Neighbourhood deconstructed: developing and contextualising methods to understand the neighbourhoods in London.

Ivann Schlosser

Supervised by Prof. Elsa Arcaute and Mateo Neira

-

A dissertation submitted in partial fulfilment

of the requirements for the degree of

Master of Science

of

University College London

-

Word Count: 9473

Module: **CASA0010**

Online: https://github.com/ischlo/Dissertation_analysis

Centre for Advanced Spatial Analysis

University College London

September 13, 2021

Abstract

”neighbourhood (neighborhood) - the area of a town that surrounds someone’s home, or the people who live in this area.”[1] is the definition of the neighbourhood in the Cambridge dictionary. It highlights both the proximity to home and the community aspects of this area in a city. At the same time, throughout academic and planning literature, there isn’t a consensus on what is an adequate definition of neighbourhoods, leading to this space in the city to be defined according to the context of the research and the limitations of the data. This work looks at this area through a local diversity index to better understand the livability of neighbourhoods as areas around home. It is anchored in the modern planning paradigm that highlights the importance of walkable access to a wide range of opportunities surrounding ones home. It verifies the mix of land uses to be an indicator of attractiveness of a place for people, as originally formulated by Jane Jacobs and widely adopted in the modern urban design practice. Finally, it investigates with an original methodology how the activity spaces of local residents overlap in proximity to their home, allowing to build a hierarchical subdivision of space in the city based on the visiting patterns, illustrating a community vision of neighbourhoods.

[1] https://dictionary.cambridge.org/fr/dictionnaire/anglais/neighbourhood

I, Ivann Schlosser, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work. It is 9473 words in length.

Acknowledgements

I would like to thank professor Elsa Arcaute for her contribution throughout this time, her inspiration and rigour in research, Mateo Neira for his sharp comments, literature suggestions, and methodology advice, Alexei Poliakov, for providing the data set, technical support and advice. Thank you to my family and friends for their support. I would also like to thank the CASA community for simply being the way it is. Finally, this work pays tribute to London, and expresses my gratitude for the opportunities and encounters that it has given me.

Contents

1 Introduction **10**

1.1 Research question and system of interest 12

1.2 Ethical considerations 12

2 Literature review **13**

2.1 Context 13

2.2 The planners role. 15

2.3 Defining substructures within the city 17

2.4 The spatial layer 17

2.5 The social layer 18

2.6 Human mobility 20

2.7 Investigate locations through users data 21

2.8 Final note 22

3 Data **23**

3.1 OpenStreetMap 23

3.2 Amenities of interest 23

3.3 Mobility data 23

3.4 Spatial partition 25

4 Neighbourhoods through amenities **26**

4.1 Methods 26

4.2 Interpretation 27

4.3 Diversity vs Quantity 29

4.4 Defining neighbourhoods 31

5 Neighbourhoods through mobility **36**

5.1 Methods 36

5.2 Identifying neighbourhoods 39

6 Discussion **42**

6.1 Diversity of visited places 42

6.2 Community through hierarchical percolation 43

6.3 Context 44

6.4 Overlap of the definitions 45

6.5 Limitations and improvements 46

7 Conclusion **46**

7.1 Reproducibility 47

A Appendix **57**

A.1 Selection of amenities 57

A.2 Clustering 58

A.3 Entropy 58

A.4 Online supplementary material 59

A.5 The grid 60

List of Figures

3.1 Overview of the data on users and their GPS traces. On the bottom plots, each colour is for a user, 3 are shown as an example. 24

4.1 Entropy index compared with density. Density is intended as the number of amenities within 7 minutes of walking distance. 28

4.2 The density of amenities in London 30

4.3 The entropy of amenities in London 31

4.4 The famous neighbourhoods found. Blue lines are the main water bodies of the city. Orange colour corresponds to high entropy amenities (S $>$ 3.1). 32

4.5 Location of the places with the highest density of amenities. The blue area corresponds to densities of less than 400 amenities within 7 walking minutes, and the yellow area are from 400 up. 33

4.6 Example of the clustering algorithm linking an amenity to it's nearest local maximum. 34

4.7 A schematic representation of the neighbourhoods found through the investigation of the AES. Each distinct coloured polygon represents a different neighbourhood. 35

5.1 The neighbourhoods based on similarity of visitors for a partition into 380 components of the hierarchical tree. 40

5.2 Top plot: size of the biggest components with respect to the partition. Vertical lines represent examined cases. The black line is for the kept threshold of 380. Bottom plot: median size of the typical distance from the distribution for each partition. 41

6.1 The visitors recorded at locations with amenities plotted against the diversity. The number of visitors is log scaled. The mobility taken into account is on the neighbourhood level, within $r = 3$km from each users home. The dark red line on 6.1a shows an exponential relation between the two variables. On 6.1b, the colour scale corresponds to the number of observation. Orange lines are the median for each quantity. . 43

6.2 On 6.2a is mapped the measured values of entropy after being transferred on the hexagonal grid. Each cell of the grid is given the mean values of the entropy index of the amenities present in it. Places without amenities are plotted with a value of -1 in the map. On figure 6.2b one can see the number of visitors observed in each cell of the grid. . . 44

A.1 Clusters of amenities in selected boroughs 58

A.2 Alternative visualisation of the neighbourhoods from amenities by size of each cluster. 59

A.3 Entropy index of amenities for selected boroughs. 60

A.4 61

A.5 Illustration of the hexagonal grid with non residential cells merged to the nearest residential one. 61

A.6 Percolation on mobility 62

List of Tables

A.1 Summary table of amenities of London according to OSM map feature tags. 57

1 Introduction

Nowadays, cities are facing great challenges, with urban population increasing more than ever and passing by the symbolic mark of 50% of the total population worldwide[1] (UN, 2019). This has direct consequences on how cities should organise themselves, in terms of food supply, public services such as transport, safety, health. While responding to those challenges, cities need to fulfil their original and main objectives of providing places of encounter, socialisation, economic, scientific and cultural growth as well as optimal conditions for well-being and happiness of their inhabitants (Glaeser 2011).

Arguably, the main challenge that humanity faces today is the climate emergency, and cities as the main emitters of pollution (Moran et al. 2018) need to be on the forefront of adapting infrastructure and lifestyles to be more sustainable. Being able to respond to those challenges means a great deal of collaboration between all the actors of the built environment industry, and the role of academics should also be emphasised as central to the study and understanding, but also communication of valuable insights.

In the modern urban sciences, cities are being studied as complex systems. This methodology, taken from math and physics, allows to effectively deconstruct cities into pieces that, despite the apparent complexity, allow a high level of accuracy of models and quantitative descriptions to be made. The tools adapted from statistical physics allow to have an accurate macroscopic understanding of the system as described by Barthelemy (2019). This work pays attention to the neighbourhood scale of the city, and the ambition is to combine existing and developed methods to understand this place from the user, citizen perspective and as a possible unit to act upon when it comes to urban design.

Bridging the urban planning and city science literature, a definition of neighbourhoods is inferred. Mostly relying on the work of Jane Jacobs, who proposed that neighbourhoods should be places of encounters, containing a diverse mix of land uses, be walkable in size in order to serve as the foundation of strong and vibrant communities. Since her work in the early 60s, a new planning paradigm has been formulated

[1] https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS

and referred to as New Urbanism (Laurence 2006).

It puts personal well-being, inclusiveness, sustainability at the centre of the debate and actions. The recently published London plan [2] highlights the importance of providing locally a healthy, safe environment, with diverse land uses, and a walkable access to amenities. At the same time, it highlights the intention to drastically reduce car use and redevelop the local environment accordingly, in a way giving back the streets to the people. Neighbourhoods could greatly benefit from all these modifications and this work provides some tools that allow to better understand the local structure of the city. First, from the planning perspective and then, through the study of local individual mobility, from which various communities are identified. Moreover, the recent pandemic has illustrated how the quality of the environment surrounding our home is important to our well-being. Hence, a better understanding of the neighbourhoods in a city can be an important step towards a better handling of such events.

[2] The London plan website.

1.1 Research question and system of interest

The questions that are aimed to be answered in this work are directly related to the identification of neighbourhoods from the perspective of planning, and from the perspective of individual mobility. First, by looking at amenities, their spatial distribution and diversity. The importance of amenities as social attributes in space that attract people and influence the creation of a lively, vibrant environment is discussed in the following sections. Second, this work is dedicated to inferring neighbourhood boundaries from a data set of human mobility, spanning 2 weeks between the 13th and 27th of February 2020. The developed methodology, consisting in a hierarchical percolation of the similarity of users in proximity to their home , allows to visualise the neighbourhoods from a community perspective at various scales. The appropriate scale, consistent with considerations from the previous section is proposed and discussed. The main research questions addressed in this work are formulated in the following way:

How to identify neighbourhoods, as places in space, in the context of modern urban planning ?

How to identify neighbourhoods from a community perspective, based on individual mobility patterns around home ?

How do these definitions relate and how can they be useful ?

1.2 Ethical considerations

The data on mobility has, in its raw format, a very high spatial and time granularity, leading to risks of identifying individuals. To remove all possible risks of misuse and to meet the ethical standards set out for academic research in the UK, it has been properly aggregated by the provider considering the regulations in place. It has been transferred, stored and processed through secure networks and devices as required by the code of ethics of the British Educational Research Association (BERA). The data set on amenities comes from the open source mapping platform OpenStreetMap and does not present any ethical issues. The data sets are introduced further in section 3.

2 Literature review

2.1 Context

In the science of cities of the 1st half of the 20th century, sociologists and planners where proposing various models, based on observations, to understand how cities are structured. They were proposing a mono-centric vision that placed in the middle the central business district (CBD), the main attractor of workforce and centre of commerce, optimising the cost of bringing the product to the consumer. Surrounding it were the residential parts of the city that would see the social class and income of their inhabitants decrease as we get further from the CBD and closer to the heavy industry cites (Alonso 1964). This model has since evolved and nowadays, digital tools and the availability of a great amount of data gathered allow to investigate and understand the various phenomena that shape our cities in greater details.

