet articles in specialised stores
3	0.63	19.20	25	477202	Retail sale of leather and travel goods
3	0.61	18.01	26	476202	Retail sale of writing materials and stationery
3	0.56	17.88	118	477805	Art trade
3	0.48	14.79	76	476100	Retail sale of books in specialised stores
3	0.44	15.97	62	475400	Retail sale of electrical household appliances in specialised stores
3	0.43	13.67	76	474300	Retail sale of audio and video equipment in specialised stores
3	0.26	15.36	60	477804	Retail sale of gifts and souvenirs
3	0.26	19.18	21	476300	Retail sale of music and video recordings in specialised stores
3	0.00	10.79	17	475300	Retail sale of carpets, rugs, wall and floor coverings in specialised stores
4	0.96	16.52	167	479900	Other retail sale not in stores, stalls or markets
4	0.89	18.21	272	478900	Retail sale via stalls and markets of other goods
4	0.68	15.03	105	474100	Retail sale of computers, peripheral units and software in specialised stores
5	0.92	13.10	60	4752	Retail sale of hardware, paints and glass in specialised stores
5	0.77	19.99	79	473000	Retail sale of automotive fuel in specialised stores
5	0.74	18.93	79	475902	Retail sale of furniture
5	0.62	26.00	64	475903	Retail sale of household utensils and fittings n.e.c.
5	0.47	15.99	81	477902	Retail sale of second-hand goods in stores
6	0.85	21.71	78	479100	Retail sale via mail order houses or via Internet
6	0.85	20.40	191	4711	Retail sale in non-specialised stores with food, beverages or tobacco predominating
6	0.78	19.05	68	472200	Retail sale of meat and meat products in specialised stores
6	0.69	13.44	48	472500	Retail sale of beverages in specialised stores
6	0.65	10.43	267	476201	Retail sale of newspapers and magazines, newspaper stands
6	0.43	14.95	24	475901	Retail sale of musical instruments
6	0.41	11.69	90	477802	Retail sale of glasses and other optical goods

9.3.4 Groups for the subset C

Gr	quotK	CatKmax	#	Category	CategoryName
1	1.00	19.72	272	478900	Retail sale via stalls and markets of other goods
1	0.95	19.90	105	478100	Retail sale via stalls and markets of food, beverages and tobacco products
1	0.57	15.91	82	4764	Retail sale of sporting equipment in specialised stores
1	0.52	7.94	267	476201	Retail sale of newspapers and magazines, newspaper stands
1	0.31	14.64	76	474300	Retail sale of audio and video equipment in specialised stores
1	0.28	12.43	81	477902	Retail sale of second-hand goods in stores
2	0.88	12.75	176	477300	Dispensing chemist in specialised stores
2	0.72	17.36	78	479100	Retail sale via mail order houses or via Internet
3	1.00	21.91	45	477103	Retail sale of babies' and children's clothing
3	1.00	17.88	268	477105	Retail sale of clothing accessories and clothing equally combined
3	0.96	15.74	239	477101	Retail sale of women's clothing
3	0.94	25.56	25	477202	Retail sale of leather and travel goods
3	0.91	17.14	49	475100	Retail sale of textiles in specialised stores
3	0.88	20.28	49	4775	Retail sale of cosmetic and toilet articles in specialised stores
3	0.87	21.02	245	477700	Retail sale of watches and jewellery in specialised stores
3	0.78	13.64	76	476100	Retail sale of books in specialised stores
3	0.76	17.08	50	477102	Retail sale of men's clothing
3	0.72	20.90	34	476500	Retail sale of games and toys in specialised stores
3	0.72	12.00	156	472401	Retail sale of bread, cakes, flour confectionery and sugar confectionery in specialised stores
3	0.72	10.90	103	477201	Retail sale of footwear
3	0.65	15.58	21	476300	Retail sale of music and video recordings in specialised stores
3	0.65	15.17	90	477802	Retail sale of glasses and other optical goods
3	0.65	18.50	62	475400	Retail sale of electrical household appliances in specialised stores
3	0.64	22.49	26	476202	Retail sale of writing materials and stationery
3	0.62	11.91	210	4729	Other retail sale of food in specialised stores
3	0.56	11.07	60	477804	Retail sale of gifts and souvenirs
3	0.49	19.18	118	477805	Art trade
3	0.44	14.10	18	4719	Other retail sale in non-specialised stores
3	0.31	17.13	66	477901	Retail sale of antiques
3	0.22	12.73	17	475300	Retail sale of carpets, rugs, wall and floor coverings in specialised stores
4	1.00	3.87	35	474200	Retail sale of telecommunications equipment in specialised stores
4	1.00	14.90	60	4752	Retail sale of hardware, paints and glass in specialised stores
4	0.86	13.90	79	473000	Retail sale of automotive fuel in specialised stores
4	0.78	20.35	105	474100	Retail sale of computers, peripheral units and software in specialised stores
4	0.51	15.46	64	475903	Retail sale of household utensils and fittings n.e.c.
4	0.28	13.13	24	475901	Retail sale of musical instruments
5	1.00	21.34	167	479900	Other retail sale not in stores, stalls or markets
5	0.72	12.61	170	477806	Other retail sale of goods in specialised stores n.e.c.
6	0.80	18.65	48	472500	Retail sale of beverages in specialised stores
6	0.78	16.02	191	4711	Retail sale in non-specialised stores with food, beverages or tobacco predominating
6	0.72	17.44	131	472402	Bakeries - tea rooms
6	0.54	9.07	22	477400	Retail sale of medical and orthopaedic goods in specialised stores
6	0.51	13.57	79	475902	Retail sale of furniture
6	0.42	11.18	99	4776	Retail sale of flowers, plants, seeds, fertilisers, pet animals and pet food in specialised stores
6	0.42	14.19	68	472200	Retail sale of meat and meat products in specialised stores

9.3.5 Groups for the subset D

Gr	quotK	CatKmax	#	Category	CategoryName
1	1.00	16.06	50	477102	Retail sale of men's clothing
1	0.96	11.59	268	477105	Retail sale of clothing accessories and clothing equally combined
1	0.95	12.99	103	477201	Retail sale of footwear
1	0.89	17.13	245	477700	Retail sale of watches and jewellery in specialised stores
1	0.86	19.84	45	477103	Retail sale of babies' and children's clothing
1	0.84	12.28	99	4776	Retail sale of flowers, plants, seeds, fertilisers, pet animals and pet food in specialised stores
1	0.82	18.17	25	477202	Retail sale of leather and travel goods
1	0.79	12.77	90	477802	Retail sale of glasses and other optical goods
1	0.75	11.67	176	477300	Dispensing chemist in specialised stores
1	0.73	16.40	239	477101	Retail sale of women's clothing
1	0.69	21.65	118	477805	Art trade
1	0.68	12.46	156	472401	Retail sale of bread, cakes, flour confectionery and sugar confectionery in specialised stores
1	0.62	9.40	170	477806	Other retail sale of goods in specialised stores n.e.c.
1	0.60	15.98	26	476202	Retail sale of writing materials and stationery
1	0.59	13.36	49	475100	Retail sale of textiles in specialised stores
1	0.54	13.14	82	4764	Retail sale of sporting equipment in specialised stores
1	0.53	14.75	76	476100	Retail sale of books in specialised stores
1	0.50	21.19	66	477901	Retail sale of antiques
1	0.50	20.03	62	475400	Retail sale of electrical household appliances in specialised stores
1	0.48	16.05	49	4775	Retail sale of cosmetic and toilet articles in specialised stores
1	0.46	10.30	81	477902	Retail sale of second-hand goods in stores
1	0.44	14.34	18	4719	Other retail sale in non-specialised stores
1	0.40	12.22	17	475300	Retail sale of carpets, rugs, wall and floor coverings in specialised stores
1	0.38	18.88	64	475903	Retail sale of household utensils and fittings n.e.c.
1	0.26	17.14	21	476300	Retail sale of music and video recordings in specialised stores
1	0.12	13.11	60	477804	Retail sale of gifts and souvenirs
2	0.83	18.42	105	478100	Retail sale via stalls and markets of food, beverages and tobacco products
2	0.49	9.96	34	476500	Retail sale of games and toys in specialised stores
3	0.90	9.24	210	4729	Other retail sale of food in specialised stores
3	0.89	11.46	79	473000	Retail sale of automotive fuel in specialised stores
3	0.68	11.79	22	477400	Retail sale of medical and orthopaedic goods in specialised stores
3	0.67	11.62	79	475902	Retail sale of furniture
3	0.56	19.46	60	4752	Retail sale of hardware, paints and glass in specialised stores
3	0.44	15.04	48	472500	Retail sale of beverages in specialised stores
3	0.42	18.58	35	474200	Retail sale of telecommunications equipment in specialised stores
3	0.31	10.94	68	472200	Retail sale of meat and meat products in specialised stores
4	0.95	24.26	272	478900	Retail sale via stalls and markets of other goods
4	0.91	11.60	131	472402	Bakeries - tea rooms
4	0.91	14.13	78	479100	Retail sale via mail order houses or via Internet
4	0.85	14.08	24	475901	Retail sale of musical instruments
5	0.94	16.02	167	479900	Other retail sale not in stores, stalls or markets
5	0.81	9.95	191	4711	Retail sale in non-specialised stores with food, beverages or tobacco predominating
5	0.57	9.65	267	476201	Retail sale of newspapers and magazines, newspaper stands
5	0.50	14.56	105	474100	Retail sale of computers, peripheral units and software in specialised stores
5	0.44	9.71	76	474300	Retail sale of audio and video equipment in specialised stores

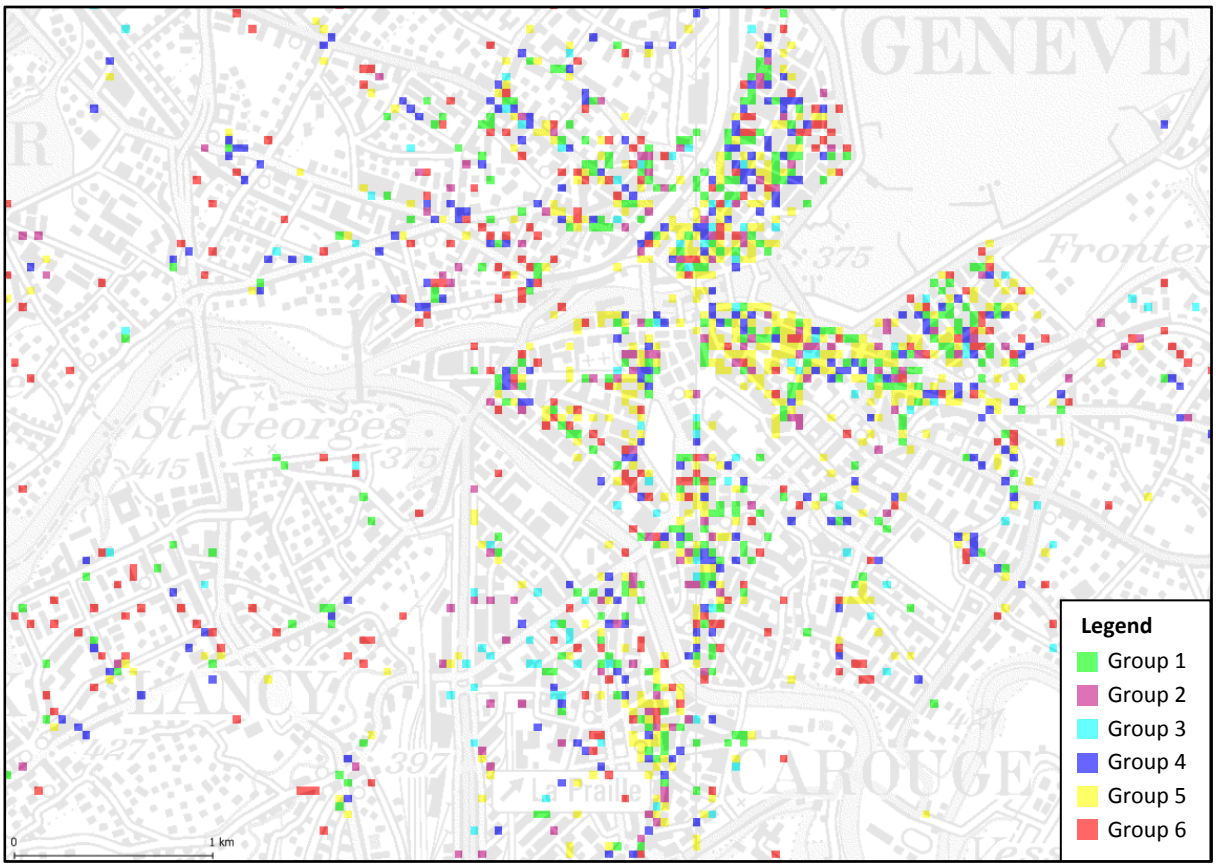
9.3.6 Groups for the set hecto-all

Gr	quotK	CatKmax	#	Category	CategoryName
1	0.92	14.67	78	479100	Retail sale via mail order houses or via Internet
1	0.82	7.92	105	474100	Retail sale of computers, peripheral units and software in specialised stores
1	0.81	11.63	24	475901	Retail sale of musical instruments
1	0.68	9.05	22	477400	Retail sale of medical and orthopaedic goods in specialised stores
2	1.00	11.22	239	477101	Retail sale of women's clothing
2	1.00	13.22	45	477103	Retail sale of babies' and children's clothing
2	1.00	11.27	268	477105	Retail sale of clothing accessories and clothing equally combined
2	1.00	10.00	103	477201	Retail sale of footwear
2	1.00	14.55	245	477700	Retail sale of watches and jewellery in specialised stores
2	0.96	5.37	76	474300	Retail sale of audio and video equipment in specialised stores
2	0.96	11.31	49	475100	Retail sale of textiles in specialised stores
2	0.95	13.89	25	477202	Retail sale of leather and travel goods
2	0.95	11.24	26	476202	Retail sale of writing materials and stationery
2	0.95	6.92	170	477806	Other retail sale of goods in specialised stores n.e.c.
2	0.90	9.64	76	476100	Retail sale of books in specialised stores
2	0.89	11.41	50	477102	Retail sale of men's clothing
2	0.85	13.43	49	4775	Retail sale of cosmetic and toilet articles in specialised stores
2	0.84	7.71	60	477804	Retail sale of gifts and souvenirs
2	0.83	11.76	66	477901	Retail sale of antiques
2	0.80	16.34	21	476300	Retail sale of music and video recordings in specialised stores
2	0.79	8.52	90	477802	Retail sale of glasses and other optical goods
2	0.77	13.00	118	477805	Art trade
2	0.74	13.03	34	476500	Retail sale of games and toys in specialised stores
2	0.74	15.50	18	4719	Other retail sale in non-specialised stores
2	0.68	6.94	99	4776	Retail sale of flowers, plants, seeds, fertilisers, pet animals and pet food in specialised stores
2	0.67	3.70	156	472401	Retail sale of bread, cakes, flour confectionery and sugar confectionery in specialised stores
2	0.65	8.50	64	475903	Retail sale of household utensils and fittings n.e.c.
2	0.58	8.15	62	475400	Retail sale of electrical household appliances in specialised stores
2	0.54	4.87	82	4764	Retail sale of sporting equipment in specialised stores
2	0.51	7.86	81	477902	Retail sale of second-hand goods in stores
3	1.00	21.84	79	473000	Retail sale of automotive fuel in specialised stores
3	1.00	13.94	60	4752	Retail sale of hardware, paints and glass in specialised stores
3	0.95	7.51	131	472402	Bakeries - tea rooms
3	0.93	7.48	48	472500	Retail sale of beverages in specialised stores
3	0.91	4.26	210	4729	Other retail sale of food in specialised stores
3	0.87	7.46	68	472200	Retail sale of meat and meat products in specialised stores
3	0.81	13.10	35	474200	Retail sale of telecommunications equipment in specialised stores
3	0.78	6.58	191	4711	Retail sale in non-specialised stores with food, beverages or tobacco predominating
3	0.73	4.04	267	476201	Retail sale of newspapers and magazines, newspaper stands
3	0.61	10.77	17	475300	Retail sale of carpets, rugs, wall and floor coverings in specialised stores
3	0.51	8.39	79	475902	Retail sale of furniture
4	0.97	17.33	272	478900	Retail sale via stalls and markets of other goods
4	0.91	15.49	105	478100	Retail sale via stalls and markets of food, beverages and tobacco products
4	0.85	14.15	167	479900	Other retail sale not in stores, stalls or markets
4	0.60	5.01	176	477300	Dispensing chemist in specialised stores