The structure of a city is the result of complicated processes that happen over space and time, and researchers have been able to identify some of the elements that are the most resistant to time, those are roads (Barthelemy 2017; Masucci, Stanilov, and Batty 2014). While buildings and their use have consistently changed and evolved other time, roads have been shown to be the most time-proof element. In fact, the growth of roads in the city has been well modelled using statistical methods, with the highest scales and centrality scores being linked to older roads, and smaller segments, to more recent additions to the road network as found by Strano et al. (2012) on the example of Milan.

The investigation of scaling laws (Clauset, Shalizi, and Newman 2009) linking economic growth, wages, cultural richness, attraction of capitals to city size has shown all the positive aspects of agglomeration in cities for their residents. These results indicate that the whole represents more than the sum of its parts when it comes to human growth in the city, especially R&D, innovation as found out by Arcaute, Hatna, et al. (2015) in England and Wales and by Bettencourt et al. (2007) in the US. At the same time, the amount of infrastructure needed to meet peoples needs has been shown to increase sub-linearly with city size (Arcaute, Hatna, et al. 2015), proving the efficiency of cities in terms of resource spending. Not only positive things, however,

result from this agglomeration of individuals as more violent crimes tend also to occur in bigger cities.

Further investigation into the distribution of businesses within the city has allowed to understand how economic activity is spread out in London on a more granular level and this time with a spatial component (Cottineau and Arcaute 2020) and on a greater scale in the case of french urban agglomerations (Cottineau, Finance, et al. 2019). In the former case, it has allowed building profiles of neighbourhoods based on the addresses of businesses registered there and outline how major changes occur in the economic activity of startups over the recent years in different neighbourhoods of London, illustrating at the same time the polycentric transition of activity clusters in the city.

With the scaling of the size of a city increases the effective cost of commute, measured either in time, distance or financial cost. Congestion that results from the increasing number of commuters in the monocentric city has made this model irrelevant in the context of contemporary cities. Recent research investigating scaling behaviour of movement indicates that the big megalopolises of the world are already showing a polycentric structure (Louf and Barthelemy 2015). The quest for the optimum between transport cost and access to the market has led people looking for alternatives to the CBD, in greater proximity to their residence. Research by Louf and Barthelemy (2013), shows how from a hypothesis of balancing attractiveness of a centre and transport cost, the polycentric transition of cities naturally emerges. The hierarchy and spatial distribution of places in cities is now more complex than ever, with new additions being made to the transport network, new developments being built and more people moving in. The multiple and intricate phenomena investigated throughout the literature presented until now and further on can provide valuable insights on human made structures not only within the field of academic research, but also help inform decisions and strategies to act on the system of interest, the city, with better informed actions.

2.2 The planners role.

To be able to respond to peoples needs, planners have drastically changed their methods over the recent decades. The new planning paradigm goes quite strongly against what has been extensively adopted in the mid 20th century, when the car-dependent user commuting to the city centre through a big highway from a suburban settlement was one of the main scenarios considered by planners. The introduction of new principles and methods, have led to the so called New Urbanism (Laurence 2006) that places individual well-being and sustainability at the centre of the debate (Iravani and Rao 2020). This paradigm puts forward a vision for the city and the neighbourhood as a place where all the necessary is easily accessible, within walking distance from home, reducing car dependence, which in turn, provides a healthier, greener, prettier environment in the neighbourhood. The various downsides that emerge from a car centric environment, such as pollution (Moran et al. 2018), less greenery, urban heat from the roads [1] to name a few are now being addressed. We are witnessing important changes in the priorities given to transport mode in cities, in London with the Ultra Low Emission Zone (ULEZ), and in Paris, with an almost total conversion of the centre (intra muros) to active travel and public transport by 2024. Recently, the city of London has selected the new redevelopment plan for Oxford Circus, a huge transit point of vehicles and people, to be completely dedicated to pedestrians [2]. In Lausanne, the tram line crossing the city was deprecated to dedicate more space in the streets for cars, it is now being rebuilt and the network extended further from it's initial route [3], with the original demolition of the network being acknowledged as an error by city officials. In the US, local high streets that during the pandemic were fully given to pedestrians in order to facilitate their movement are now being voted to stay this way by locals and officials [4]. The distribution of amenities and their diversity naturally arises in this context as a mean to provide all the necessary places in the environment to fulfil the needs of a diverse group of individuals in proximity to

[1] Urban heat in the US
[2] https://osd.london/project/oxford-circus/
[3] https://www.tramway-lausannois.ch
[4] Can 'Open Streets' Outlast the Pandemic?

their residence and wherever they go. The wider use of participatory and inclusive planning has given the voice[5] to people that were silenced before (Listerborn 2007). They were able to formulate and express their needs. This has shed light on the actual variety of uses that the urban environment is subject to, pushing planners to build a more inclusive, safe, accessible environment for everyone. This sort of interventions, that are now widely adopted in Europe, require a high level of understanding of local specifics and occur on a much smaller spatial scale that the interventions of the mid 19th and 20th century. We are unlikely to see a project of the scale of baron Haussmann's reconstruction of Paris in any of the European big cities in the near future. Nowadays, urban design goes more in line with the propositions of Jane Jacobs. In her book*The death and life of great American cities*,(Jacobs 1962), she puts at the heart of lively, flourishing cities small walkable blocks in which there would be a diversity of land uses and amenities that provide the opportunity to form communities, socialise and accommodate a variety of lifestyles. The direct impact of the built environment on health has also wide and growing attention and some occurrences of illnesses, mental or physical, have been linked to the environment in which individuals live. Some very comprehensive research done on the north American continent have found links between the diversity of available transport modes and the obesity and diabetes rates (Glazier et al. 2014). Individuals living in areas segregated from the point of view of transport and relying mostly on cars to move around are more likely to develop diabetes and be obese (Creatore et al. 2016).

On the other hand, those living in areas where a diversity of modes of transport is available and the the environment is active travel friendly, where showing a better health condition on average. The widespread use of cars has also proven to be not only harmful for nature and the driver's health, but also to people around. By studying pollution from vehicles in Geneva, Fleury et al. 2021 have found similarities between pollution clusters and Parkinson disease occurrences of people living within them.

This kind of research leads to a understanding of the importance that good planning can have on the health and well being of individuals. While the propositions exposed

[5] The queer city: how to design more inclusive public space

in her book relied mostly on her own impressions and experiences of the city and were not verified at the time, urban science seems to be taking the side of Jane Jacobs against Robert Moses and Le Corbusier in what constitutes a vibrant environment to live in according to research done by Sung, S. Lee, and Cheon (2015),De Nadai et al. (2016) on Seoul and a set of Italian cities. In the same time most of the work of the later two is now considered as the remains of a peculiar period in urban planning and architecture that the actors of the industry are trying to move away from.

2.3 Defining substructures within the city

There seems to be a certain acceptance of what are the boundaries of a city, whether administrative, or perceptive, these boundaries are sometimes intentionally maintained, like the greenbelt in London. Research by Arcaute, Molinero, et al. (2016) using percolation on the length of street segments and building a hierarchy of road clusters came independently to city boundaries that are similar to commonly observed ones and administratively recorded.

There is, however, much less agreement on what are the different parts that the city is made of and what are their boundaries. This work aims to combine certain criteria in order to define a substructure of the city, the neighbourhood. The motivation comes from the importance of this place for the citizen, but also for the planner, developer, city mayor, especially in the New Urbanism paradigm. The local understanding and ability to act on the right parameters of the system can lead to better global outcomes, whether in terms of well being of individuals, safety, sustainability, or even disease spread. The following sections overview how, on a smaller scale, the city has been investigated from the pure perspective of space 2.4, from the perspective of the social attributes 2.5.

2.4 The spatial layer

When taking a closer look at the cities spatial organisation, the main literature is divided into various approaches, for example looking at locations of buildings simplified into polygons in 2D space, or looking at streets and intersections, modelled as a spatial network (Barthélemy 2011). When it comes to neighbourhoods, some re-

search uses distance between buildings to differentiate substructures of the city and define neighbourhoods at a certain «human scale». In this case Aras, Ouellette, and Jain (2021) produced, for various thresholds, partitions of 3 big American cities into neighbourhoods, revealing different patterns of distances across the cities as well as the natural or man-made barriers that sometimes impose a boundary on an area. They highlight the role of wide, car oriented streets and highways, in the formation of boundaries that could increase exclusion and segregation. The resulting heterogeneous layout bears links to the way the environment was shaped by master plans or local interventions and throughout the history of the city.

By investigating the shape factor of the blocks that are formed by closed loops of streets in the network, Louf and Barthelemy (2014) find some spatial patterns that allow to separate different parts of New York into neighbourhoods. This methodology allows as well to differentiate between the types of urban layout, with planned neighbourhoods showing a more regular grid and homogeneous shape factor, while areas resulting from decades or centuries of localised interventions and constructions showed a much greater diversity of block shapes.

2.5 The social layer

The social layer relates to human activity in the city, which gives a meaning to the spatial layer, creating places in space, as described by Park (1936) and Tuan (2011). A growing attention is taken on how those two interact, for example when it comes to public health (Creatore et al. 2016; Glazier et al. 2014), in this context Weiss et al. (2007) have proposed a methodology to define neighbourhoods that are relevant for health research. More generally, Chaskin (1997) has provided a great overview of the role and importance of neighbourhood in the literature. It highlights the importance of this area near ones home in the formation of stable communities and safe places to live in.

Additionally, the visual perception on the environment with which the citizen interacts has been studied over the decades and with ever evolving technological tools. A great example from the past is Lynch (1960), in his famous book *The image of the city*, the mental map of a selected area in 3 American cities is built from the information ob-

tained through interviews with local workers or residents. Based on their description of their daily routine, it has allowed to better understand the local perception of space and the recognisable landmarks that people rely on when navigating a city. This work illustrates how, on the one hand, each of us has a very personal vision of space, and our mental map is something intimately linked to the way we interact with the environment on a daily basis. On the other hand, by studying a sample of people that visit certain places, we can construct a more general mental image that the visitors of an area have. A research that asked people in various parts of Seattle to define their neighbourhood found the most important elements to be space and community (Guest and B. A. Lee 1984).

More recently, using the variety of digital and modelling tools at hand, Salesses, Schechtner, and Hidalgo (2013) were able to build a map of the perception of neighbourhoods safety in New York based on a comparison of street photographs. They found links between the perception of safety in the urban layout presented on pictures and occurrences of violent crime. This sort of research creates bridges between urban sciences and planning and provides valuable insights into possible causal relations between the urban forms and social variables.