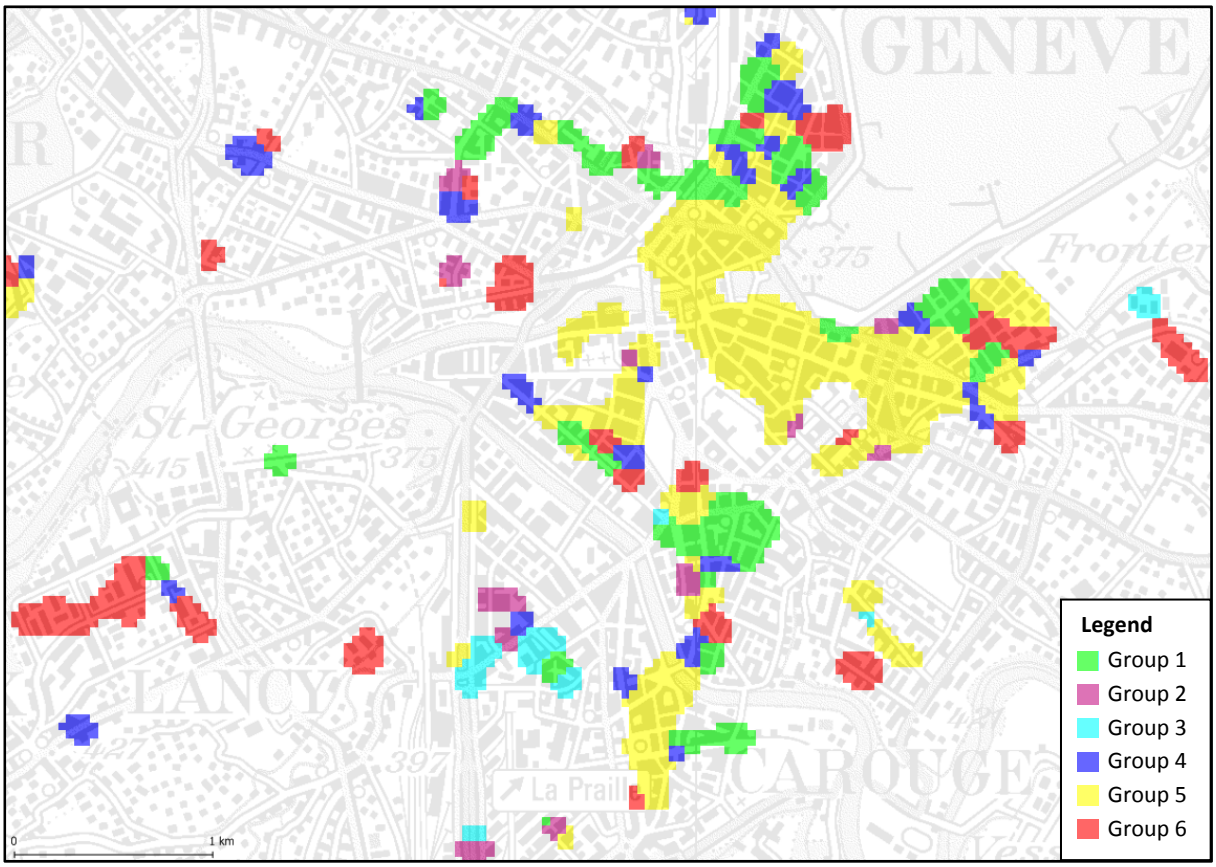
9.3.7 Groups for the subset hecto-A

Gr	quotK	CatKmax	#	Category	CategoryName
1	1.00	12.01	176	477300	Dispensing chemist in specialised stores
1	0.94	19.00	79	473000	Retail sale of automotive fuel in specialised stores
1	0.89	11.22	191	4711	Retail sale in non-specialised stores with food, beverages or tobacco predominating
1	0.59	18.44	60	4752	Retail sale of hardware, paints and glass in specialised stores
2	0.89	17.73	78	479100	Retail sale via mail order houses or via Internet
2	0.81	17.31	105	474100	Retail sale of computers, peripheral units and software in specialised stores
3	1.00	25.00	105	478100	Retail sale via stalls and markets of food, beverages and tobacco products
3	1.00	16.50	272	478900	Retail sale via stalls and markets of other goods
3	0.95	18.00	167	479900	Other retail sale not in stores, stalls or markets
3	0.55	8.95	267	476201	Retail sale of newspapers and magazines, newspaper stands
4	1.00	19.13	45	477103	Retail sale of babies' and children's clothing
4	1.00	17.63	245	477700	Retail sale of watches and jewellery in specialised stores
4	0.92	18.00	49	4775	Retail sale of cosmetic and toilet articles in specialised stores
4	0.89	16.27	103	477201	Retail sale of footwear
4	0.87	18.40	50	477102	Retail sale of men's clothing
4	0.82	12.35	90	477802	Retail sale of glasses and other optical goods
4	0.82	14.32	170	477806	Other retail sale of goods in specialised stores n.e.c.
4	0.81	19.70	76	476100	Retail sale of books in specialised stores
4	0.81	19.43	118	477805	Art trade
4	0.80	18.00	268	477105	Retail sale of clothing accessories and clothing equally combined
4	0.79	19.03	49	475100	Retail sale of textiles in specialised stores
4	0.79	9.01	156	472401	Retail sale of bread, cakes, flour confectionery and sugar confectionery in specialised stores
4	0.75	17.16	25	477202	Retail sale of leather and travel goods
4	0.68	18.93	66	477901	Retail sale of antiques
4	0.66	16.72	60	477804	Retail sale of gifts and souvenirs
4	0.66	18.29	26	476202	Retail sale of writing materials and stationery
4	0.60	15.04	239	477101	Retail sale of women's clothing
4	0.56	9.03	76	474300	Retail sale of audio and video equipment in specialised stores
4	0.52	17.23	48	472500	Retail sale of beverages in specialised stores
4	0.44	14.24	79	475902	Retail sale of furniture
4	0.43	14.34	18	4719	Other retail sale in non-specialised stores
4	0.36	12.81	62	475400	Retail sale of electrical household appliances in specialised stores
4	0.36	14.34	22	477400	Retail sale of medical and orthopaedic goods in specialised stores
4	0.15	17.68	21	476300	Retail sale of music and video recordings in specialised stores
5	0.73	15.64	35	474200	Retail sale of telecommunications equipment in specialised stores
5	0.72	13.35	82	4764	Retail sale of sporting equipment in specialised stores
5	0.72	10.71	81	477902	Retail sale of second-hand goods in stores
5	0.70	15.15	131	472402	Bakeries - tea rooms
5	0.69	15.44	64	475903	Retail sale of household utensils and fittings n.e.c.
5	0.68	11.59	210	4729	Other retail sale of food in specialised stores
5	0.60	16.60	99	4776	Retail sale of flowers, plants, seeds, fertilisers, pet animals and pet food in specialised stores
5	0.46	11.88	24	475901	Retail sale of musical instruments
5	0.34	14.04	34	476500	Retail sale of games and toys in specialised stores
5	0.34	15.46	68	472200	Retail sale of meat and meat products in specialised stores
5	0.33	12.32	17	475300	Retail sale of carpets, rugs, wall and floor coverings in specialised stores

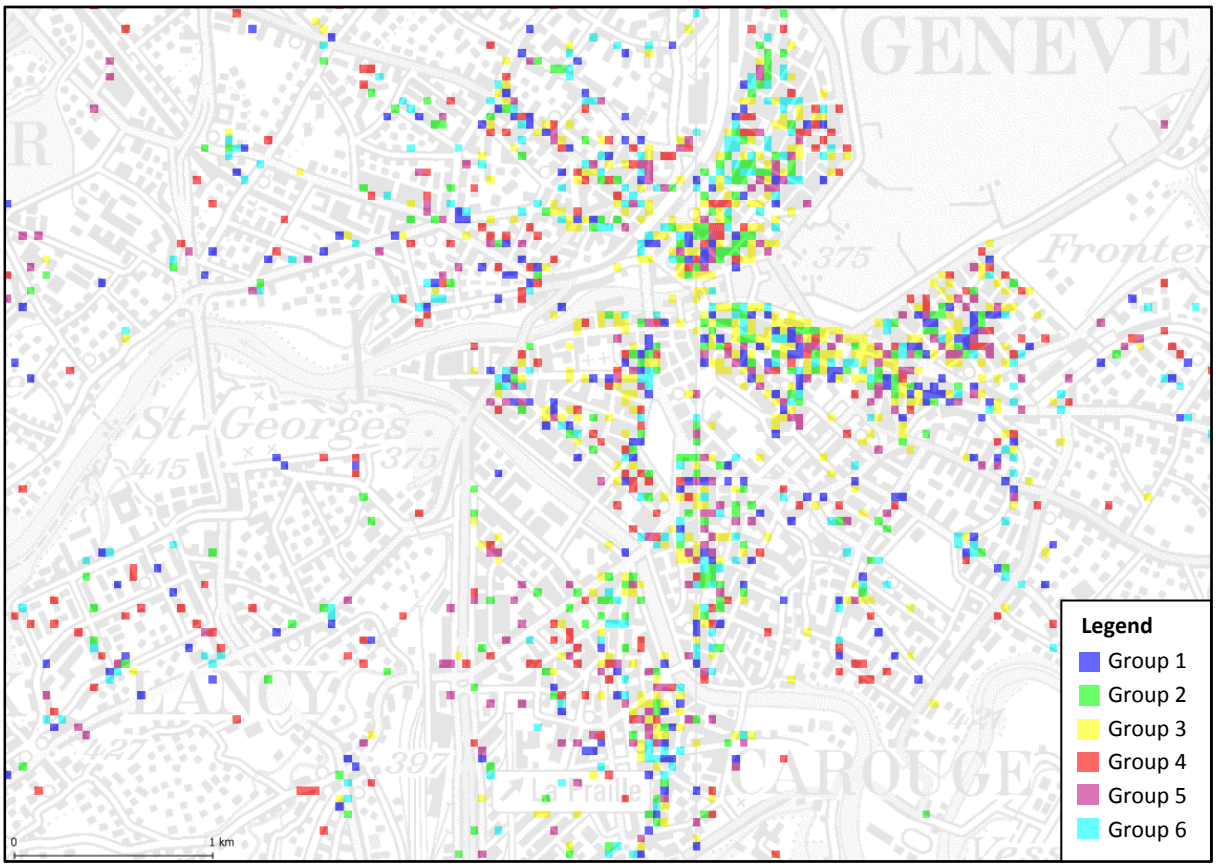
9.3.8 Groups for the subset A represented on 40m-grid



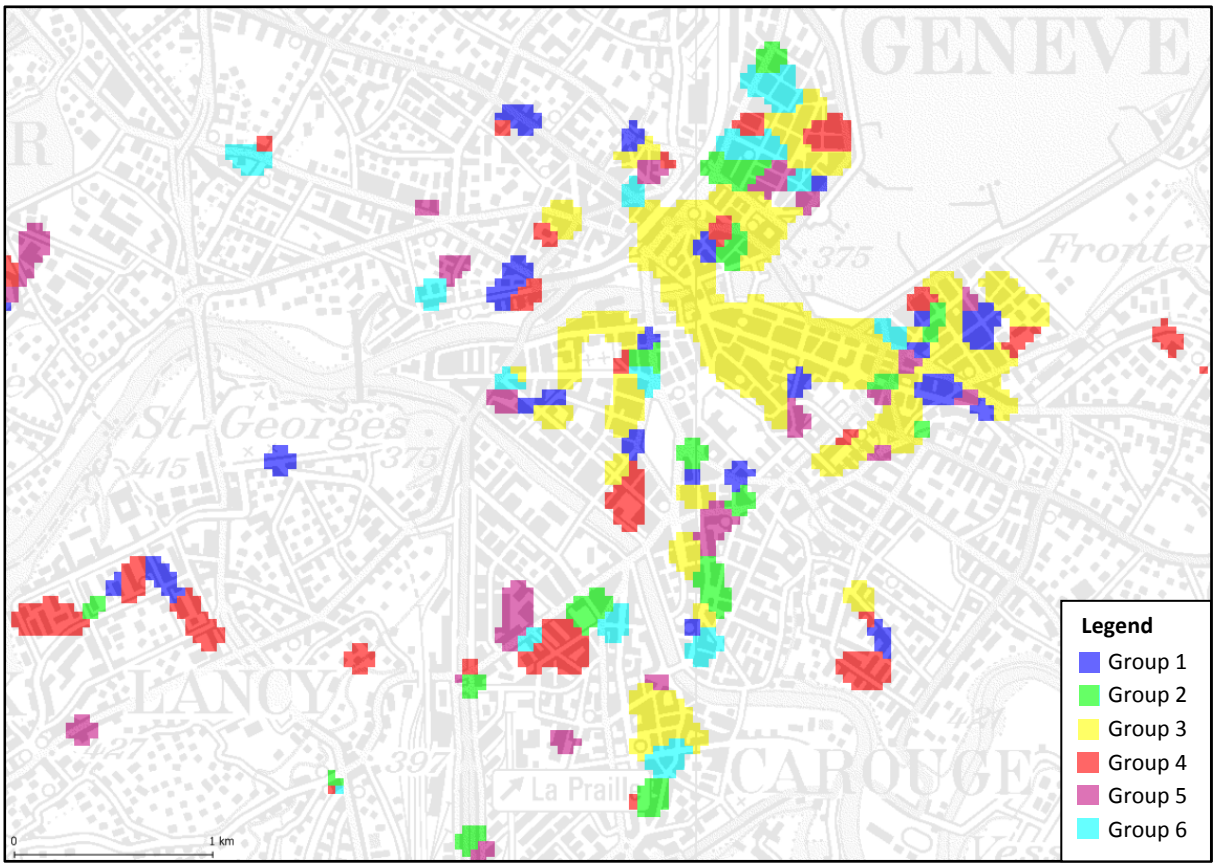
Filtred and simplified:



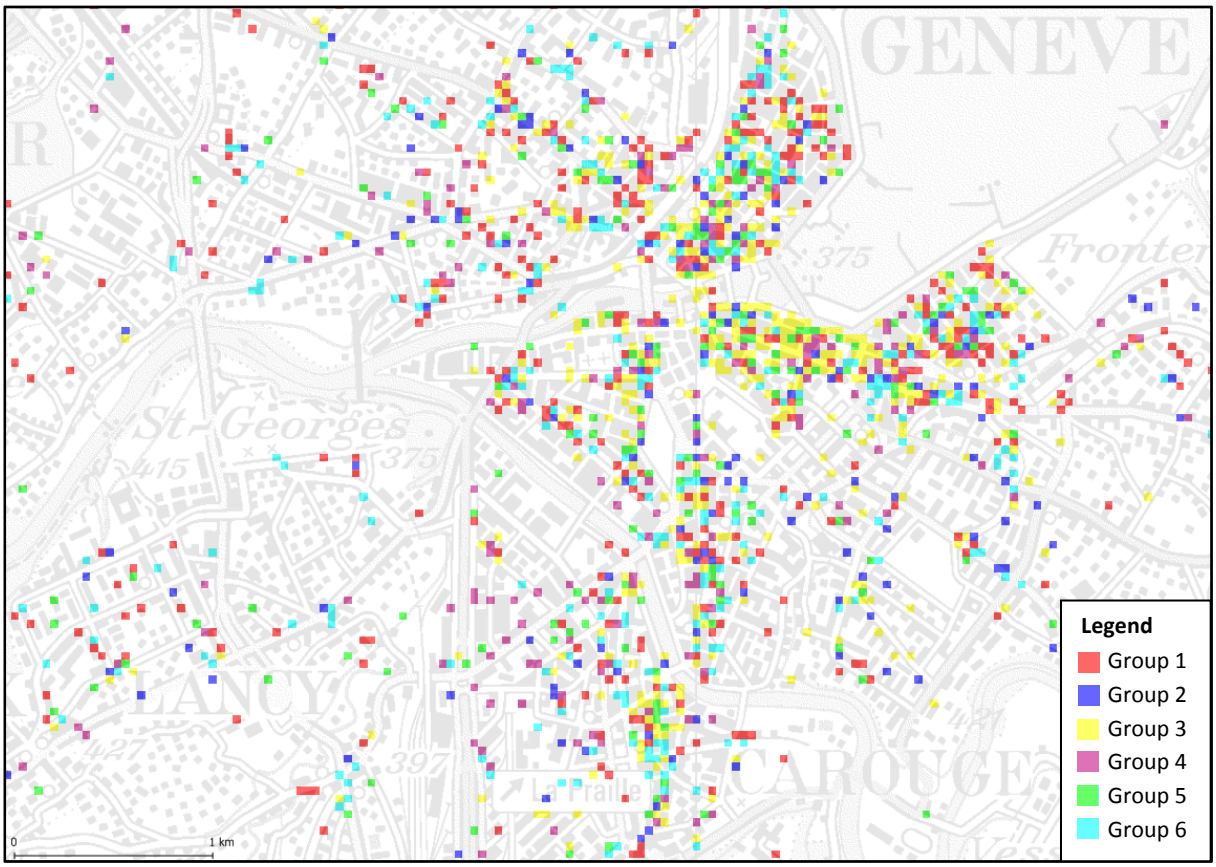
9.3.9 Groups for the subset B represented on 40m-grid