In the context of defining substructures within a city based on social attributes, quite often we see amenities being used as a proxy between the spatial and the social and as a mean to better understand how locally the environment is used by people and businesses. By studying locations and numbers of amenities in Boston, Hidalgo, Castañer, and Sevtsuk (2020) were able to build a model describing the co existence of amenities and found that the density distribution in space allows to identify the various neighbourhoods of the city as seen from a citizen perspective. Dong, Ratti, and Zheng (2019), with certain limits, modelled socioeconomic attributes of neighbourhoods based only on the types of food that was available at the local restaurants.

Additionally, Xu et al. (2020) demonstrate locally the inequalities of access to some basic amenities in 6 big cities of the world, by comparing with the case of optimally distributed amenities and the effective distance travelled by residents to reach them. Similarly, but on the scale of the US and Brazil, Barbosa et al. (2020) have outlined

the differences in transport habits of multiple social classes of people in their neighbourhoods, and found links between greater distances travelled, and fewer local available amenities. This extensive comparative research on Brazilian and American cities looked at the diversity of amenities and the distance to access them. It found opposing results for the two countries that were attributed to differences in their level of development. In Brasil, the effective distance travelled to amenities increased with higher income groups. And opposite results were found for the US. In general, cities with a more developed public transport system showed a more homogeneously distributed access to amenities through income groups, while more segregated ones had opposing results.

2.6 Human mobility

With smartphones invading almost every aspect of our life in the last 15 years, it has become possible to use the data they generate to better understand the behaviour of people, their points of attraction in the underlying spatial structure. The science of cities has recently gained great insights into the structure of cities and mobility through the study of GPS signals emitted by devices. Louail et al. (2015) have deconstructed mobility patterns of people in the biggest cities in Spain, introducing the dilatation index that allows to measure how polycentric is a city, showing how the number of hot-spots grows sub linearly with population and how their location evolves with time throughout the day. Other authors adopt tools from statistical physics to model patterns of trips, with periods of displacement of a certain distance followed by stops of certain time (Jiang, Yin, and Zhao 2009; Bazzani et al. 2010; Giannotti et al. 2011; Riccardo, Armando, and Sandro 2012).

The possibility to verify various mobility models through the use of such data has generated a great body of literature. Recently, Schläpfer et al. (2021) found in the mobility patterns of users in Boston a conservation law, similar to the conservation of energy in physics, linking the distances travelled from home to the frequency of such travels. In the context of this work, it is interesting to interpret this result as an indicator that on the neighbourhood level, we are likely to see many short trips done by users. This result seems very intuitive, with a user doing multiple short trips from home to

the shops, market, cinema, jogging throughout the week, but going far, for example to see friends, or visit another part of town or even another country less frequently. The scale of various containers of human activity has further on been investigated by Alessandretti, Aslak, and Lehmann (2020), where a model for generating user traces is compared to actual data and shows high statistical significance.

By exploring in more details the movement of individuals Alessandretti, Sapiezynski, et al. (2018) have found that although locations visited by an individual change, the number of distinct locations at any point in time remains constant, indicating a conservation of ones capacity to visit places in a given time. In the same paper, it is concluded that the average number of locations in the activity spaces of people is about 21. Other authors find persistent patterns in the daily movement of individuals that show that the first two preferred locations take up to 40 $\%$ of a users time (González, Hidalgo, and Barabási 2008), and that most of the time people can be found in about 5 locations. This work models an individuals mobility through the probability density function $\Phi_a(x,y)$ of finding the individual a at location (x,y).

2.7 Investigate locations through users data

By studying tap in and outs from Oyster cards on the London transport network and geo tagged tweets, the attractiveness, vibrancy of London has been investigated (Sulis et al. 2018). Building up on Jane Jacobs notion of vitality, vibrancy of a neighbourhood, and using individual mobility data, it is possible to link the intensity of users activity with high spatial and temporal precision, to various places and deduce their overall attractiveness on the city level.

Dong, Chen, et al. (2017) find strong links between recorded occupations of various shopping areas and economic, commercial activity that was declared in a set of commercial areas in China. This outlines the possibilities that high granularity mobility data can have in the understanding of lifestyles of people and their points of attraction for various uses. This new type of data has gained wide commercial use as well, and location intelligence is now an important area of activity of the services, housing, marketing industries [6].

[6] https://www.centreforcities.org/high-streets/

2.8 Final note

Strict neighbourhood boundaries are malleable and varying. From the personal perspective, we build a mental map based on the places we visit and the things we see, which depend on the gender, age, interests, occupation, lifestyle(Guest and B. A. Lee 1984). Throughout the research, neighbourhood boundaries depend on the topic being addressed. And because of public transport, these boundaries are easily extended, the neighbourhood is thus an open system that is part of a much greater whole. There is no consensus as to what are neighbourhoods and this work intends to propose a potentially useful interpretation of this space that can be applied to understand cities on a smaller, more human scale and be able to identify where the citizens activity tend to agglomerate on the scale of the area around their home. The neighbourhood is seen as the system of interest for understanding the local behaviour of people, and as the unit on which planners and urban designers can effectively rely on to act on the environment.

3 Data

3.1 OpenStreetMap

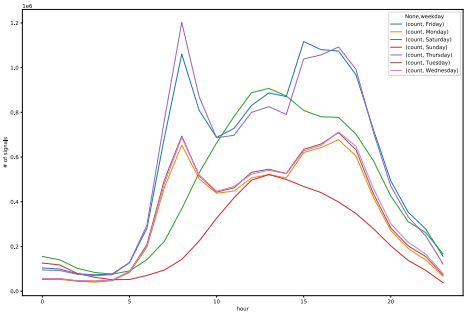
The open source mapping service OpenStreetMap (OSM) is used for studying the amenities. It consists in an active community of contributors from all over the world that categorise and map every possible feature of the environment. The data is very rich especially for central and western Europe and provides an up to date and precise representation of cities in terms of infrastructure, services, amenities, shops etc...

3.2 Amenities of interest

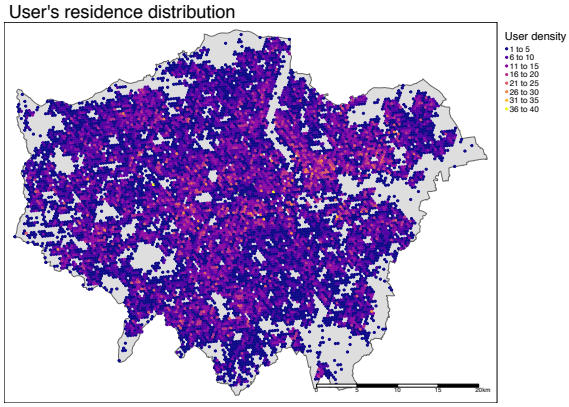
The amenities of interest can be chosen in various ways, and various categories are investigated throughout literature Hidalgo, Castañer, and Sevtsuk 2020,Xu et al. 2020,Sevtsuk 2014. Here, the full set of amenities is investigated except those representing insignificant importance for a neighbourhoods livability, vibrancy. Amenities that are chosen are the ones that serve as attractors of people and thus are fundamental to neighbourhood life, like restaurants, museums, galleries, cafes, shops, or represent the essentials of public services such as transport, police stations, post office etc... The full set of amenities selected is provided in the appendix. It consists of a joint set based on the OSM map feature tags ”amenity”, ”shop” and ”tourism” for a total amount of 96 different categories. The full set is shown in appendix A.1.

3.3 Mobility data

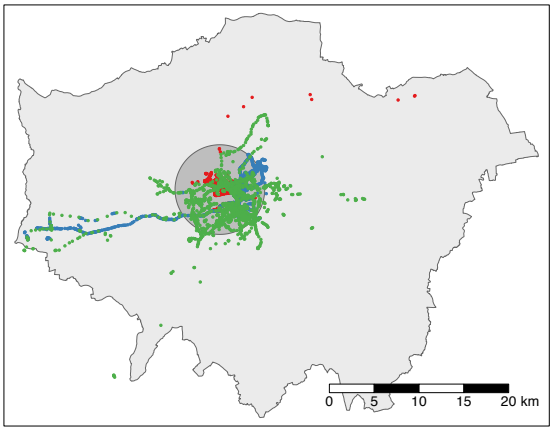
The mobility data is composed of GPS signals emitted by mobile devices. Each signal has it’s exact latitude-longitude coordinates, the date and hour at which it was emitted and the id of the emitting device. For each emitting device, an aggregated residence and work location is known for its user, assumed to be 1, it is deduced from an averaged position of the signals during the night and during the working hours respectively. The period of observation spans between 13 and 27 of February 2020. Examples of traces of 3 devices for the period of observation are given in figure 3.1c,3.1d. The spatial distribution of the users residence is shown of figure 3.1b, and the profile of the intensity of signal emissions is shown with respect to the hour of the day and for each day of the week on figure 3.1a.



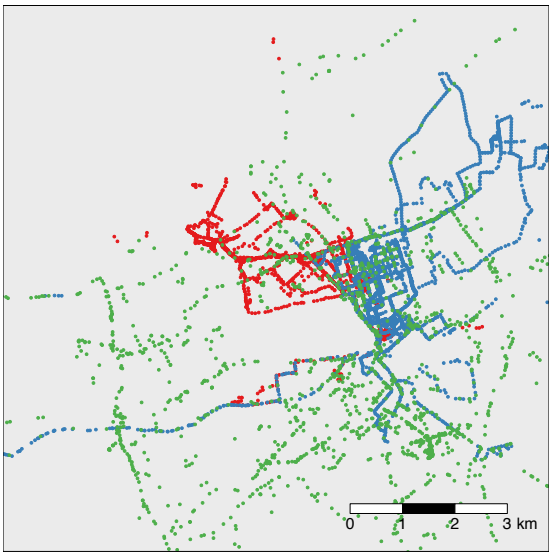
(a) Intensity of signals with respect to hour of the day. Colors correspond to weekdays.



(b) Number of users distributed by their residential location.



(c)



(d)

Figure 3.1: Overview of the data on users and their GPS traces. On the bottom plots, each colour is for a user, 3 are shown as an example.