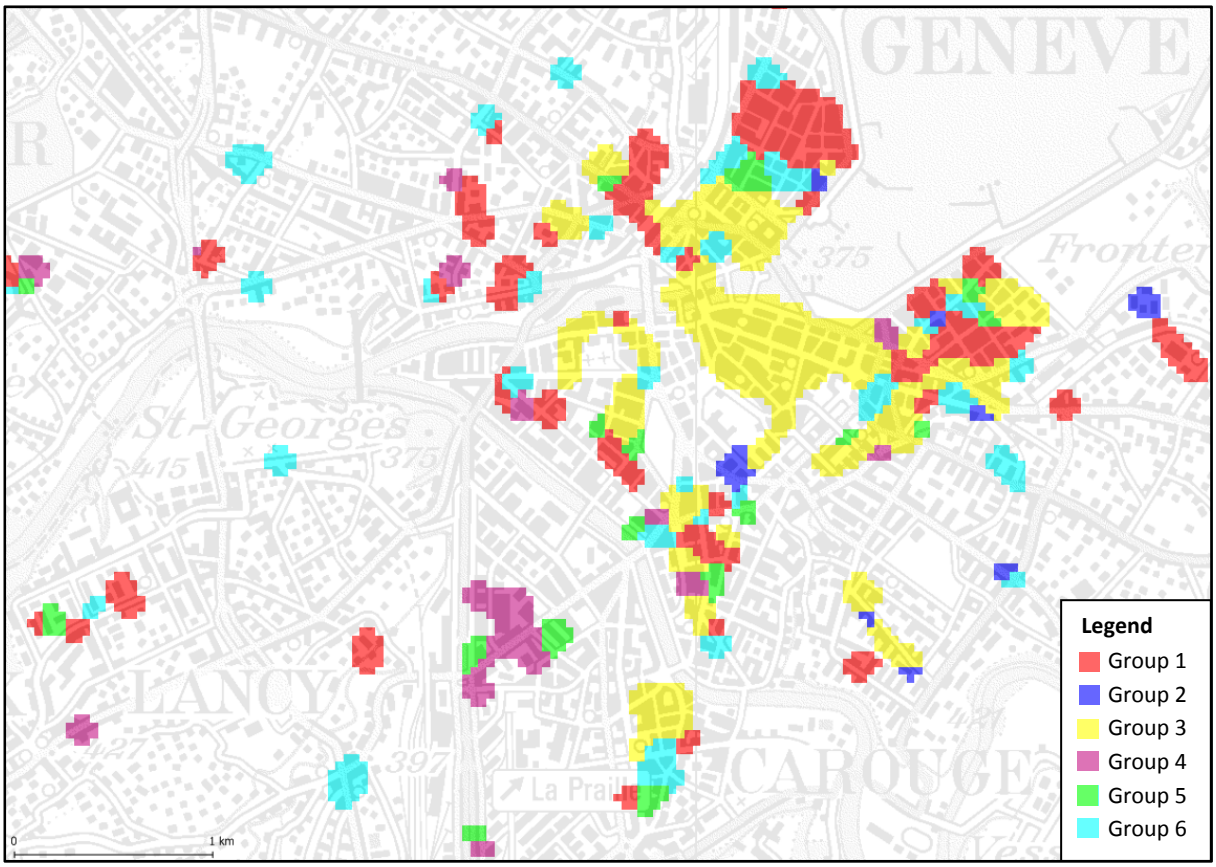
Filtred and simplified:



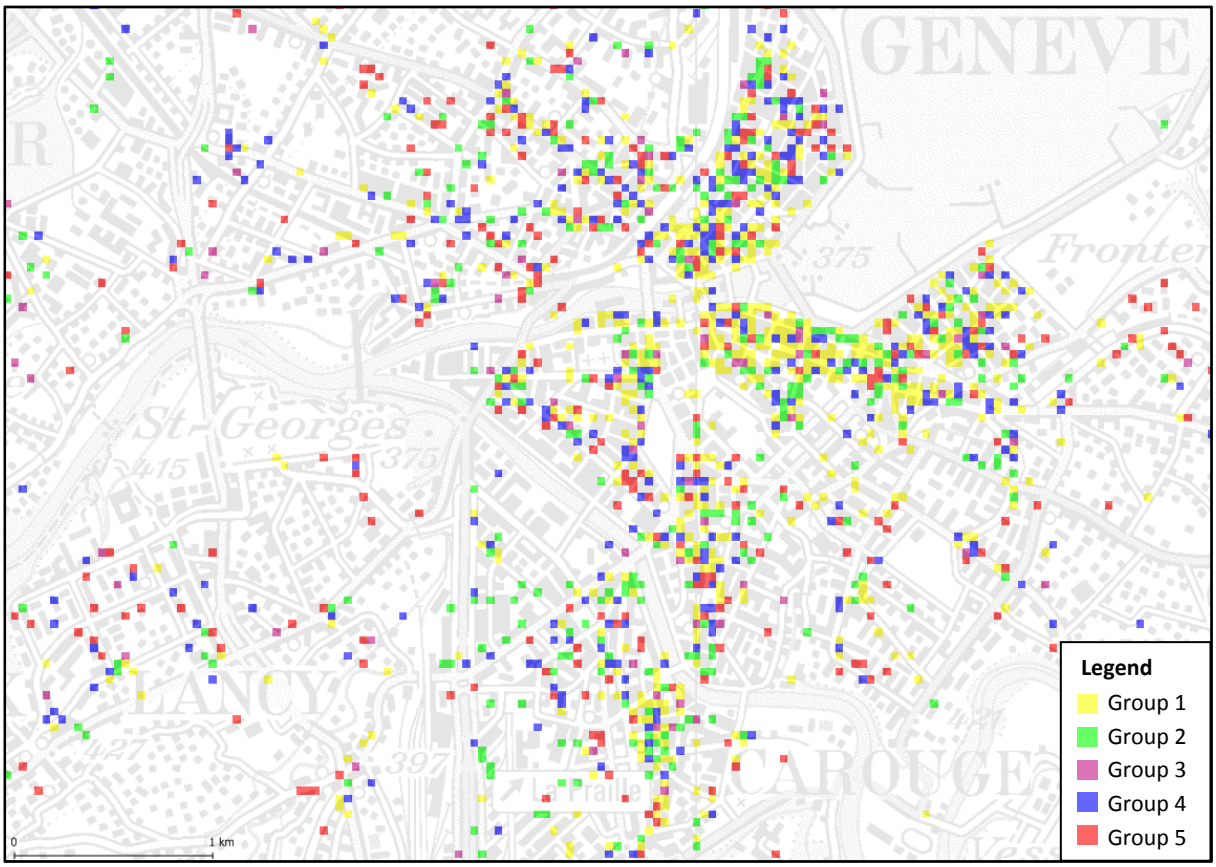
9.3.10 Groups for the subset C represented on 40m-grid



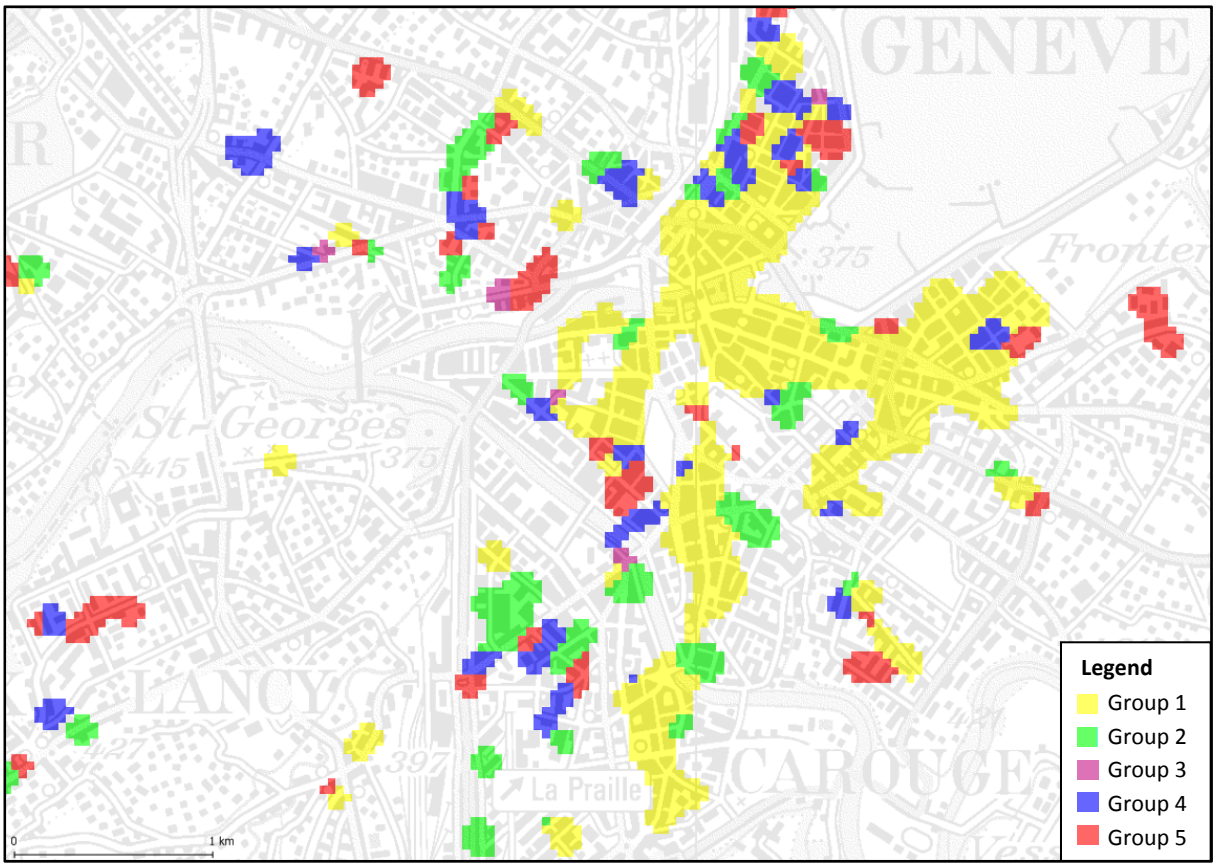
Filtred and simplified:



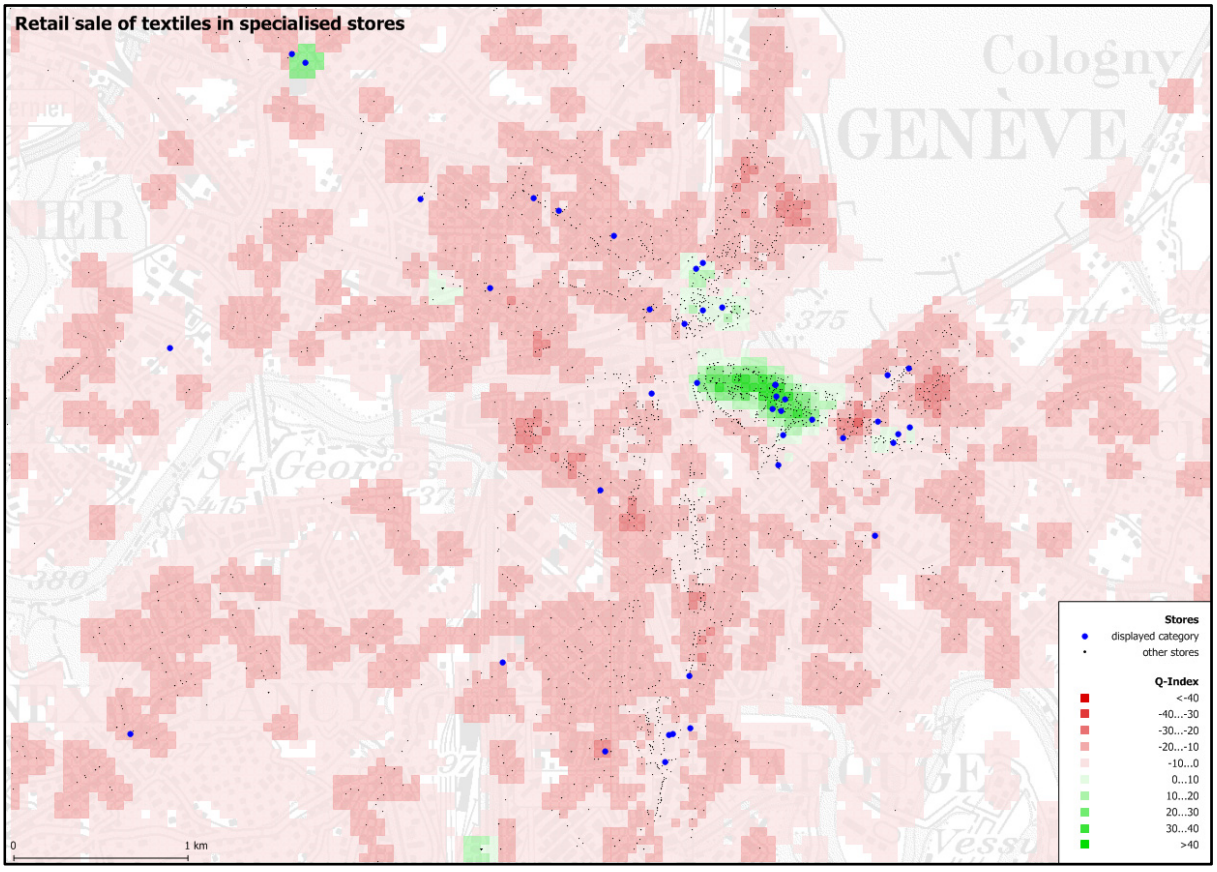
9.3.11 Groups for the subset D represented on 40m-grid

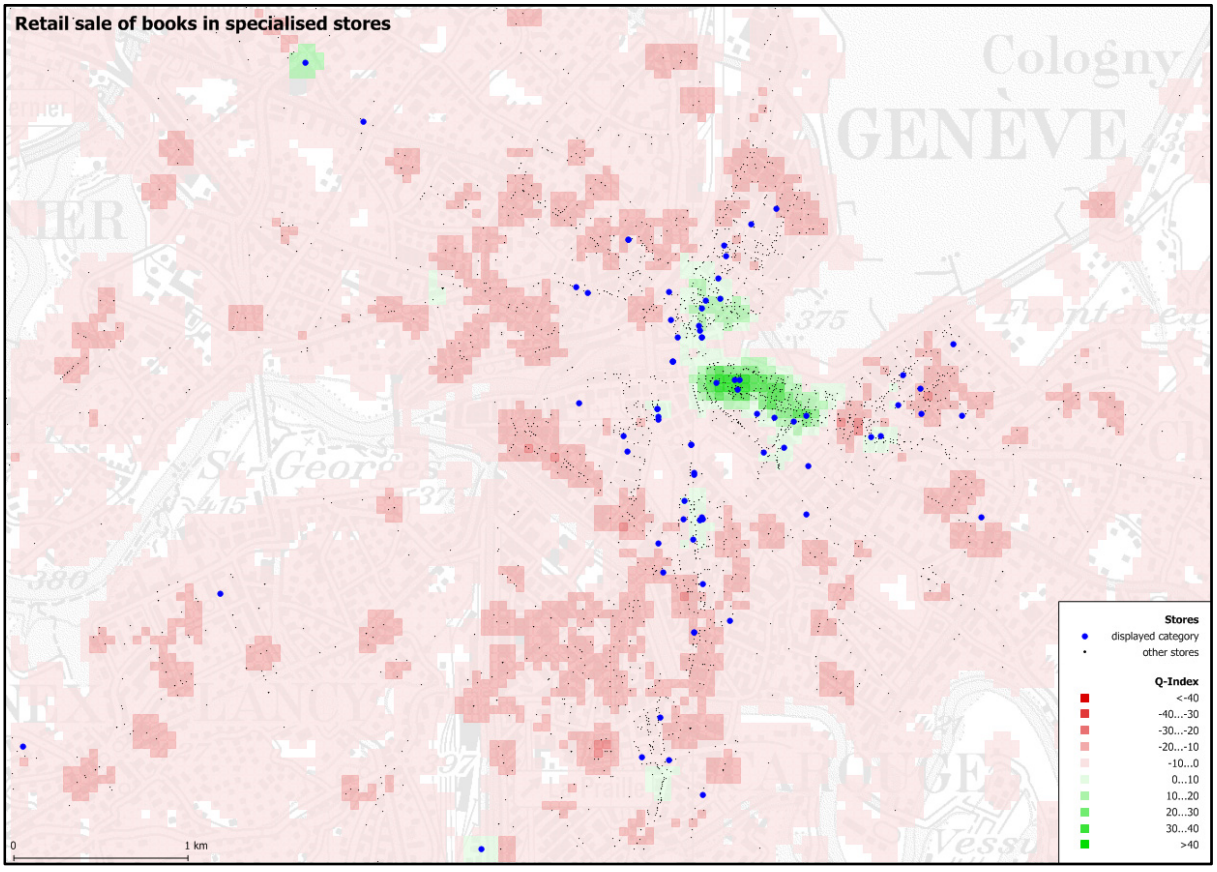


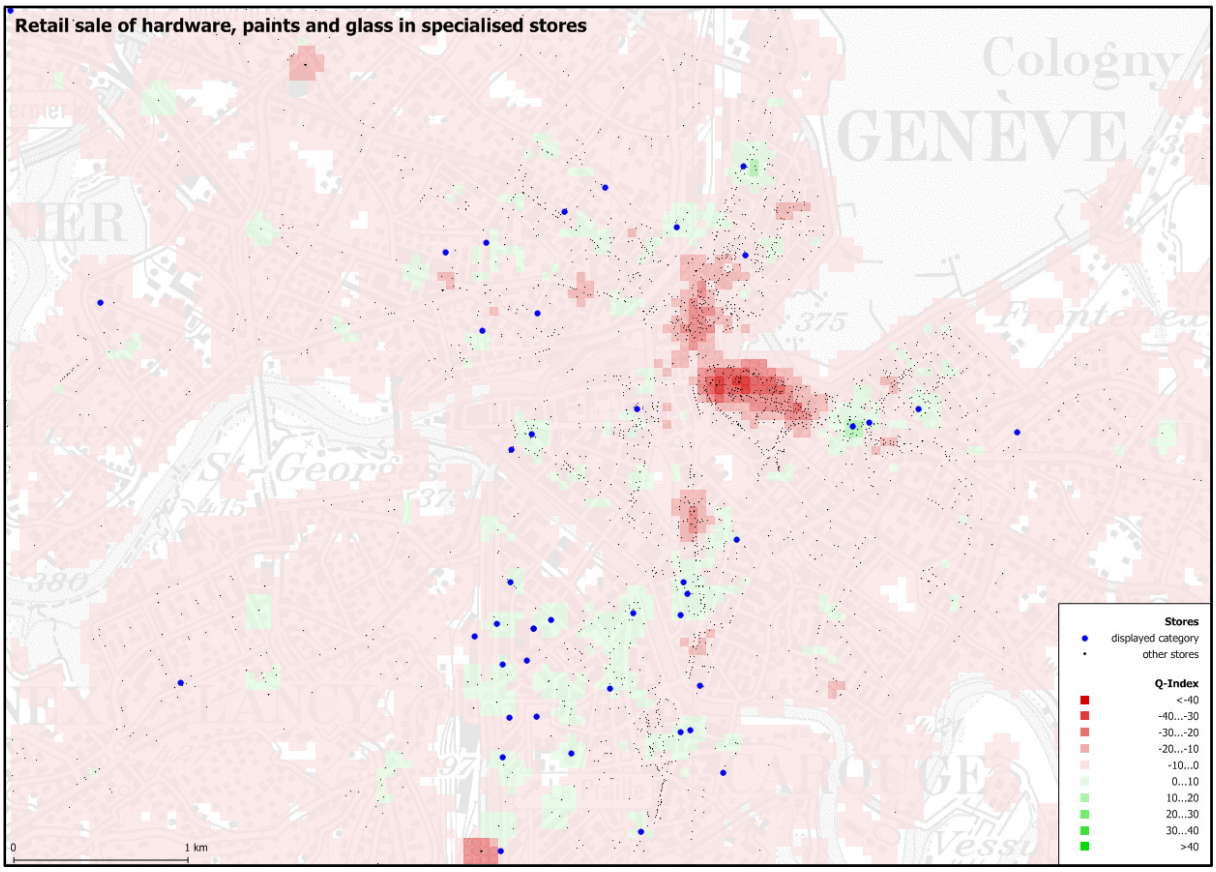
Filtred and simplified:

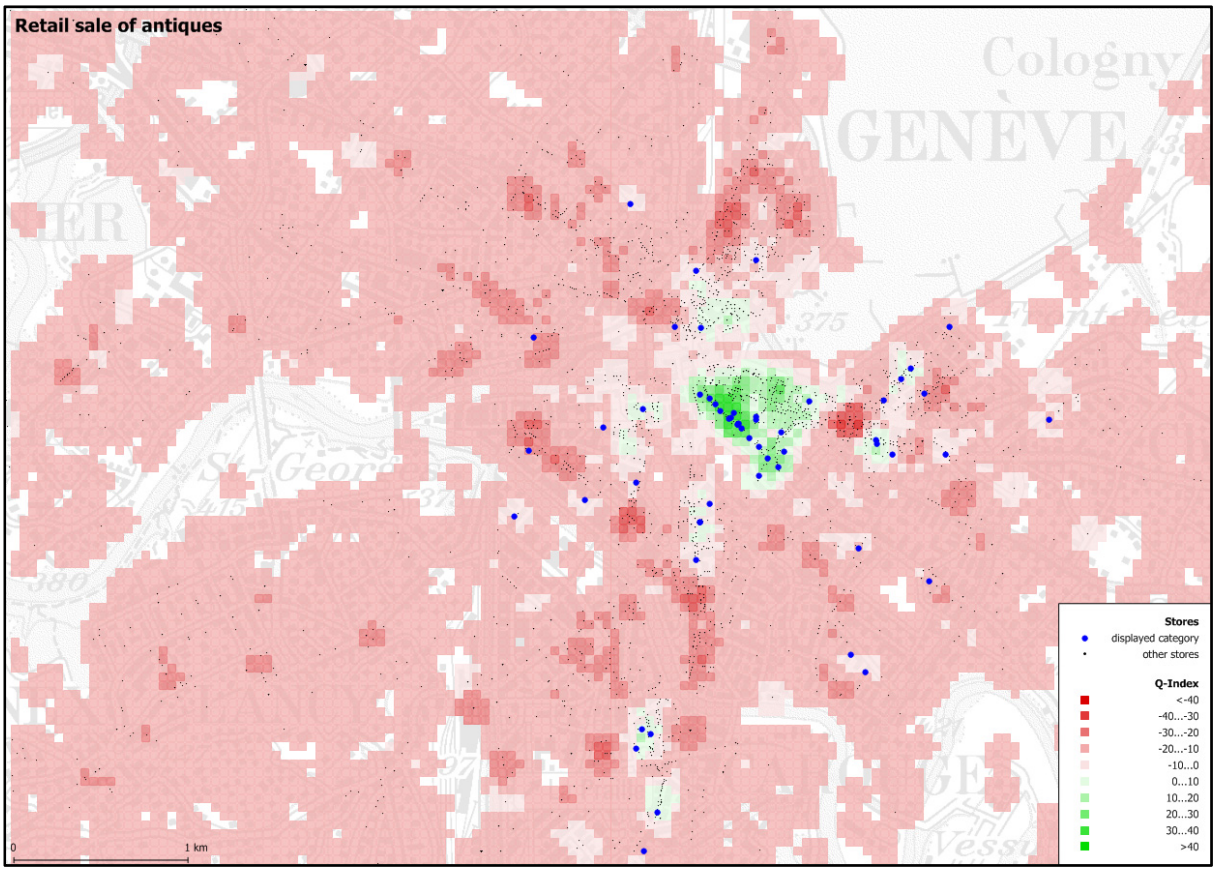


9.4 Q-Index-maps for ‘Geneva retail stores’






Retail sale of hardware, paints and glass in specialised stores
Cologny
GENÈVE
Stores
displayed category
other stores
Q-Index
<-40
-40...-30
-30...-20
-20...-10
-10...0
0...10
10...20
20...30
30...40
>40
0
1 km


Retail sale of antiques
Cologny
GENÈVE
Stores
displayed category
other stores
Q-Index
<-40
-40...-30
-30...-20
-20...-10
-10...0
0...10
10...20
20...30
30...40
>40
0
1 km

9.5 Classification results for 'Geneva activities'

9.5.1 Groups for the set 'all'

Gr	quotK	CatKmax	#	Category	CategoryName
1	0.95	4.45	2614	F0	CONSTRUCTION
1	0.92	2.07	1751	N0	ADMINISTRATIVE AND SUPPORT SERVICE ACTIVITIES
1	0.87	5.58	1640	H0	TRANSPORTATION AND STORAGE
1	0.79	2.45	1800	J0	INFORMATION AND COMMUNICATION
2	1.00	2.81	3139	S0	OTHER SERVICE ACTIVITIES
2	0.73	7.34	344	U0	ACTIVITIES OF EXTRATERRITORIAL ORGANISATIONS AND BODIES
2	0.67	5.33	536	O0	PUBLIC ADMINISTRATION AND DEFENCE; COMPULSORY SOCIAL SECURITY
2	0.66	3.77	1273	P0	EDUCATION
2	0.66	4.05	3108	Q0	HUMAN HEALTH AND SOCIAL WORK ACTIVITIES
2	0.62	3.88	853	R0	ARTS, ENTERTAINMENT AND RECREATION
2	0.48	1.13	2545	I0	ACCOMMODATION AND FOOD SERVICE ACTIVITIES
2	0.47	3.06	6105	M0	PROFESSIONAL, SCIENTIFIC AND TECHNICAL ACTIVITIES
3	0.90	12.51	427	A0	AGRICULTURE, FORESTRY AND FISHING
3	0.85	3.30	1919	C0	MANUFACTURING
3	0.84	7.76	134	E0	WATER SUPPLY; SEWERAGE, WASTE MANAGEMENT AND REMEDIATION ACTIVITIES
3	0.71	1.80	8371	G0	WHOLESALE AND RETAIL TRADE; REPAIR OF MOTOR VEHICLES AND MOTORCYCLES
3	0.70	3.44	1026	L0	REAL ESTATE ACTIVITIES
3	0.54	2.06	2745	K0	FINANCIAL AND INSURANCE ACTIVITIES

9.5.2 Groups for the subset A

Gr	quotK	CatKmax	#	Category	CategoryName
1	1.00	14.27	427	A0	AGRICULTURE, FORESTRY AND FISHING
1	1.00	0.92	3139	S0	OTHER SERVICE ACTIVITIES
1	0.90	2.46	1751	N0	ADMINISTRATIVE AND SUPPORT SERVICE ACTIVITIES
1	0.83	2.39	2545	I0	ACCOMMODATION AND FOOD SERVICE ACTIVITIES
2	0.83	3.35	2614	F0	CONSTRUCTION
2	0.81	6.37	1919	C0	MANUFACTURING
2	0.80	6.84	1640	H0	TRANSPORTATION AND STORAGE
3	1.00	9.56	134	E0	WATER SUPPLY; SEWERAGE, WASTE MANAGEMENT AND REMEDIATION ACTIVITIES
3	0.93	6.50	344	U0	ACTIVITIES OF EXTRATERRITORIAL ORGANISATIONS AND BODIES
3	0.92	5.86	536	O0	PUBLIC ADMINISTRATION AND DEFENCE; COMPULSORY SOCIAL SECURITY
3	0.69	1.99	1800	J0	INFORMATION AND COMMUNICATION
4	1.00	2.92	2745	K0	FINANCIAL AND INSURANCE ACTIVITIES
4	1.00	2.14	6105	M0	PROFESSIONAL, SCIENTIFIC AND TECHNICAL ACTIVITIES
4	0.84	6.09	1273	P0	EDUCATION
4	0.83	5.84	1026	L0	REAL ESTATE ACTIVITIES
4	0.81	3.83	3108	Q0	HUMAN HEALTH AND SOCIAL WORK ACTIVITIES
4	0.80	5.41	853	R0	ARTS, ENTERTAINMENT AND RECREATION
4	0.72	1.40	8371	G0	WHOLESALE AND RETAIL TRADE; REPAIR OF MOTOR VEHICLES AND MOTORCYCLES

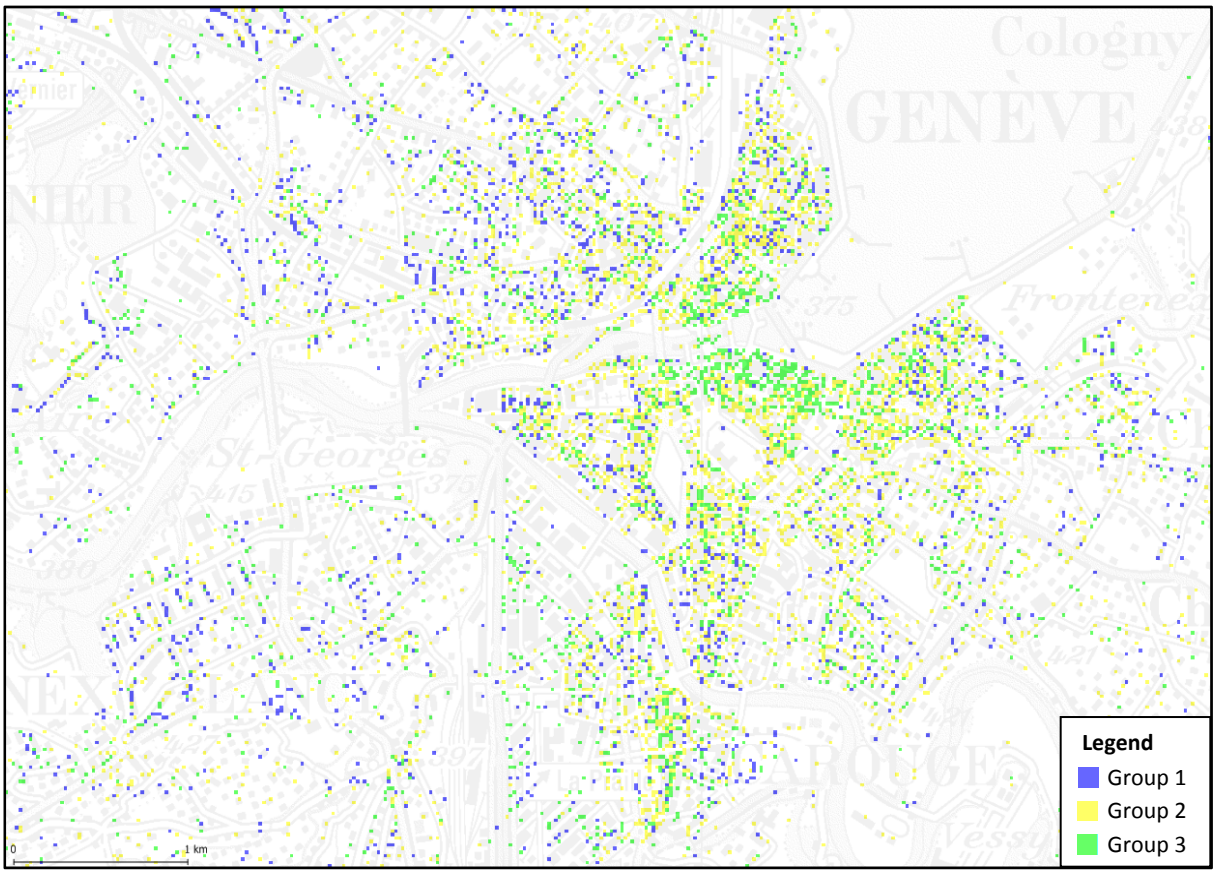
9.5.3 Groups for the subset B

Gr	quotK	CatKmax	#	Category	CategoryName
1	1.00	1.89	3139	S0	OTHER SERVICE ACTIVITIES
1	0.96	4.83	344	U0	ACTIVITIES OF EXTRATERRITORIAL ORGANISATIONS AND BODIES
1	0.88	4.15	853	R0	ARTS, ENTERTAINMENT AND RECREATION
1	0.88	4.56	3108	Q0	HUMAN HEALTH AND SOCIAL WORK ACTIVITIES
1	0.88	6.05	536	O0	PUBLIC ADMINISTRATION AND DEFENCE; COMPULSORY SOCIAL SECURITY
1	0.77	4.23	2545	I0	ACCOMMODATION AND FOOD SERVICE ACTIVITIES
1	0.50	7.00	2745	K0	FINANCIAL AND INSURANCE ACTIVITIES
2	1.00	14.92	427	A0	AGRICULTURE, FORESTRY AND FISHING
2	1.00	9.35	134	E0	WATER SUPPLY; SEWERAGE, WASTE MANAGEMENT AND REMEDIATION ACTIVITIES
2	0.77	3.89	1273	P0	EDUCATION
2	0.76	3.63	1919	C0	MANUFACTURING
2	0.65	5.03	1026	L0	REAL ESTATE ACTIVITIES
2	0.17	1.68	6105	M0	PROFESSIONAL, SCIENTIFIC AND TECHNICAL ACTIVITIES
3	1.00	2.61	1800	J0	INFORMATION AND COMMUNICATION
3	0.99	7.28	1640	H0	TRANSPORTATION AND STORAGE
3	0.99	6.44	2614	F0	CONSTRUCTION
3	0.78	5.44	1751	N0	ADMINISTRATIVE AND SUPPORT SERVICE ACTIVITIES
3	0.22	2.19	8371	G0	WHOLESALE AND RETAIL TRADE; REPAIR OF MOTOR VEHICLES AND MOTORCYCLES