3.4 Spatial partition

The spatial analysis of mobility data and it's links with the work on amenities is done by aggregating values on a hexagonal grid, consisting in hexagons with a side of 190 meters. The mobility data is distributed on it, and for each cell, the users that reside in it are known. This allows to aggregate sufficiently the home locations of the devices to remove potential ethical issues and provides a granular enough spatial division to investigate the system in great details. One additional manipulation is done on the original hexagonal grid which consists in distributing the cells that do not contain residents (non-residential) to it's nearest lying residential cells. The consequence of this is the non populated areas of London, like the Lea and Thames basins, but also parks and sometimes big railway junctions showing an irregular hexagonal pattern compared to the rest of the map. A sample of the empty grid can be found in appendix A.5.

4 Neighbourhoods through amenities

The presence and accessibility of amenities is emphasised as essential in the formation of vibrant, strong neighbourhoods and communities and the modern urban resident seems to be strongly attached to the presence of a variety of opportunities nearby his home in the form of amenities. However, the existing literature seems to be limited when it comes to adopting a quantitative description of quality of the urban environment in terms of it's amenities. The land use mix (LUM), an index based on Shannon's entropy and applied to the land uses seems too limited to illustrate the diversity of environemnt. The reason for that lies, arguably, on the rather small set of possibilities to which it is applied to, referred to as the primary use categories. They are residential buildings, offices, industrial facilities, entertainment places, education facilities, recreation facilities, museums, libraries, and galleries (Sung, S. Lee, and Cheon 2015; De Nadai et al. 2016). Others have used this index on a set of 8 categories of amenities (Barbosa et al. 2020), or 10 activity types, such as going home, eating, social, recreation etc... (Zhong et al. 2017). But to actually make the best use of the entropy function, it is worth remembering that in the statistical theory of gases, Boltzmann introduced it on systems containing enormous amounts of configurations. Hence, the bigger the set of possibilities, the more revealing can be the values of the entropy.

4.1 Methods

The index is measured at the location of each amenity and consists in taking the Shannon entropy (Shannon 1949) of the amenities within t minutes of walking from a given one. The index is defined as follows:

$$S_i = -\sum_j p_{ij} ln(p_{ij}) \tag{4.1}$$

where S_i is the entropy of amenity i, $p_{ij} = \frac{m_{ij}}{M_i}$ is the number of amenities of category j divided by the total number of amenities within t minutes walking from amenity i. The literature suggests that a time of $t = 7$ minutes, which corresponds to about 500m walking distance, is a reasonable value on the neighbourhood scale and in accordance with literature (Hidalgo, Castañer, and Sevtsuk 2020),(Sulis et al. 2018)., The

implementation of this method consists in using isochrones around the location of each amenity, it allows to better understand the walking accessibility and increases the precision of the index as opposed to, for example, taking buffers. The method is especially useful when natural or man made boundaries are present in the form of railways, highways with no crossing, the Thames or the Lea rivers. The entropy value is dependent on the number of categories in the set of amenities that are considered, in this study it is $N_c = 96$. This gives an upper limit to the entropy index corresponding to the presence of all the amenities at once. The maximum value is $S_{max} = ln(N_c)$ when every category of amenity is present and in this case $S_{max} = 4.56$. The minimum value of the index $S_{min} = 0$ is obtained when only a single type of amenities is present at the location.

4.2 Interpretation

The index is intended to translate the diversity of the local urban environment in terms of social attributes, businesses, sustenance, leisure, nightlife, education, cultural opportunities, basic needs, medical care, sport etc... (see appendix A.1). It also illustrates the accessibility within reasonable walking time, and the different tendencies to cluster that amenities show depending on the area. It can be interpreted from the information theory perspective as translating the diversity of the ground floor activity of the urban environment, an aspect of the environment praised by Jane Jacobs and nowadays considered essential for a vibrant neighbourhood life. The greater the value, the greater the range of different amenities present at the location. Categories of amenities correspond to their OpenStreetMap classification described in section 3.2. More diverse or denser places could be linked to a phenomenon known in economics as competitive and complementary clustering, the first one being when businesses of a similar profile, such as restaurants, collocate in a place of interest to compete with similar businesses for the visitors of the place. The former occurs when various businesses collocate in a place in which their target audience coincides, but not necessarily the activity they do. Complementary clustering is also explained by the preference of users to visit less locations while shopping, saving time and money on transport. This happens in places with high passing flows of people. It allows busi-

nesses to be exposed to more visitors, increasing their chance to be seen and attract clients, and benefit from spillovers from surrounding places (Berry 1967). From figure 4.1, one can see how high values of the diversity index are reached relatively quickly for an increasing density, illustrating even more this tendency of different amenities to collocate.

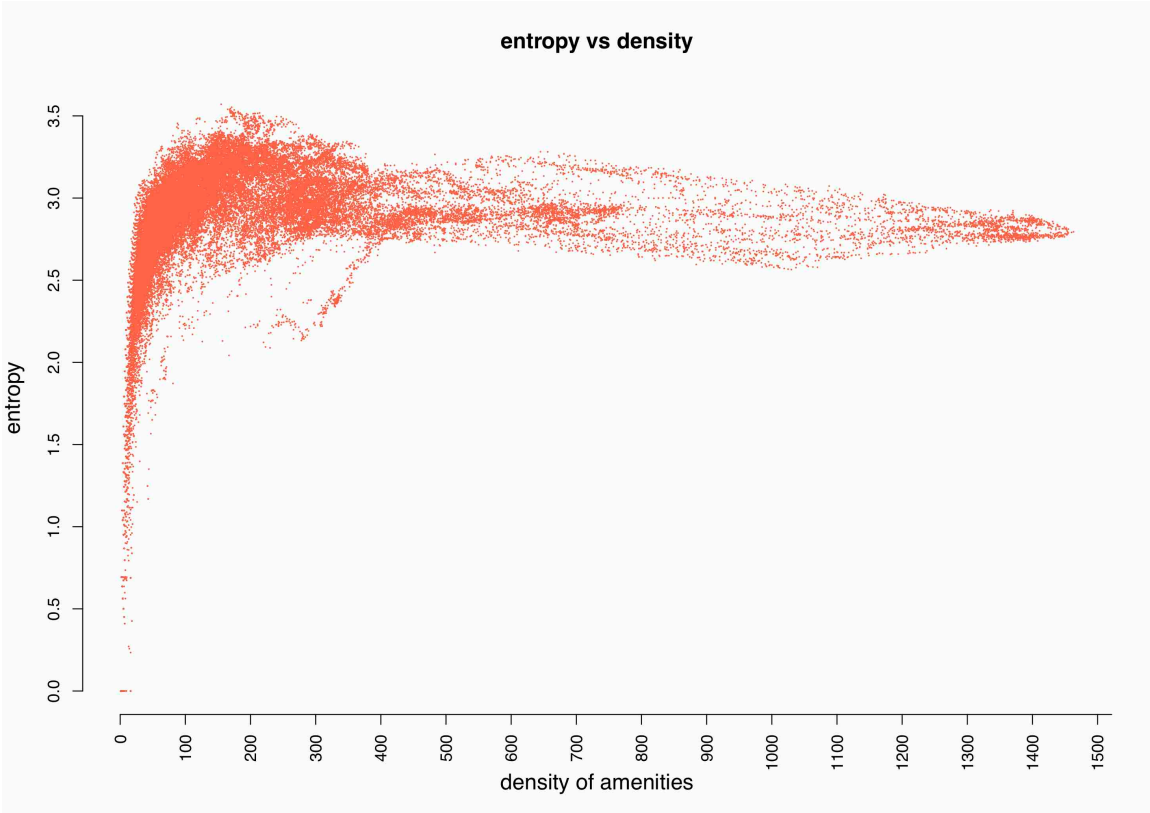


Figure 4.1: Entropy index compared with density. Density is intended as the number of amenities within 7 minutes of walking distance.

The accessibility and diversity of amenities was also found to be a driver of prices in the local real estate market (Jang and Kang 2015) and an attractor of individuals, such as the creative class (Florida 2012). The agglomeration of diverse amenities in proximity to each other also creates walkable neighbourhoods, that have gained high praise among urban residents in the last decades (Ehrenhalt 2012). While the neighbourhood has always been around, the car-centric, barely walkable environment that the planning practice of the mid 20th century omitted it from it's considerations. Nowadays, the priorities have changed and the neighbourhood is often put at

the centre of the debate and redevelopment activity. The recent sanitary events have drastically accelerated some trends that were already present, such as active travel by bike or foot and work from home. Moreover, highly visited streets that have been fully given to pedestrians to facilitate access, were voted to be kept this way by local administrations and residents in most of US cities [1]. The visible drop of importance of the CBD[2] as the main attractor of people will arguably have an effect of redistributing people [3] into places in greater proximity to their home, highlighting even more the increasing importance of the neighbourhood in the modern society.

4.3 Diversity vs Quantity

The results show quite a varying profile of entropy with respect to the number of amenities, the maximum value of $S_{max} = 3.57$ is reached with a relatively small density. In the same time, higher densities ($>$ 400) show lower entropy values as can be seen on figure 4.1. From the figure 4.2, it can be seen how the highest concentrations of amenities are actually spatially very centred around the area of Mayfair and spreading slightly to the east to the City. Interestingly, the entropy scores measured in the highest density places are lower than the peak entropy measured at around 150 amenities of density in other areas. On figure 4.5 one can see the different areas in term of density for a threshold of 400 amenities. The map 4.4 shows the amenities with a threshold of $S_{min} = 3.1$, and one can see on it some of the well known neighbourhoods of London emerge among others, such as Marylebone, Bloomsbury, Kensington (see figure 4.4). The diversity index allows to shed the light on the quantity and spatial distribution of local centres of diversity throughout the city. These places bear a strong local identity and are quite recognisable in the London city scape, they are famous for their leisure opportunities such as cafes, restaurants, shops, but present a wider set of available amenities like grocery stores, theatres, cinemas, museums, art in the street, post offices, schools etc... Further away from the central and famous areas of London, the entropy index allows to find the other local neighbourhood centres as can be seen of figure 4.3. While figure 4.2 seems to be illustrating a monocentric distribu-

[1] Link to the article.
[2] The Death and Life of the Central Business District
[3] City of London Plans for Life After Covid as Its Offices Empty

tion of the density of amenities, figure 4.3 shows a much more complicated, spatially distributed set of places in the city, further more illustrating it's polycentricity.

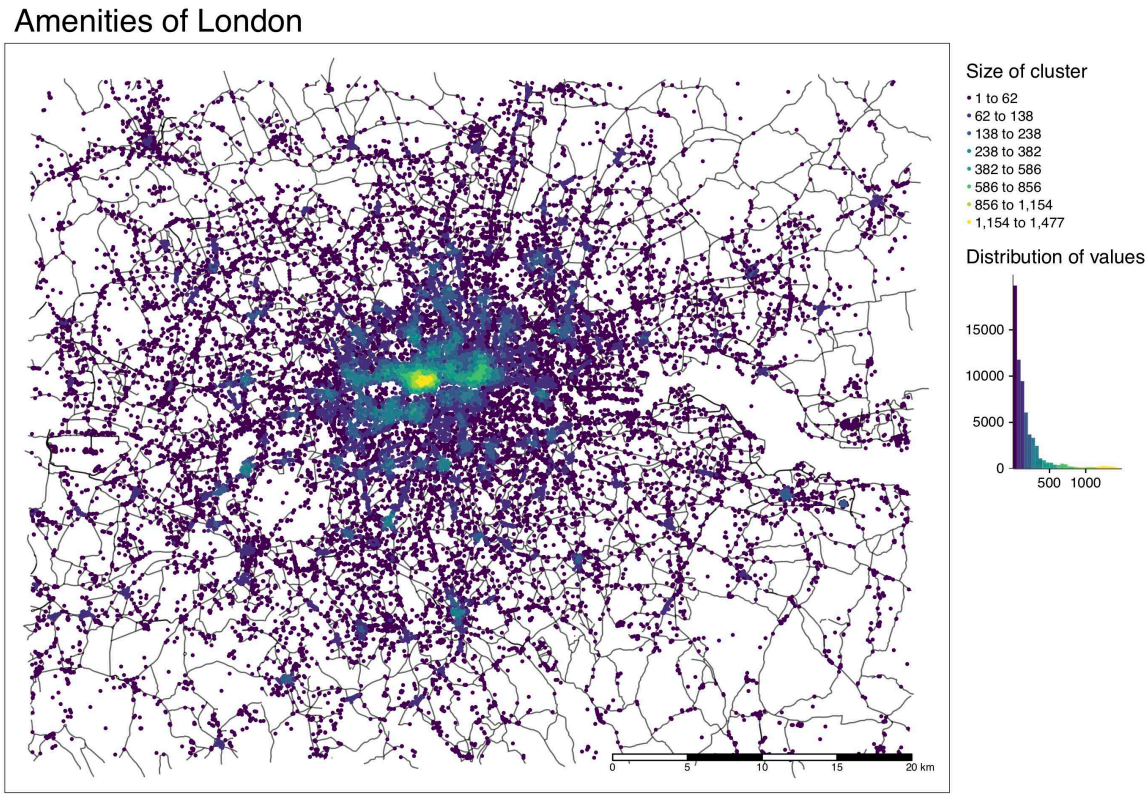


Figure 4.2: The density of amenities in London

The lower diversity of high density places, illustrated on figure 4.5 and on the plot 4.1, could be linked to a phenomenon of agglomeration as described earlier in this paragraph, illustrating a specific profile of the amenities. The areas that emerge are Mayfair and the City, which are well known places of tourism, commerce and business, with a possible greater need in specific types of amenities like restaurants, pubs, boutiques, thus reducing their overall entropy score. Other famous places, such as Shoreditch, that are found to be the most vibrant in London in terms of leisure and night life (Sulis et al. 2018), do not appear to contain a diverse mix of amenities in this model, highlighting the orientation of the area towards a specific kind of activity. Adapting a more general point of view of the index, the values obtained can be interpreted as the measurements done in the location of each amenity of a continuous planar function $S(x,y)$, that we call the amenity entropy of space (AES). When

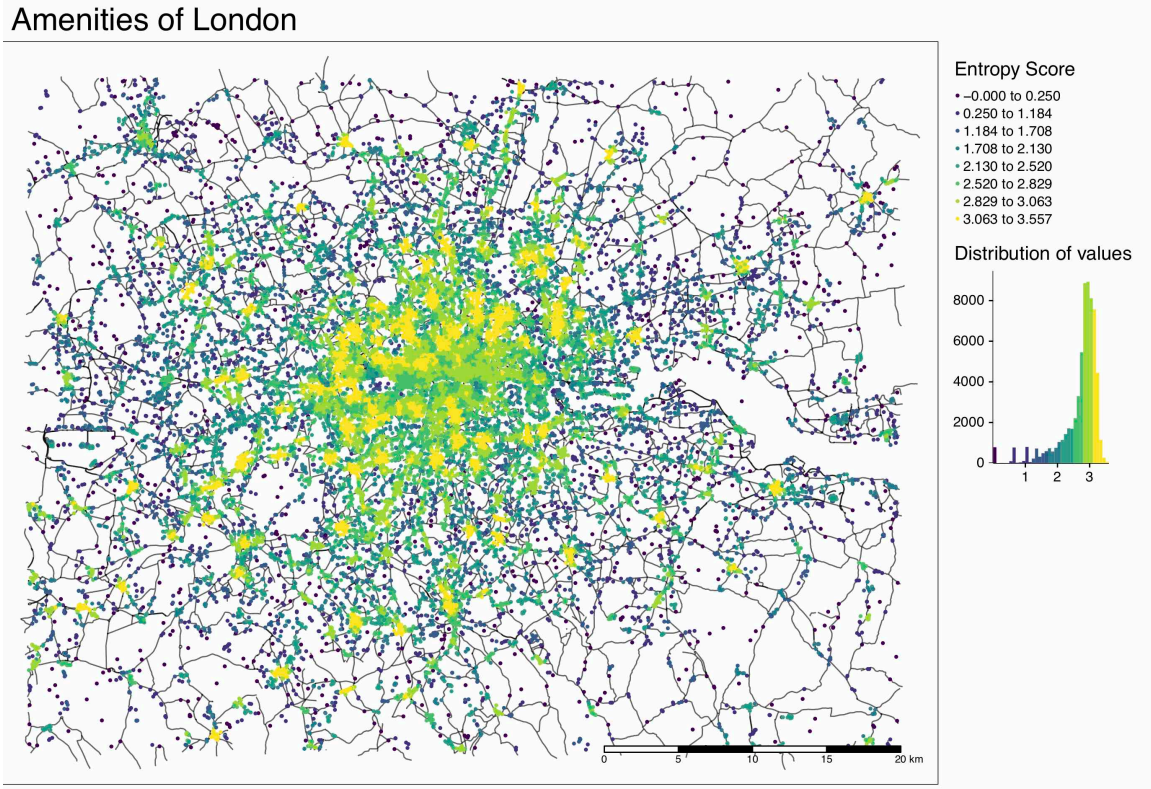


Figure 4.3: The entropy of amenities in London

attempting to define neighbourhoods as a unit that contains as diverse amenities as possible and maintaining a human-scale, it is of interest to look for the local maximum of the index and define it to be the centre of the neighbourhood. To assign each amenity to a different centre, it is useful to look for the nearest local maximum and cluster them together. The next section describes how this is implemented.

4.4 Defining neighbourhoods

The algorithm reproduces a step process during which, starting from a given amenity, the maximal entropy within 7 minutes walking time is found, then, until the maximum found corresponds to the amenity itself the process is iterated. Finally, the original amenity from which the step wise process was started is assigned to the local maximum found. This method assumes that the entropy function is relatively smooth, allowing to climb up it's gradient through this process. More details can be found in the appendix A.2. This allows to assign each amenity, to a local centre and then cluster them into neighbourhoods. This algorithm is taking into account the underlying spa-

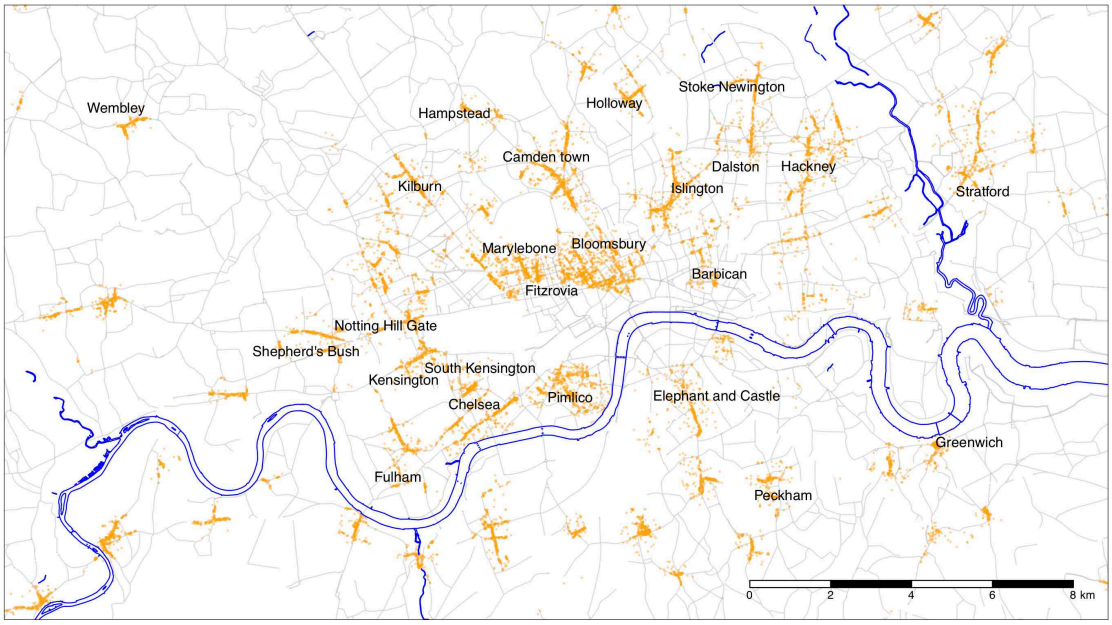


Figure 4.4: The famous neighbourhoods found. Blue lines are the main water bodies of the city. Orange colour corresponds to high entropy amenities ($S > 3.1$).

tial structure through the isochrones and as a consequence is intended to translate the way a user will be using the space when moving towards the neighbourhood centre, going uphill along the gradient of the AES function and minimising the time to access to the local maximum. This method illustrates the logic of Jane Jacobs that sees walkability and diversity as key factors of a vibrant neighbourhood. The process of clustering is illustrated on the example of an amenity in Marylebone on figure 4.6. With the locations where entropy is maximum as the centres of the neighbourhoods, the other amenities around are assigned to a specific neighbourhood based on its proximity. Figure 4.7 shows with differently coloured polygons the neighbourhoods that emerged from the iterative process. This vision of the neighbourhood, motivated by the New Urbanism practice, agglomerates neighbourhoods around the local peak of diversity of amenities and considers the pedestrian accessibility.