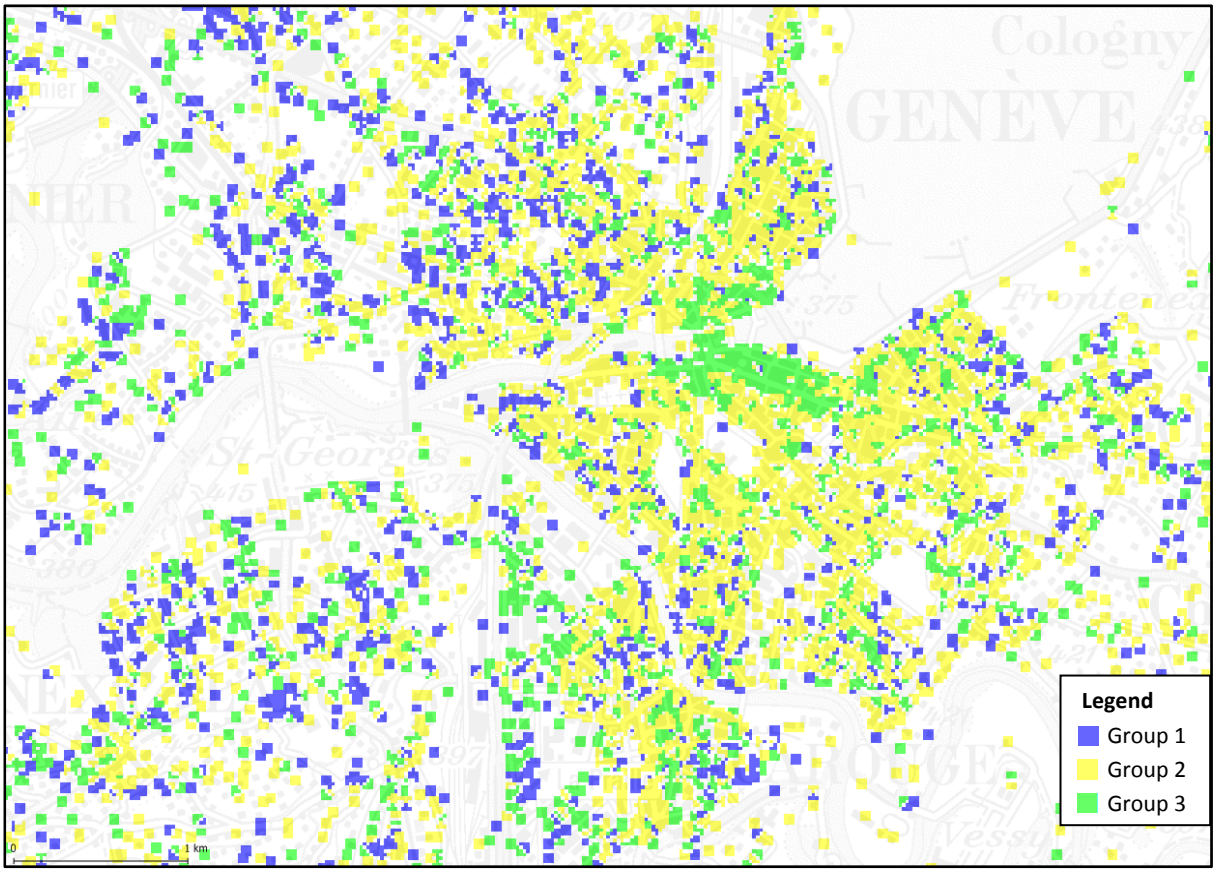
9.5.4 Groups for the subset C

Gr	quotK	CatKmax	#	Category	CategoryName
1	1.00	4.97	8371	G0	WHOLESALE AND RETAIL TRADE; REPAIR OF MOTOR VEHICLES AND MOTORCYCLES
1	1.00	4.53	2545	I0	ACCOMMODATION AND FOOD SERVICE ACTIVITIES
1	1.00	2.01	1800	J0	INFORMATION AND COMMUNICATION
1	1.00	4.61	2745	K0	FINANCIAL AND INSURANCE ACTIVITIES
1	1.00	1.01	6105	M0	PROFESSIONAL, SCIENTIFIC AND TECHNICAL ACTIVITIES
1	0.92	4.76	1026	L0	REAL ESTATE ACTIVITIES
1	0.89	4.49	3139	S0	OTHER SERVICE ACTIVITIES
2	1.00	10.82	427	A0	AGRICULTURE, FORESTRY AND FISHING
2	1.00	3.75	1751	N0	ADMINISTRATIVE AND SUPPORT SERVICE ACTIVITIES
2	0.95	4.24	344	U0	ACTIVITIES OF EXTRATERRITORIAL ORGANISATIONS AND BODIES
2	0.95	3.70	1640	H0	TRANSPORTATION AND STORAGE
3	1.00	5.70	1273	P0	EDUCATION
3	0.98	5.29	3108	Q0	HUMAN HEALTH AND SOCIAL WORK ACTIVITIES
4	1.00	5.14	1919	C0	MANUFACTURING
4	1.00	9.99	134	E0	WATER SUPPLY; SEWERAGE, WASTE MANAGEMENT AND REMEDIATION ACTIVITIES
4	1.00	3.49	2614	F0	CONSTRUCTION
5	1.00	6.60	536	O0	PUBLIC ADMINISTRATION AND DEFENCE; COMPULSORY SOCIAL SECURITY
5	1.00	3.31	853	R0	ARTS, ENTERTAINMENT AND RECREATION

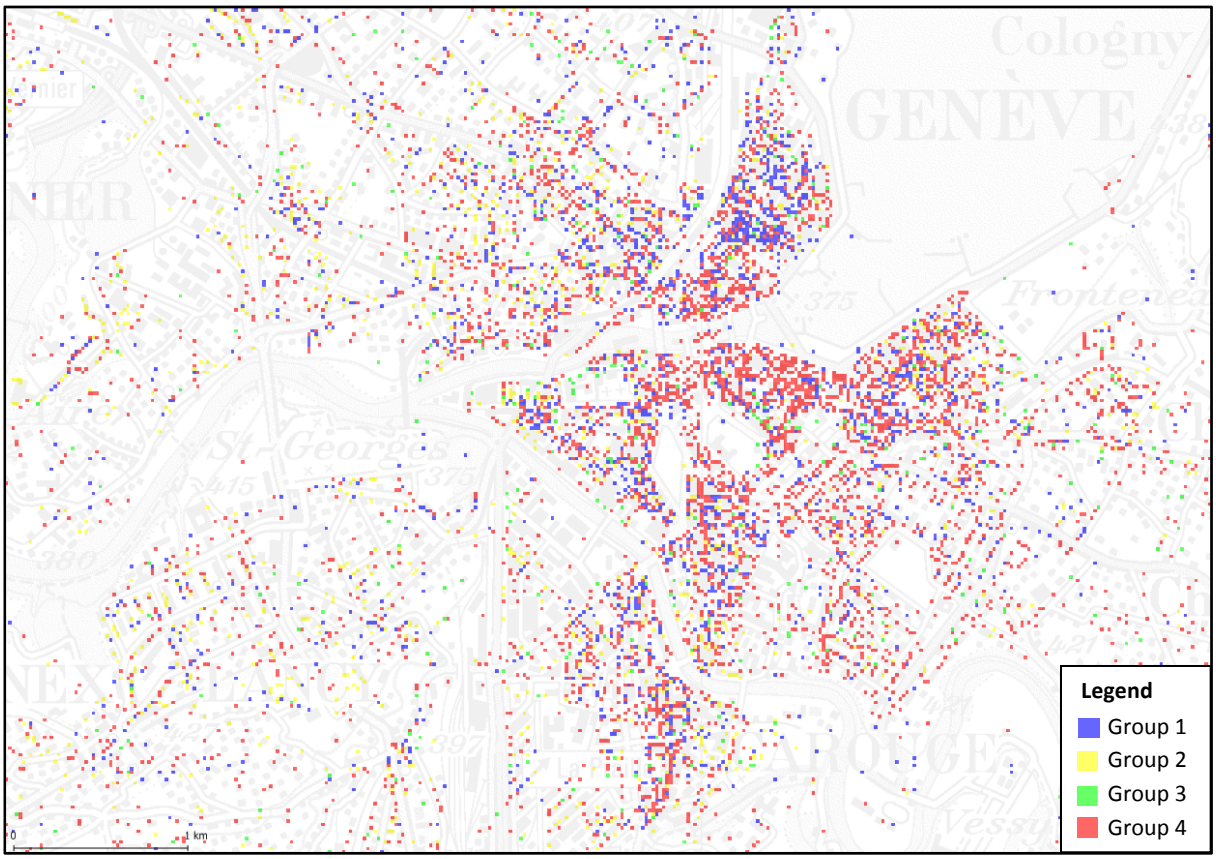
9.5.5 Groups for the set ‘all’ represented on 20m-grid



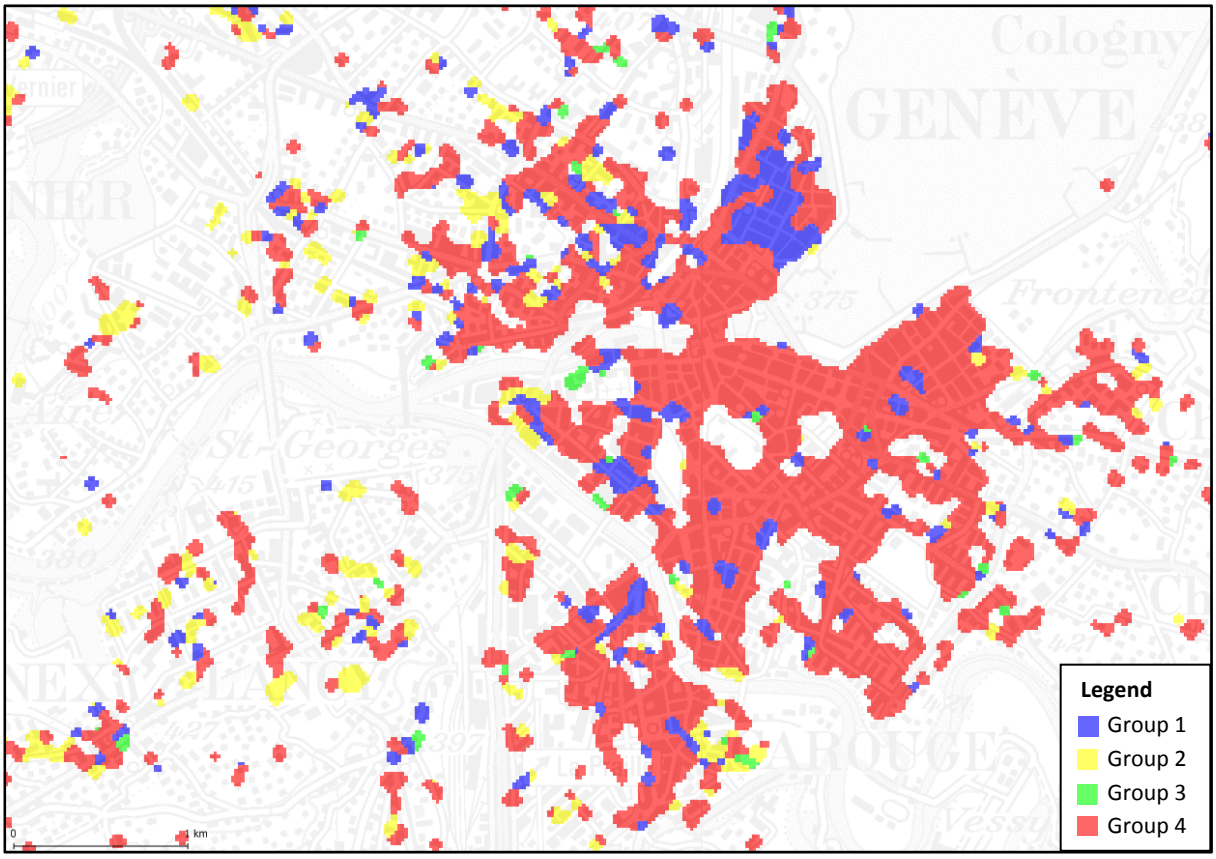
Filtered:



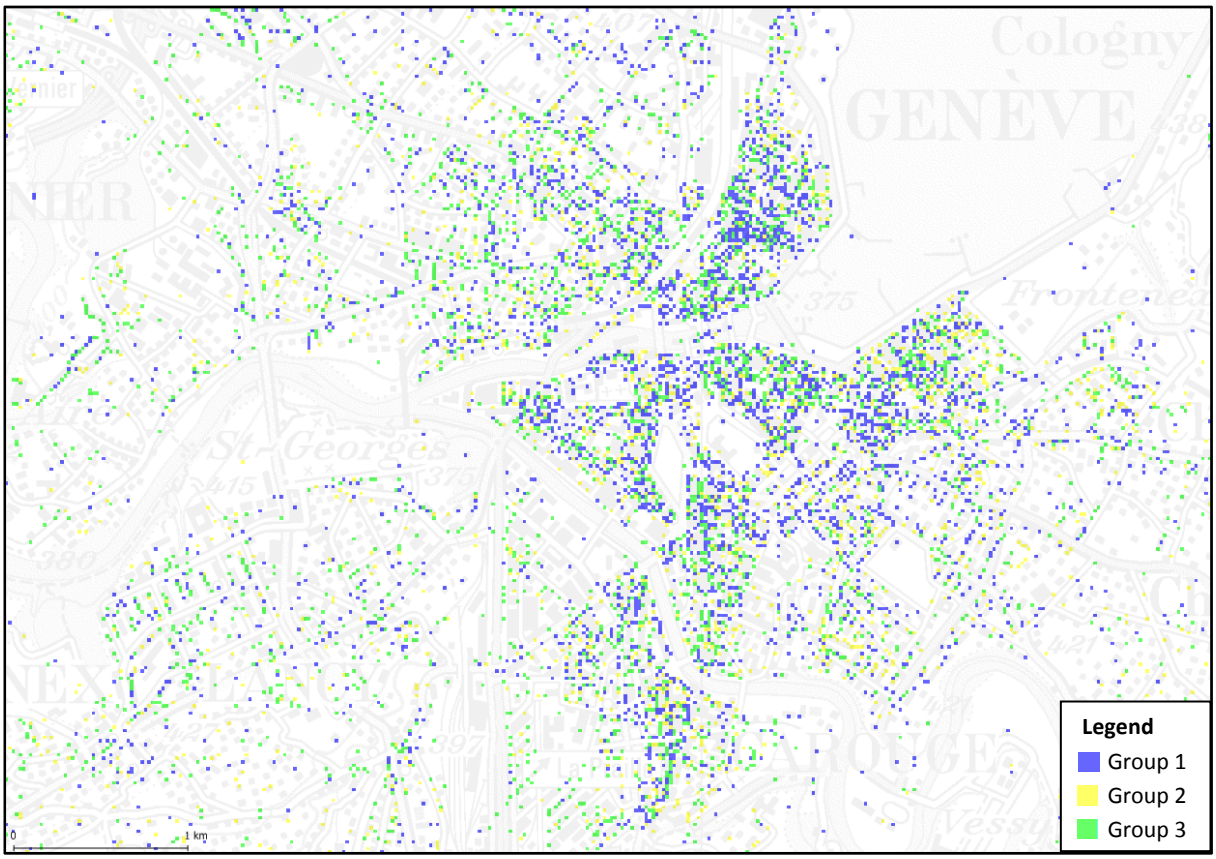
9.5.6 Groups for the subset A represented on 20m-grid