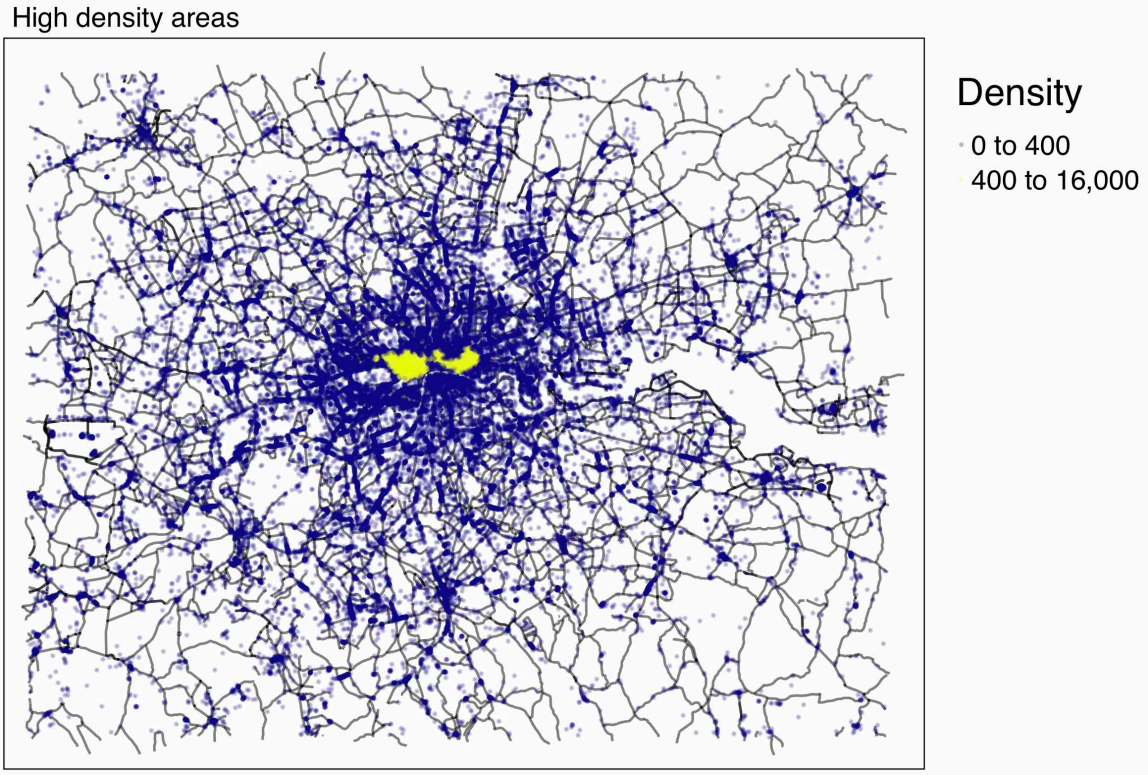


Figure 4.5: Location of the places with the highest density of amenities. The blue area corresponds to densities of less than 400 amenities within 7 walking minutes, and the yellow area are from 400 up.

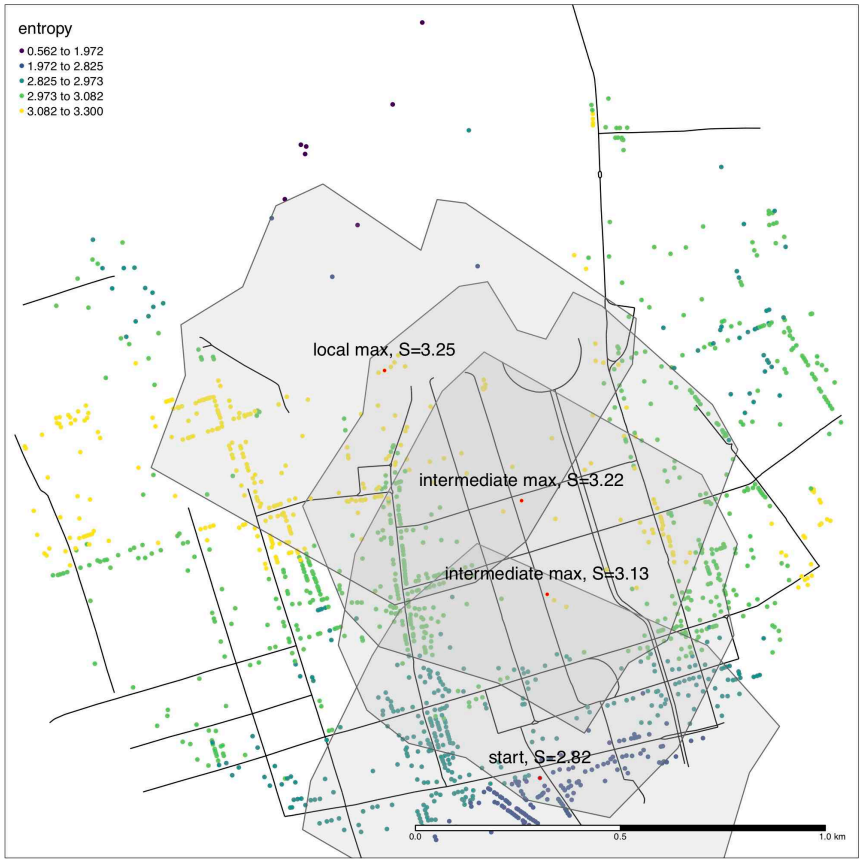


Figure 4.6: Example of the clustering algorithm linking an amenity to it's nearest local maximum.

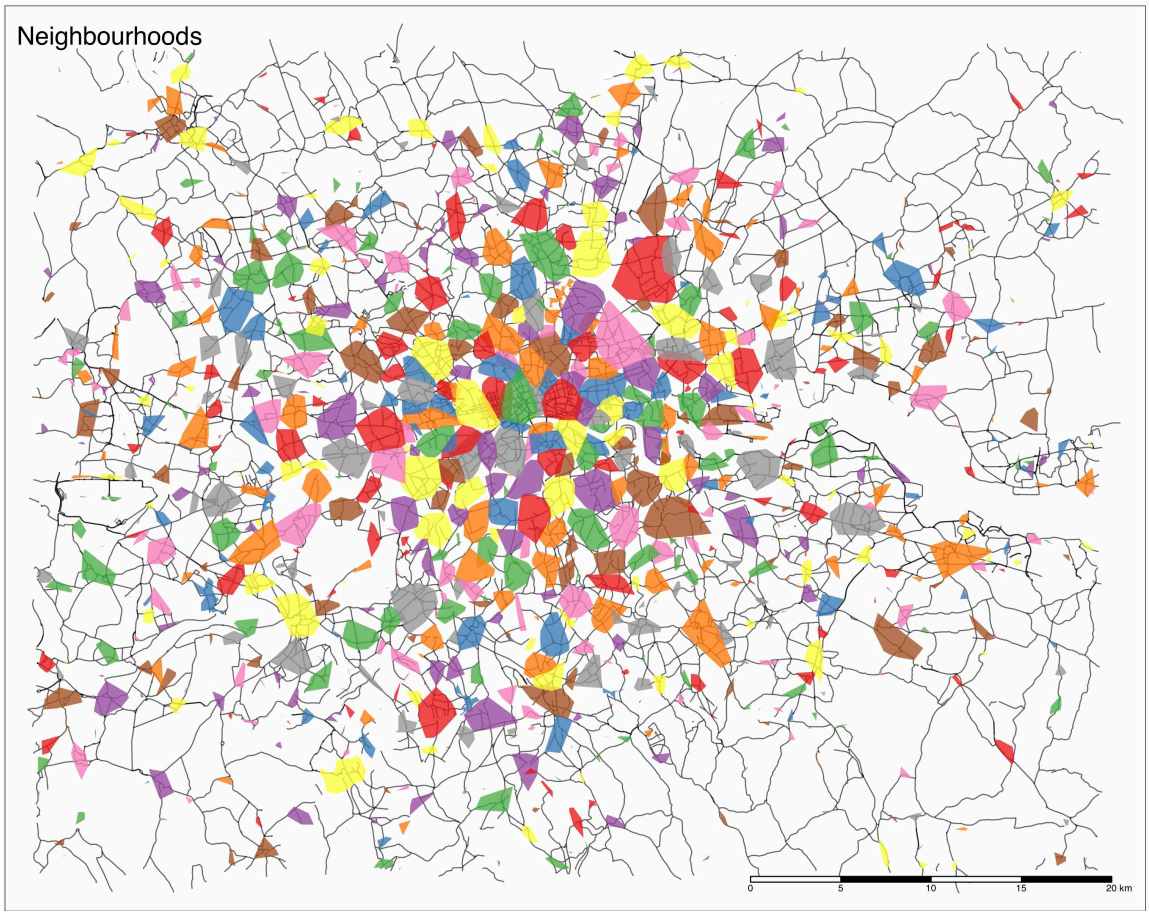


Figure 4.7: A schematic representation of the neighbourhoods found through the investigation of the AES. Each distinct coloured polygon represents a different neighbourhood.

5 Neighbourhoods through mobility

The literature covered in the section 2.6 outlines the various scientific methods that are applied to user mobility data and the increasing understanding of general patterns in individual human mobility that result from it. The summary inspection of the present data set highlights similar trends to the ones observed in the literature, and this following section provides a methodology that allows to aggregate human mobility by visited places and overlap of individuals activity spaces. The following section aims to addresses the problem of understanding the behaviour of people around their residence location and identify a subdivision of space that would illustrate the tendencies to visit exhibited by local inhabitants in the vicinity of their home. The understanding of these aspects of human mobility could help better inform local planning decisions and actions, in terms of road redevelopment to create safer, nicer places for local residents to move around their preferred areas in the neighbourhood, or to identify the areas segregated in terms of access to amenities and containing low diversity.

5.1 Methods

From the user perspective, (González, Hidalgo, and Barabási 2008) introduce the presence density function $\Phi(x,y)$, which corresponds to the probability of finding an individual at (x,y). This function can be used as a mathematical representation of the activity space of a user, where visited locations have a non zero density value. Similarly, we can define the density function on the discrete set of locations available to us through the hexagonal grid, and the activity space of a user i as the locations where $\Phi_i(h_k)$ is not 0:

$$A_i = \{h_k, \Phi_i(h_k) > 0\} \tag{5.1}$$

where h_k is an element of the grid.The value of the function gives an estimate of the importance of the cell for the user, in this case represented as the fraction of the total activity (number of signals) emitted from within it. Only movement happening between 5am and 11pm is looked at. This formalism allows now to investigate space through the presence of users in it. The overlap of the activity spaces of two users i and j can be modelled as:

$$A_{ij} = \{h_k, \Phi_i(h_k) * \Phi_i(h_k) > 0\} \tag{5.2}$$

The following assumption is taken to facilitate the identification of substructures of the city. It consists in simplifying the density function only to return a value of 1 if the location is visited, and 0 if not. It can thus be defined as follows:

$$\Phi_i(h_k) = \begin{cases} 1, \text{if } h_k \in A_i \\ 0, \text{if } h_k \notin A_i \end{cases} \tag{5.3}$$

This approximation removes the differentiation between the importance of places for a user. It is assumed that on the local scale of the neighbourhood, where a majority of trips are done by active travel means (foot, bike etc) and on the ground, we care similarly about the destination as we do about the road to get there, whether for safety reasons, or aesthetic ones. This assumption most likely does not hold for movement beyond the neighbourhood, where a public transport or car trip removes our considerations of the space we are passing by and the final destination matters most of all. Moreover, from the perspective of planners, it is important to look at the whole activity space, rather than the strict origin and final destination, in order to better inform strategies of local redevelopment and adaptation of infrastructure for local movements.

It is now proposed to see the city as a set of locations $H = \{h_k\}_{k=1}^{N_H}$, where N_H is the total number of locations that the grid includes. Each location is modelled by a hexagon in the grid and can further on be represented as the set of users that visited it, as found through the density function determined from the positions of GPS signals of users. Hence $h_i = \{k, \Phi_k(h_i) = 1\}$. A metric is then defined on the locations set H in order to investigate the similarity of it's components.

$$\sigma_{ij} = \frac{|h_i \bigcap h_j|}{|h_j|} \tag{5.4}$$

where $|.|$ denotes the cardinal of the set and $\sigma \in [0,1]$. A value of 0 means there is no similarity in the visitors, and 1 means that all the visitors of h_j also visited h_i. The matrix σ is asymmetric, $\sigma_{ij} \neq \sigma_{ji}$, as opposed to a more common definition of similarity like the Dice similarity coefficient (Cao et al. 2020) that would be $DSC = \frac{2|h_i \bigcap h_j|}{|h_i|+|h_j|}$. Using an asymmetric similarity measure allows to more effectively differentiate between

locations, allowing to better understand the hierarchy of various places and their respective asymmetric relations. Simply speaking, it is possible that a big fraction of users from h_j will visit a location h_i, $\sigma_{ij} \to 1$, but they might only represent a small fraction of the users from h_i that will also be visiting h_j, $\sigma_{ji} \to 0$. This asymmetry allows to better highlight the hierarchy of places in the urban environment and effectively merge them. Following the example, h_j can be a typical residential location, and h_i its local neighbourhood centre, or a place where users activity spaces overlap a lot, like important road links, transport hubs or town centres. Various distance thresholds are presented throughout planning and academic literature when it comes to neighbourhoods. Sometimes, it is imposed by the nature of the data, for example in the study of signals from cell phone antennas (González, Hidalgo, and Barabási 2008), where the neighbourhoods are introduced as the Voronoi tesselation, or catchment areas of each antenna reducing the spatial precision to about $r = 2$km. In other papers, the initial subdivision of the city is into square grids of a certain side length on which all the variables studied are aggregated (Louail et al. 2015; Gallotti, Bertagnolli, and De Domenico 2021; Bassolas et al. 2019; Barbosa et al. 2020). Other works rely on the public transport network and breakdown the city into areas of interest by taking the Voronoi tesselation of the stations set (Prieto Curiel et al. 2021). In planning, with the New Urbanism principles or the 15 minutes neighbourhood suggest that distances of about $1.5 - 2$km are reasonable to be covered by active travel means, whether bike or foot. Recent fundamental results on mobility obtained by Alessandretti, Aslak, and Lehmann (2020), suggest that human mobility can be well modelled when distributed into separate containers of different characteristic sizes common to most of the users. The size limits of the smallest container introduced to be the neighbourhood are found for an average of $3km$, this area is also found to contain 40% of each individuals activity. The average container size is kept as a distance threshold for this study. In this case, the proportion of the original data set of activity corresponding to this threshold is 36$\%$, bearing a certain similarity to findings from this work.

5.2 Identifying neighbourhoods

Using the similarity matrix introduced in the previous section and by performing a hierarchical clustering on the euclidean distance between each row, the respective relations between places in terms of their visitors are revealed across London.
On figure 5.1, one can see the map with the partition into 380 neighbourhoods of the city, from figure 5.2 one can see that this structure appears after a discontinuous increase in size of the biggest component when navigating upwards the hierarchical tree (omitted here because of it's size). The discontinuities, as can be seen on the figure, are followed by stable periods of the biggest component, during which only minor changes occur in the structure of the system. The investigation of discontinuities in the functions describing physical systems is a key method for identifying critical behaviour under a changing parameter, in this case the similarity of the areas in terms of visitors, defined in equation 5.4. The vertical lines represent the critical states of the system, corresponding to transitions where the biggest component undergoes major changes in size. In order to be coherent with the rest of this study, the critical value of 380 is investigated further as at this state, the median typical distance of the partition is $L = 1680$m, which corresponds to reasonable, walkable neighbourhood sizes. The typical distance of a single neighbourhood being the square root of it's area $l_i = \sqrt{n_i}$, n_i is the area of neighbourhood i and L is the median of the set $\{l_i\}$ for a given partition. The partitions for the other selected thresholds can be found in the appendix A.6 and could be further investigated in future works.

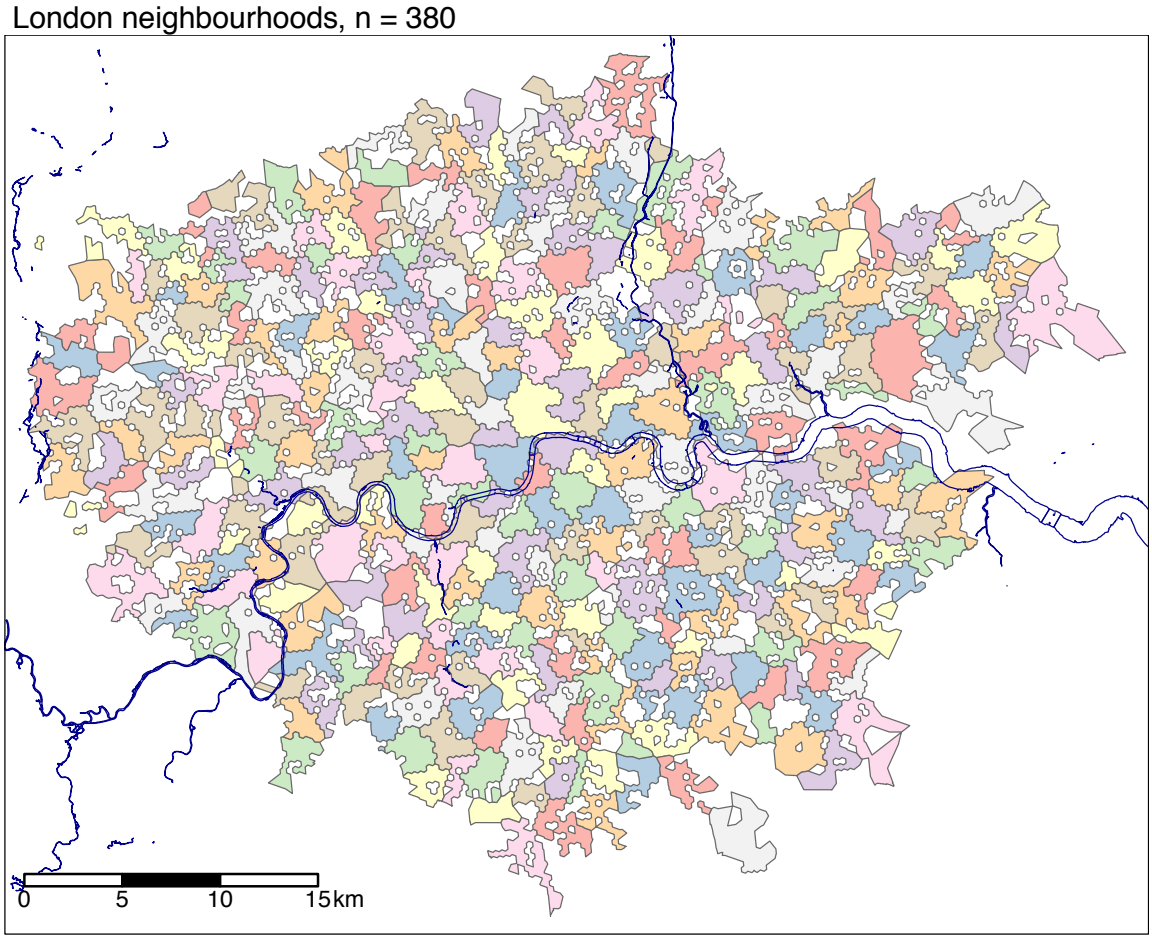


Figure 5.1: The neighbourhoods based on similarity of visitors for a partition into 380 components of the hierarchical tree.