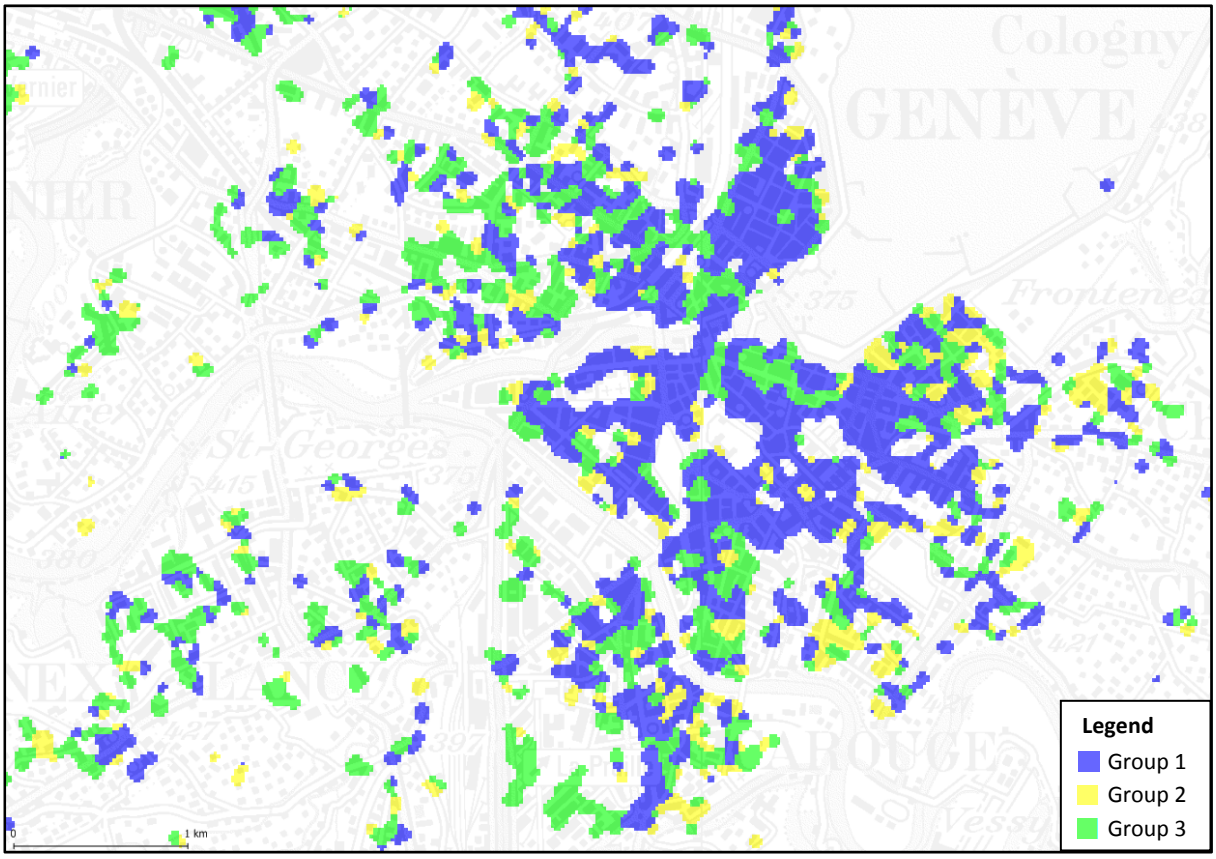
Filtered and simplified:



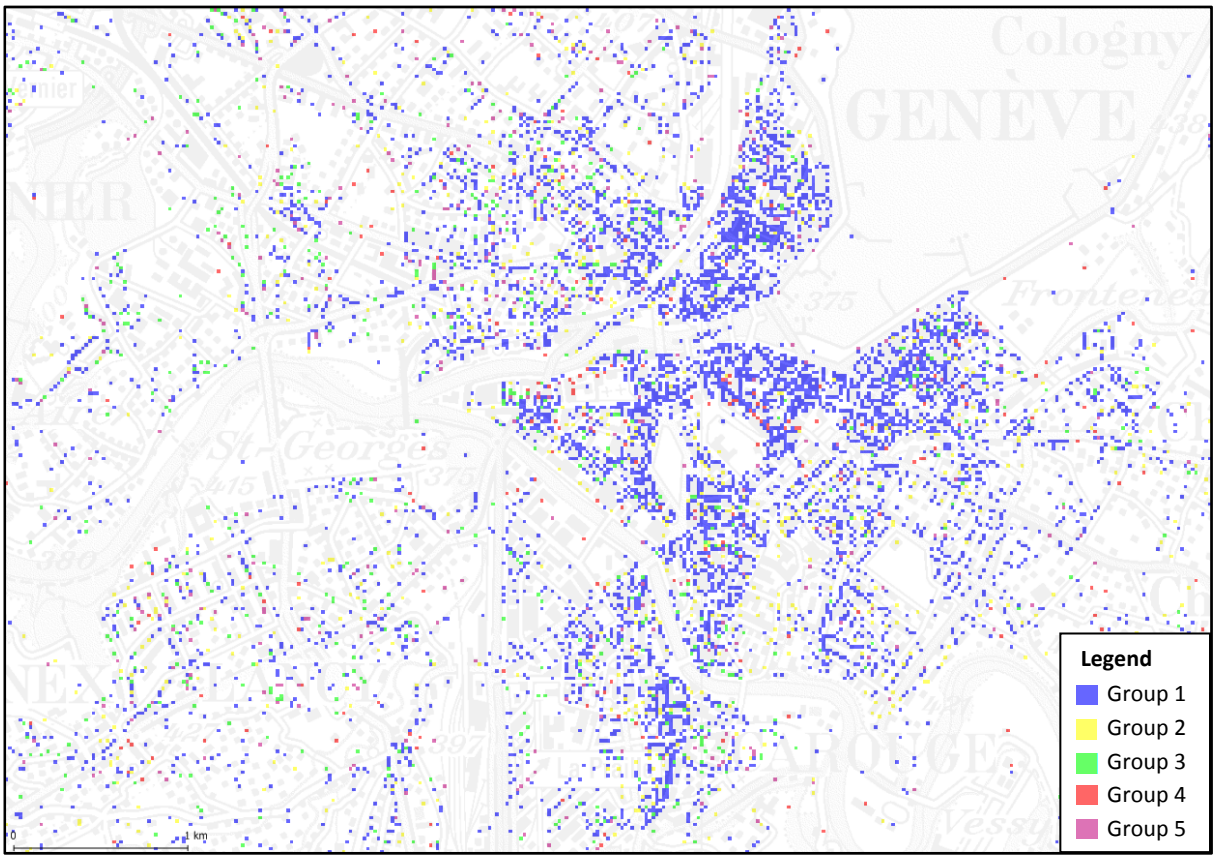
9.5.7 Groups for the subset B represented on 20m-grid



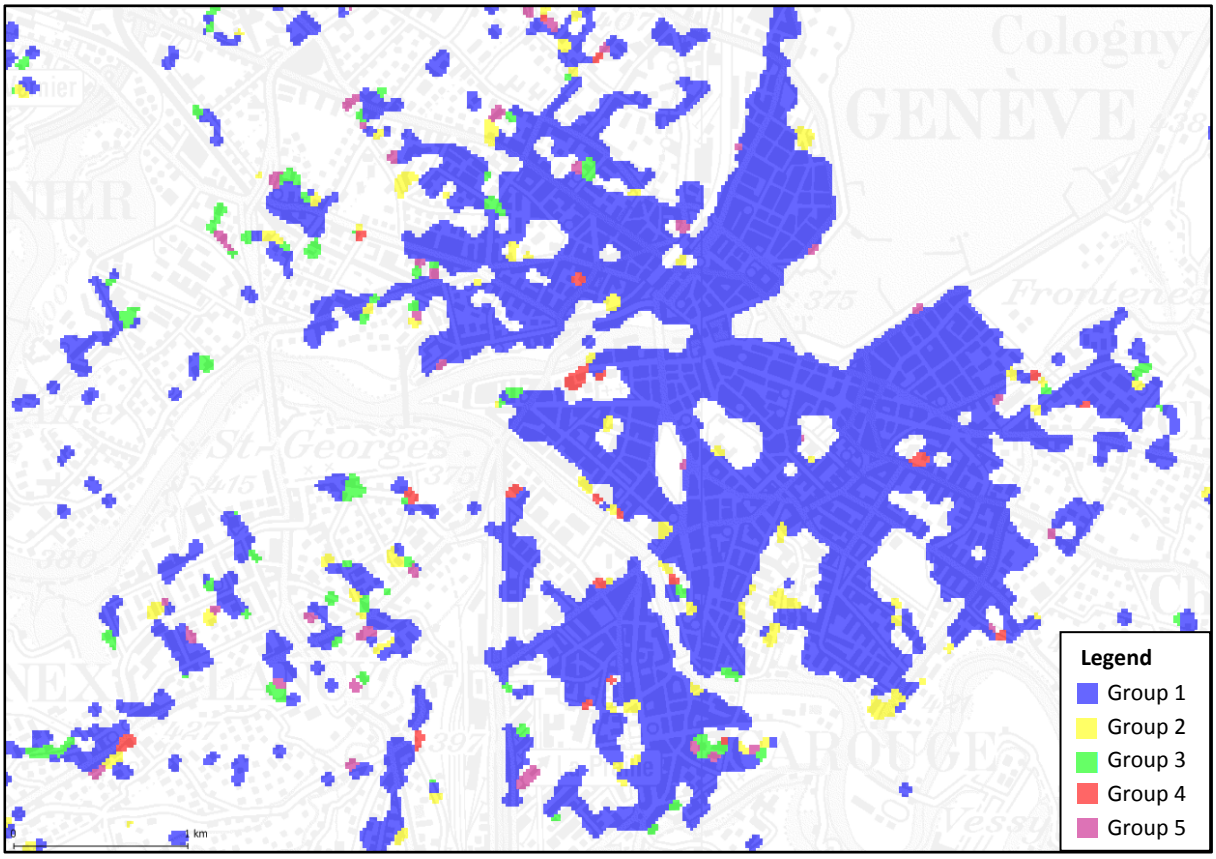
Filtered and simplified:



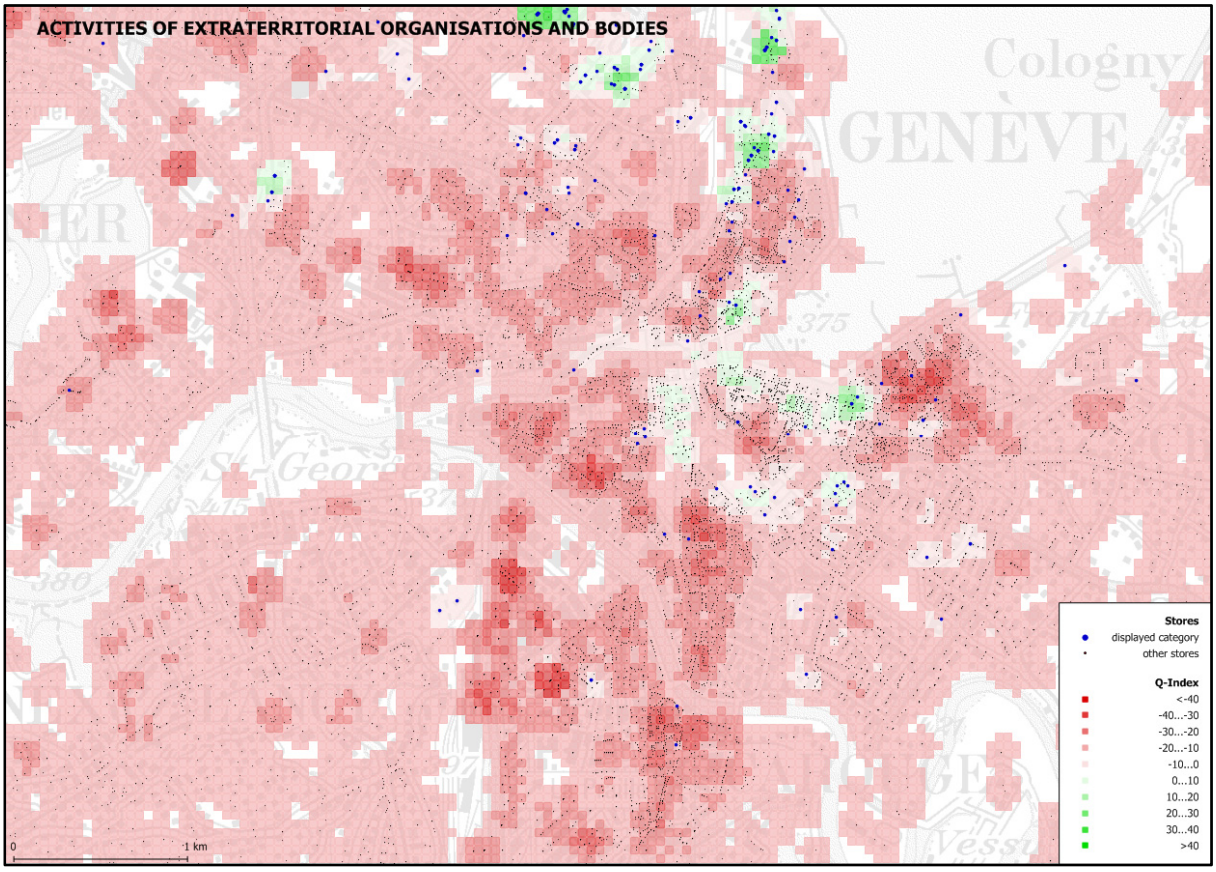
9.5.8 Groups for the subset C represented on 20m-grid

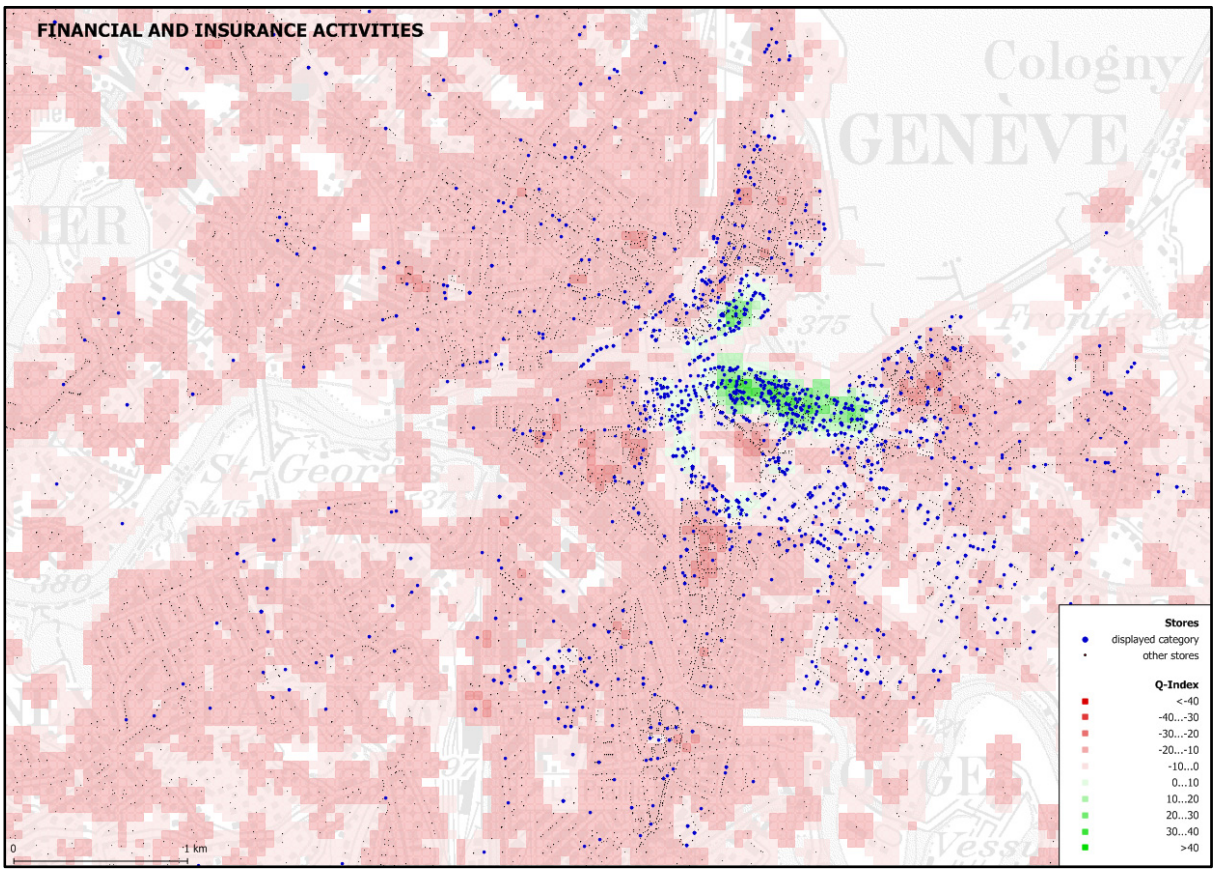


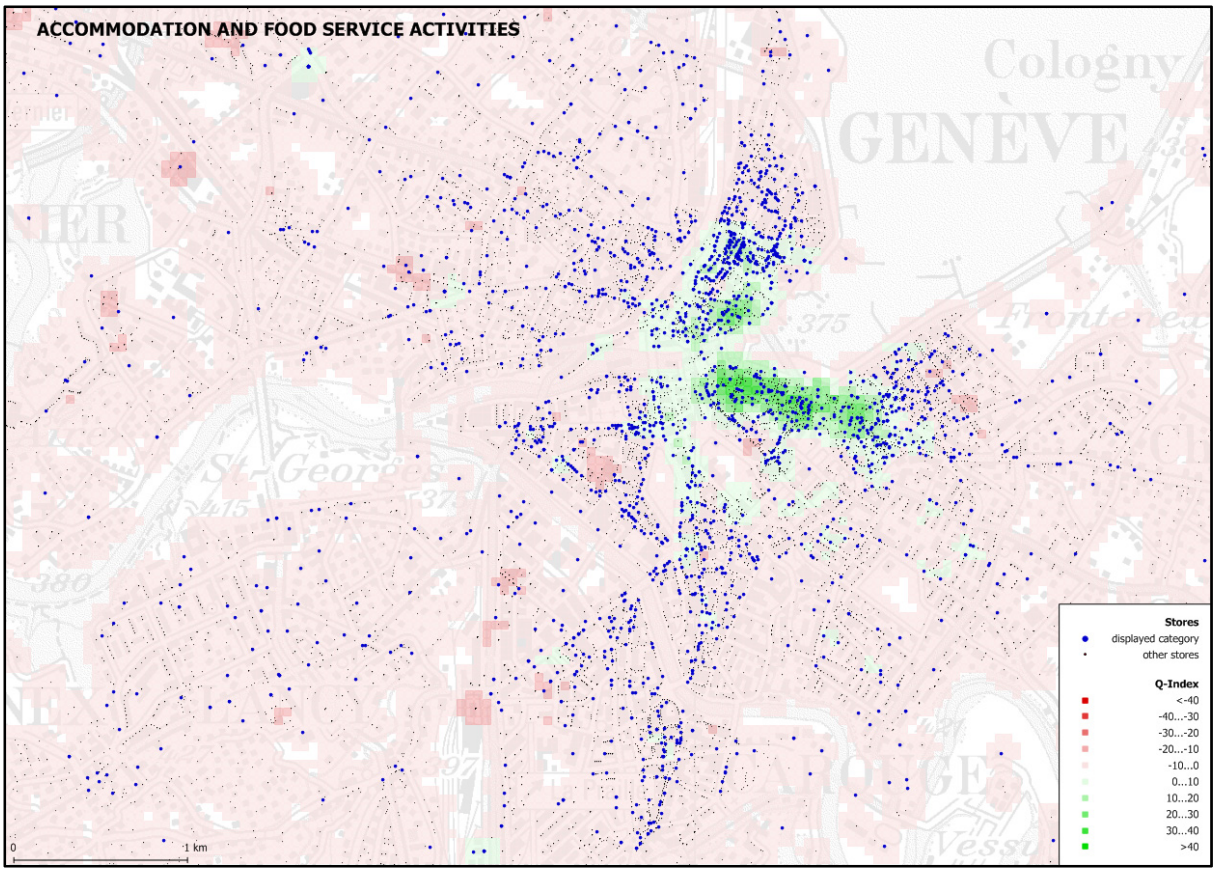
Filtered and simplified:

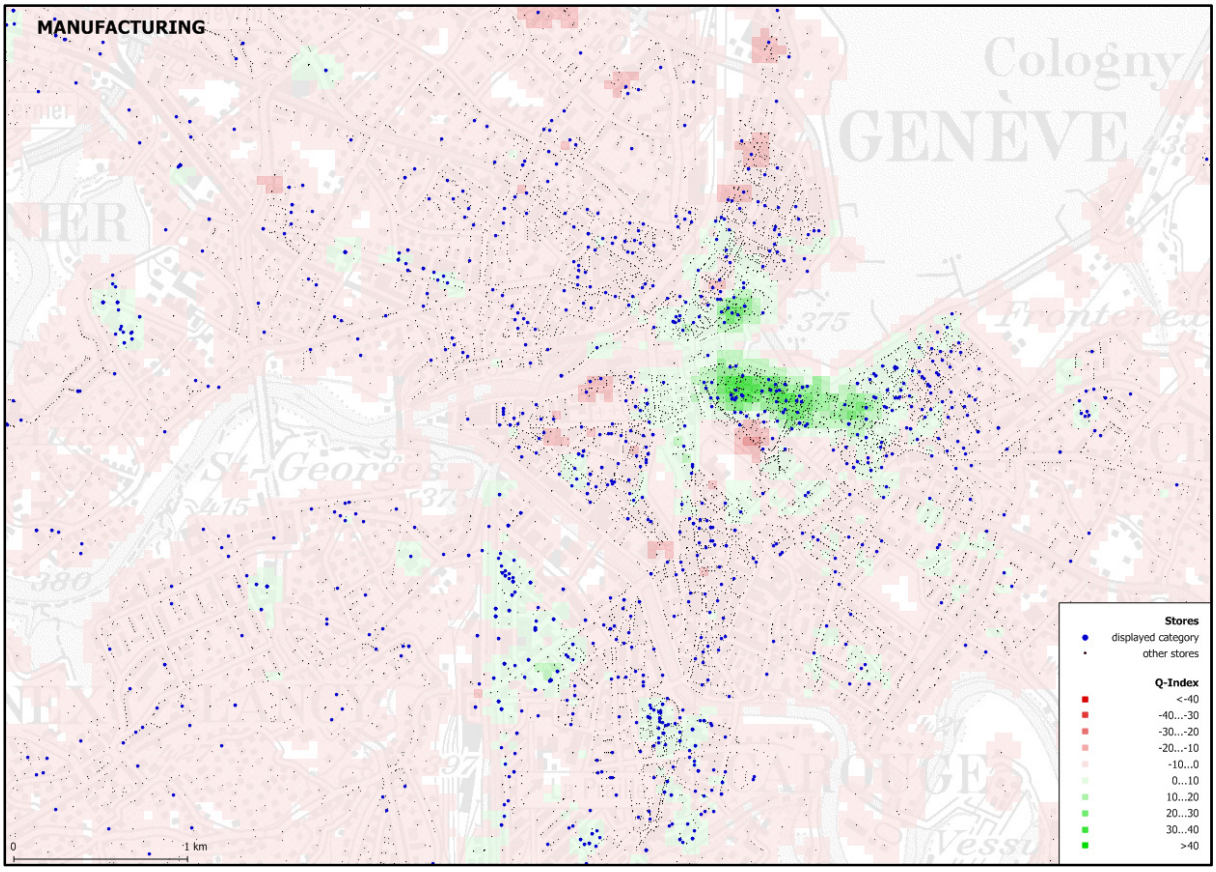


9.6 Q-Index-maps for ‘Geneva activities’






ACCOMMODATION AND FOOD SERVICE ACTIVITIES
Cologny
GENÈVE
Stores
displayed category
other stores
Q-Index
<-40
-40...-30
-30...-20
-20...-10
-10...0
0...10
10...20
20...30
30...40
>40
0
1 km

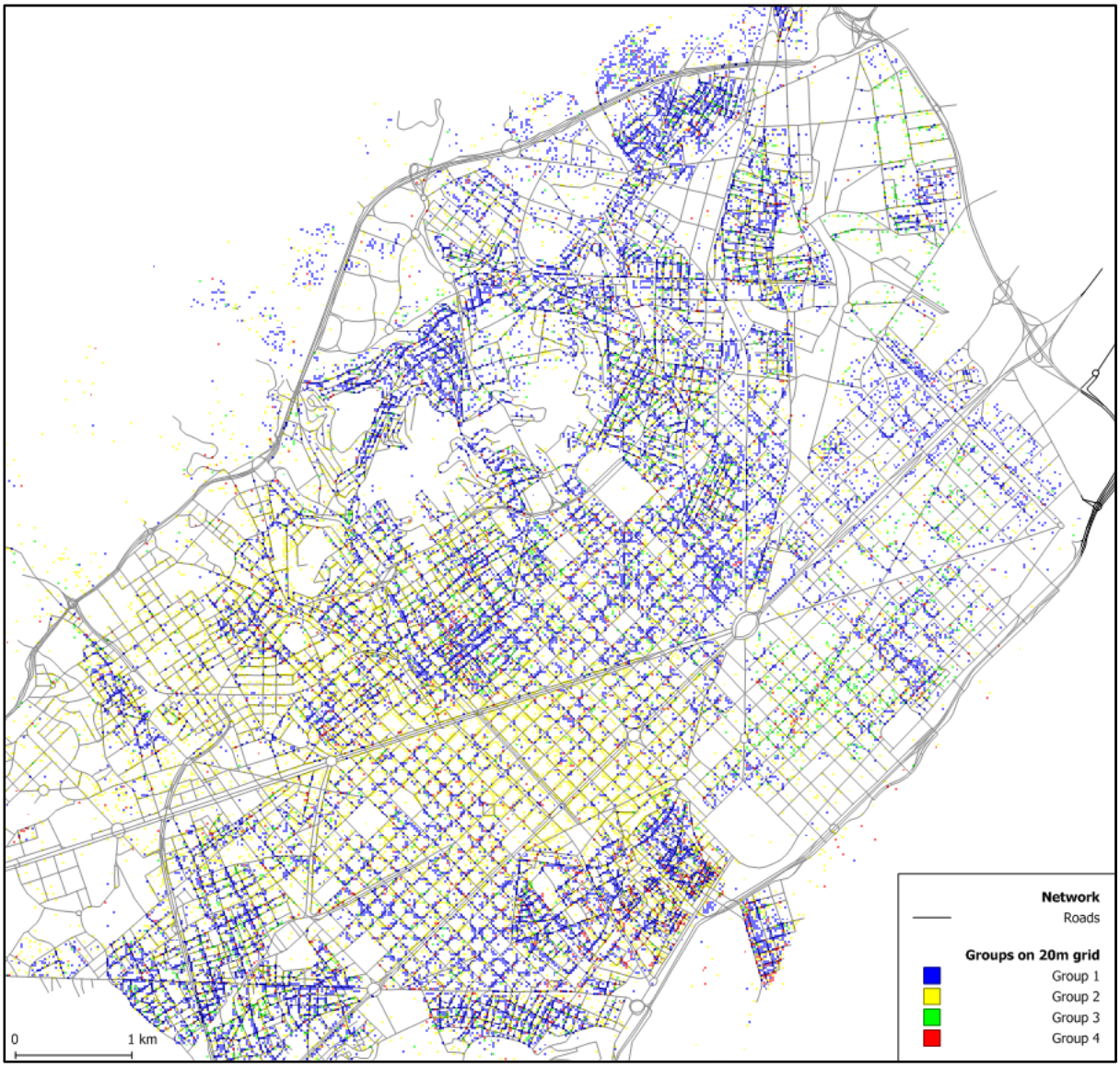

MANUFACTURING
Cologny
GENÈVE
Stores
displayed category
other stores
Q-Index
<-40
-40...-30
-30...-20
-20...-10
-10...0
0...10
10...20
20...30
30...40
>40
0
1 km

9.7 Classification results for Barcelona

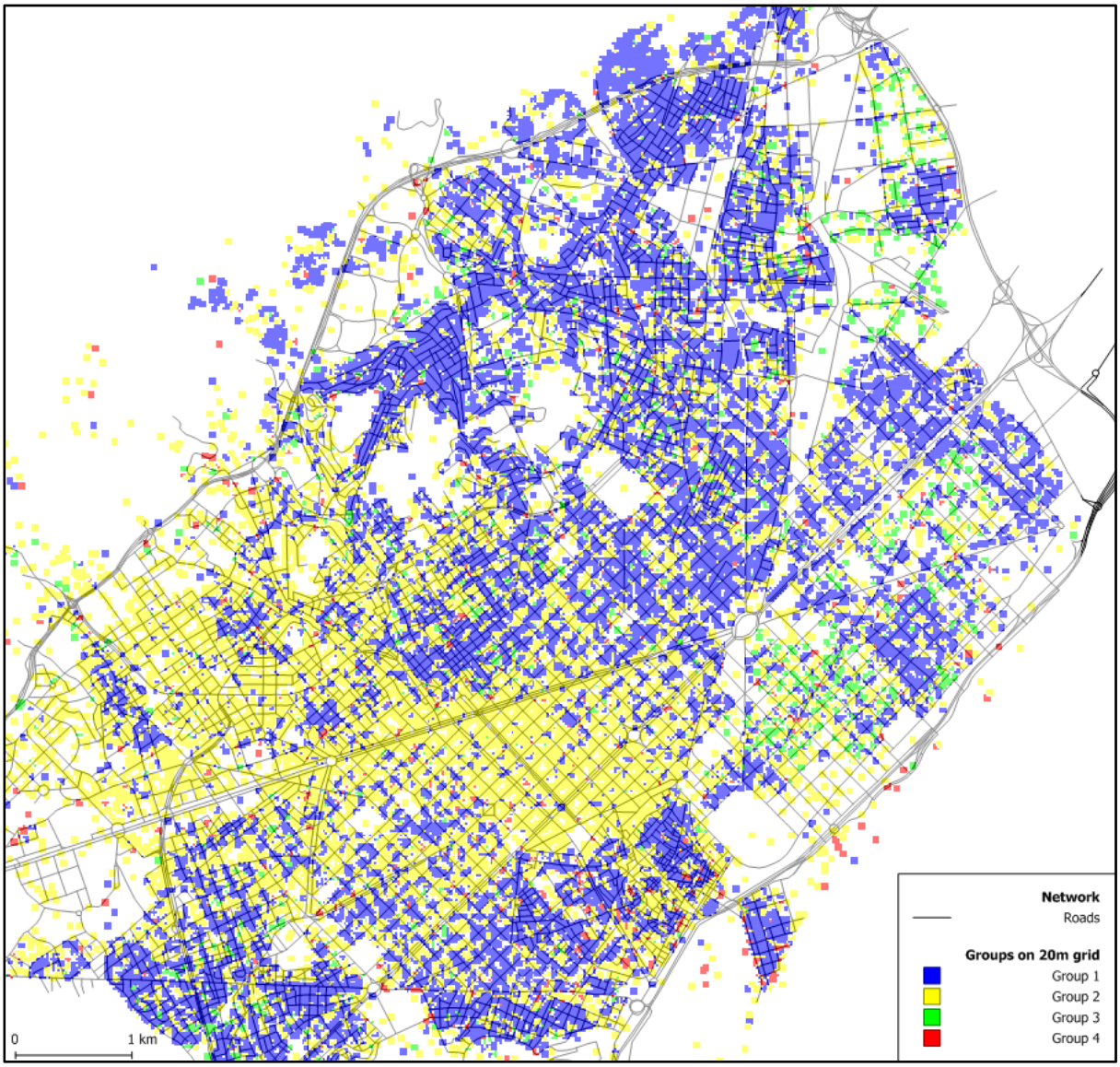
9.7.1 Groups for the set 'Barcelona-all'

Gr	quotK	CatKmax	#	Category	CategoryName
1	0.97	6.23	36310	G52	Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods
1	0.95	4.23	6492	O93	Other service activities
1	0.92	18.56	16097	I60	Land transport; transport via pipelines
1	0.51	6.12	161	D30	Manufacture of office machinery and computers
2	0.96	10.04	2961	I63	Supporting and auxiliary transport activities; activities of travel agencies
2	0.95	10.42	17189	K74	Other business activities
2	0.95	10.47	4598	J65	Financial intermediation, except insurance and pension funding
2	0.94	16.87	58	EE	Electricity, gas and water supply
2	0.93	19.05	2966	L75	Public administration and defence; compulsory social security
2	0.93	16.76	202	K73	Research and development
2	0.92	11.98	563	J67	Activities auxiliary to financial intermediation
2	0.92	29.81	87	Q99	Extra-territorial organizations and bodies
2	0.92	12.59	10343	K70	Real estate activities
2	0.91	9.64	4727	N85	Health and social work
2	0.90	6.98	4655	M80	Education
2	0.88	9.05	5721	O91	Activities of membership organizations n.e.c.
2	0.84	13.40	657	J66	Insurance and pension funding, except compulsory social security
2	0.83	28.74	72	I62	Air transport
2	0.83	7.82	515	O90	Sewage and refuse disposal, sanitation and similar activities
2	0.82	25.64	61	I61	Water transport
2	0.82	4.20	1110	K71	Renting of machinery and equipment without operator and of personal and household goods
2	0.78	5.11	866	I64	Post and telecommunications
2	0.74	8.65	9162	O92	Recreational, cultural and sporting activities
2	0.68	12.18	138	K72	Computer and related activities
2	0.59	6.02	12723	G51	Wholesale trade and commission trade, except of motor vehicles and motorcycles
2	0.52	16.90	79	D35	Manufacture of other transport equipment
2	0.37	10.02	115	P95	Activities of households as employers of domestic staff
3	0.99	31.80	1390	D28	Manufacture of fabricated metal products, except machinery and equipment
3	0.98	26.73	308	D25	Manufacture of rubber and plastic products
3	0.98	21.91	138	D32	Manufacture of radio, television and communication equipment and apparatus
3	0.98	15.03	483	D33	Manufacture of medical, precision and optical instruments, watches and clocks
3	0.98	20.57	969	D36	Manufacture of furniture; manufacturing n.e.c.
3	0.97	17.32	451	D20	Manufacture of wood and of products of wood and cork, except furniture
3	0.97	24.84	246	D31	Manufacture of electrical machinery and apparatus n.e.c.
3	0.97	24.02	58	D27	Manufacture of basic metals
3	0.95	25.54	310	D29	Manufacture of machinery and equipment n.e.c.
3	0.92	18.11	163	D21	Manufacture of pulp, paper and paper products
3	0.91	19.05	160	D26	Manufacture of other non-metallic mineral products
3	0.91	6.15	3375	G50	Sale, maintenance and repair of motor vehicles and motorcycles; retail sale of automotive fuel
3	0.90	18.85	103	D19	Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear
3	0.89	8.52	479	D15	Manufacture of food products and beverages
3	0.88	22.72	80	D34	Manufacture of motor vehicles, trailers and semi-trailers
3	0.87	6.13	1640	D18	Manufacture of wearing apparel; dressing and dyeing of fur
3	0.84	17.70	275	D17	Manufacture of textiles
3	0.73	10.45	86	CC	Mining and quarrying
3	0.72	12.60	344	D24	Manufacture of chemicals and chemical products
3	0.21	8.72	3802	D22	Publishing, printing and reproduction of recorded media
4	0.75	5.77	12758	H55	Hotels and restaurants

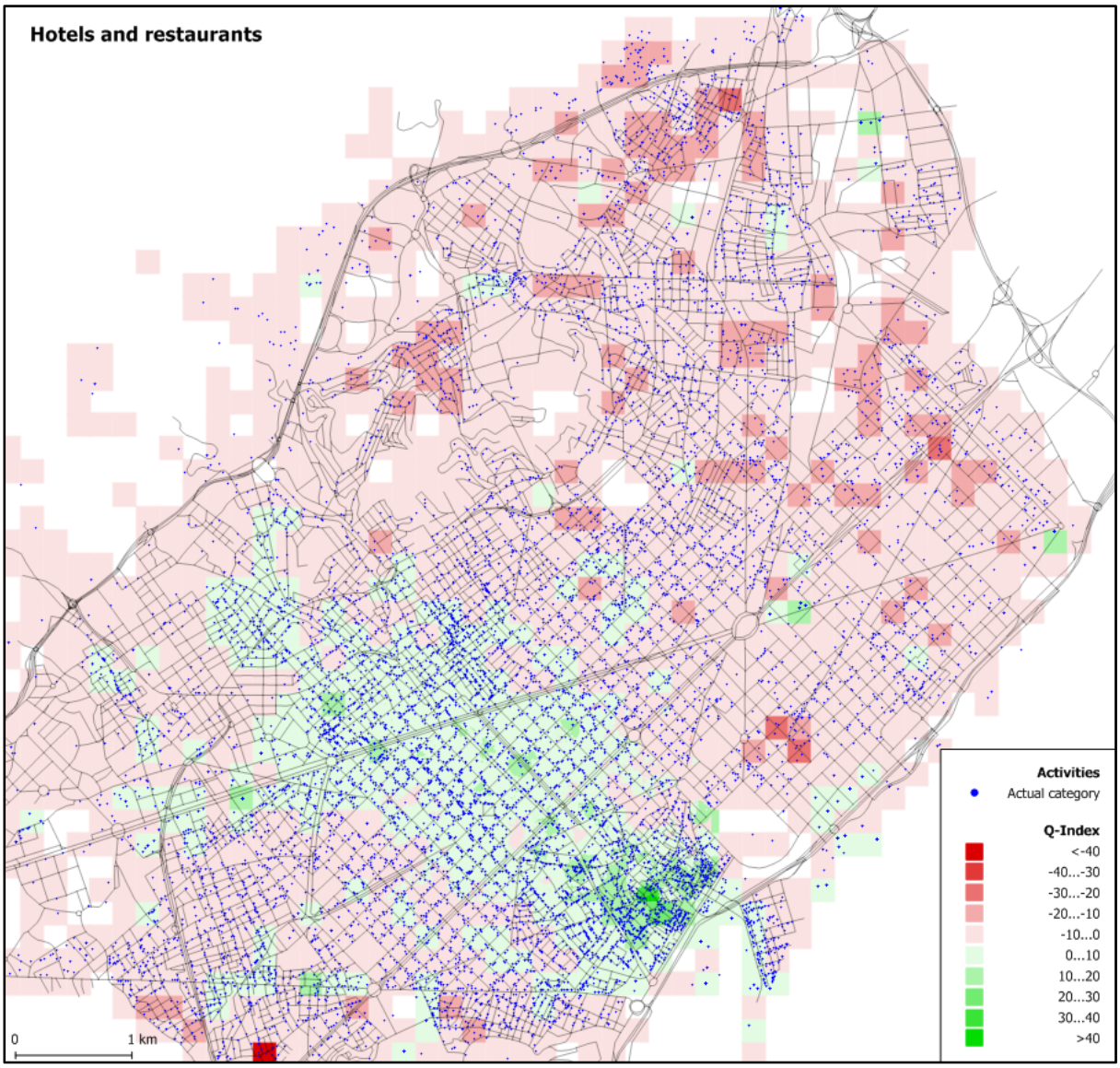
9.7.2 Groups for the set ‘Barcelona-all’ on 20m-grid

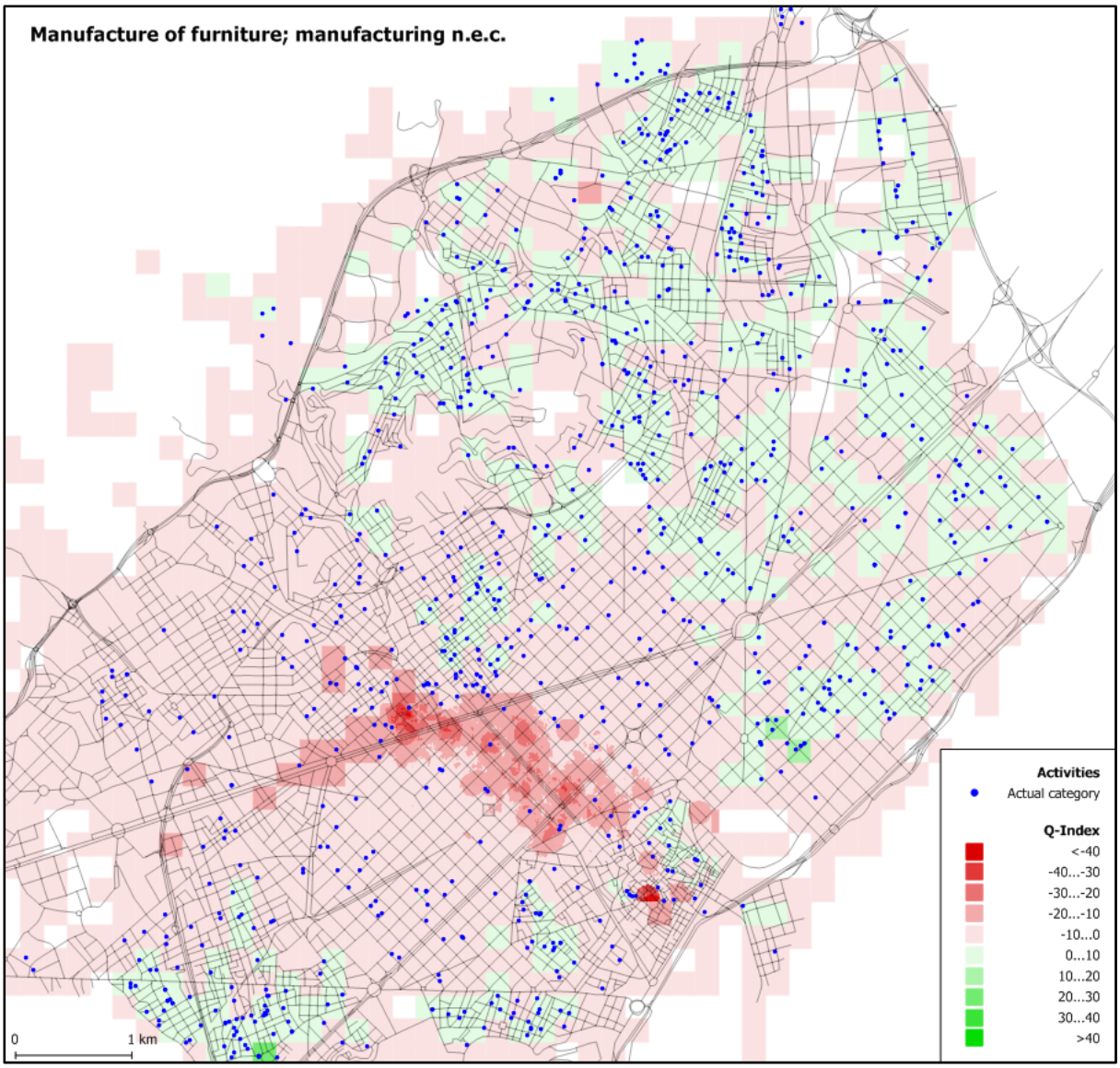


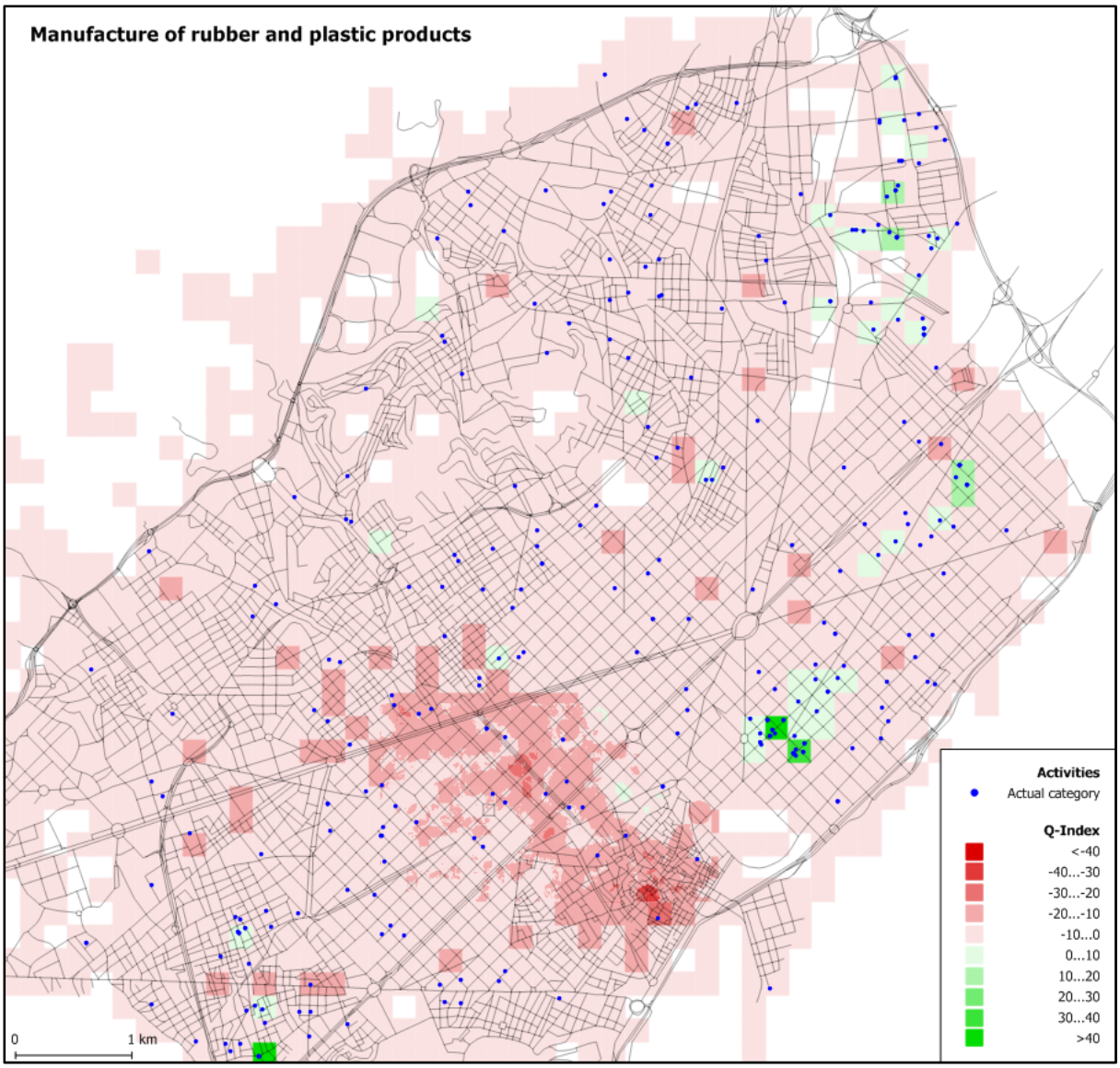
9.7.3 Groups for the set ‘Barcelona-all’ on 20m-grid – filtered



9.8 Q-Index maps for Barcelona

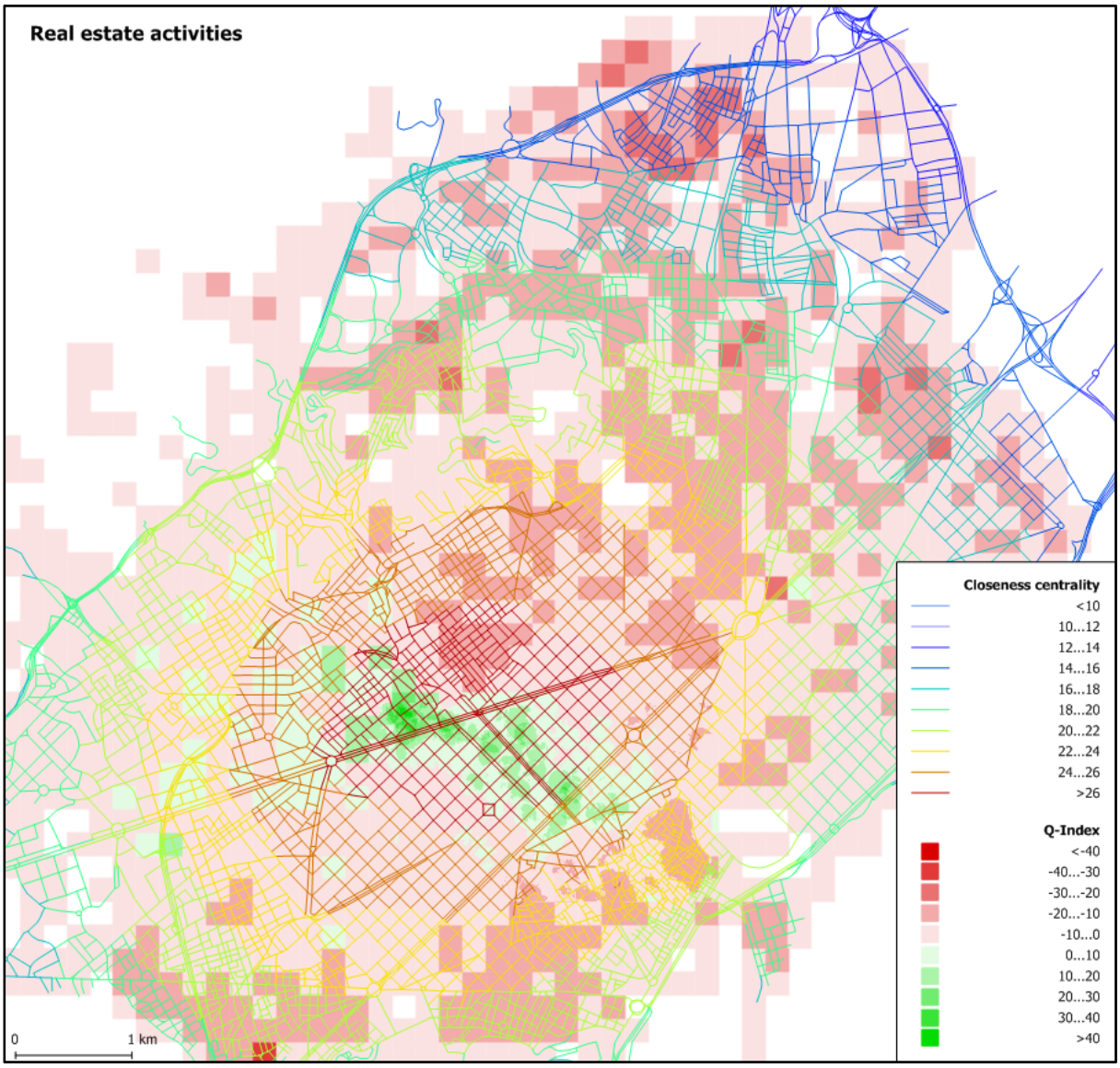



Manufacture of furniture; manufacturing n.e.c.
Activities
Actual category
Q-Index
<-40
-40...-30
-30...-20
-20...-10
-10...0
0...10
10...20
20...30
30...40
>40
0
1 km


Manufacture of rubber and plastic products
Activities
Actual category
Q-Index
<-40
-40...-30
-30...-20
-20...-10
-10...0
0...10
10...20
20...30
30...40
>40
0
1 km

9.9 Comparison centrality – Q-Index

9.9.1 Map with global closeness centrality and Q-Index for real estate activities



9.10 Program code and data

The program code, the data and an electronic version of this report are on a CD delivered with this report.