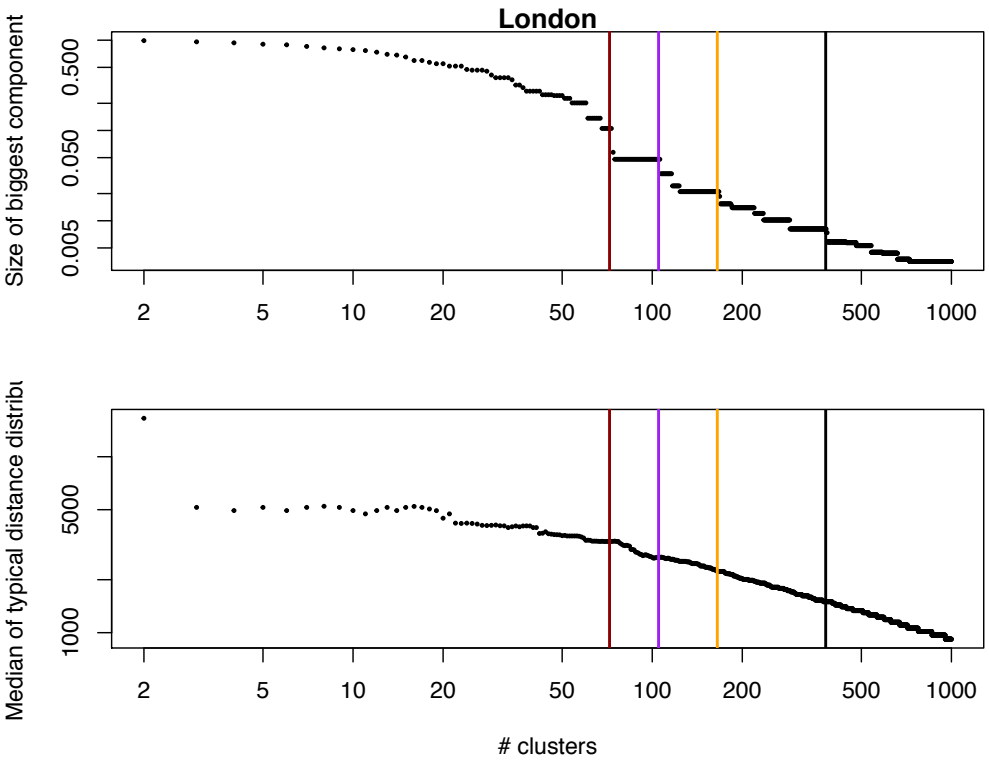


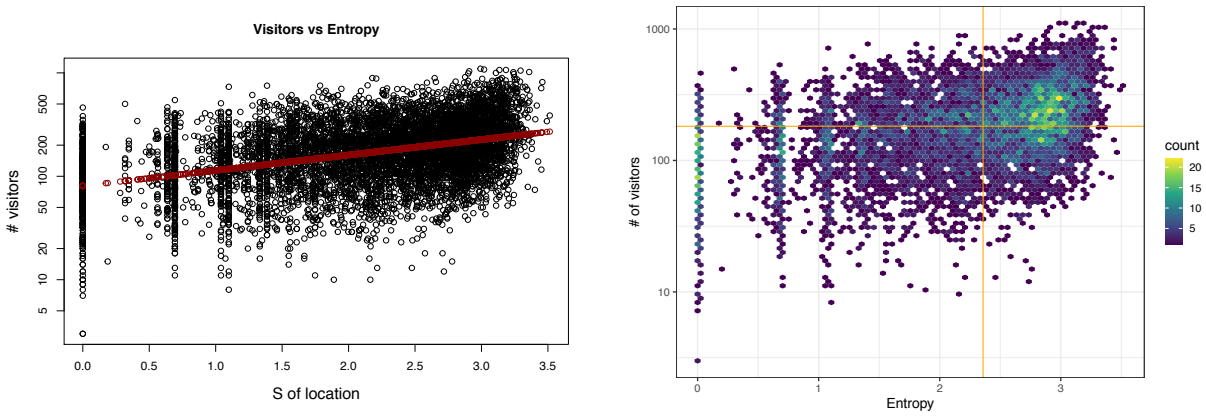
Figure 5.2: Top plot: size of the biggest components with respect to the partition. Vertical lines represent examined cases. The black line is for the kept threshold of 380. Bottom plot: median size of the typical distance from the distribution for each partition.

6 Discussion

The first part of this work focuses on the neighbourhood from the perspective of space in which amenities are located and serve as indicators of local neighbourhood life. The ideas of Jane Jacobs on the mix land uses and walkability as fundamental aspects of vibrant neighbourhoods are combined into the amenity entropy of space (AES) function that illustrates the pedestrian accessibility and the diversity of available amenities on the local scale throughout the city. The local maximums of this function are further on proposed to be seen as local centres and the surrounding amenities are merged together to form neighbourhoods. In a more general sense, it illustrates the zones of attraction to local maximums of diversity by active travel means.
The second part of this work shows how, through a methodology derived from physics, it is possible to extract from the mobility patterns of users around their home a substructure of the city. Bearing in mind the walkable size of the neighbourhoods and by looking at the level of similarity between visitors of a location, it was possible to find a substructure of the city that is intended to represent the neighbourhood as seen by a local resident. This method uses percolation on a parameter of similarity of visitors in different locations to construct neighbourhoods from a community perspective.

6.1 Diversity of visited places

When looking at the visited places that contain amenities within 3km from home, we see that higher entropy places tend to get more visitors.
By further investigating the data at locations where amenities are present, an exponential relation is found between the number of visitors and the entropy of the place. $V(s) \propto e^{\beta s}$, where $\beta = 0.34$, V is the number of visitors and s the entropy. This model gives an $R^2 = 0.97$, indicating that high entropy of amenities of a given location can explain the presence of an increased number of users, showing consistency with the agglomeration of activity phenomenon found in literature and introduced in section 4.2. Additionally, this result verifies the ideas of Jane Jacobs in the context of neighbourhoods. More diverse places contain more individuals, and the methods and results from chapter 4 could further be applied to develop our understanding of the urban en-



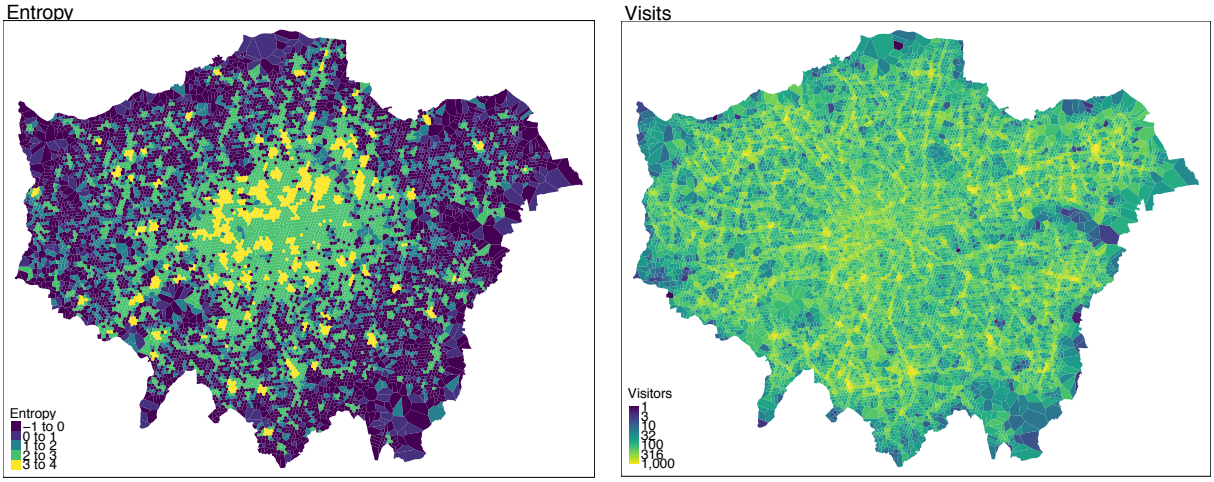
(a) (b)

Figure 6.1: The visitors recorded at locations with amenities plotted against the diversity. The number of visitors is log scaled. The mobility taken into account is on the neighbourhood level, within $r = 3$km from each users home. The dark red line on 6.1a shows an exponential relation between the two variables. On 6.1b, the colour scale corresponds to the number of observation. Orange lines are the median for each quantity.

vironment locally.

6.2 Community through hierarchical percolation

The method presented in section 5.1 can have future applications in the understanding of human mobility and local communities. It's strength lies in its pure reliance on mobility data from users portable devices that allows to extract spatially coherent results, as demonstrated in the chapter 5. It can have future implications on the way the urban environment is mapped and studied and can be especially useful in places where exact cartographic information is absent, allowing to identify communities from the pure perspective of users locations with respect to each other. The mobility data from personal devices is relatively easy to obtain and available for almost every inhabited corner of the globe because of the very wide and frequent use of smartphones by people. Moreover, the basic aspect of it makes it easy to collect and transfer.



(a) (b)

Figure 6.2: On 6.2a is mapped the measured values of entropy after being transferred on the hexagonal grid. Each cell of the grid is given the mean values of the entropy index of the amenities present in it. Places without amenities are plotted with a value of -1 in the map. On figure 6.2b one can see the number of visitors observed in each cell of the grid.

6.3 Context

It is interesting to look at the results of this work, in the context of research on the vibrancy of places in London (Sulis et al. 2018; Roth et al. 2011). One observation is that places found to be the most important and attractive for users on the city level did not necessarily show high local diversity values. Areas such as Shoreditch, the City, Liverpool streeet, Mayfair, St James Park are known to be very busy, but show up with lower diversity in the context of this work as opposed to such places as Marylebone, Hampstead, Islington, Pimlico or Fulham. The methodology proposed in this work finds the various villages that London is made of, while the previous works illustrate the places that link the people globally, in the context of economic, governance, or leisure or cultural activities. The results of this work thus seem complementary to the wider body of research done on the interactions of people on the whole city scale by providing this decentralised vision of centres of local activity, introduced here as neighbourhoods.

6.4 Overlap of the definitions

Finally, this work would be incomplete without an attempt to measure the level of overlap between the two definitions of neighbourhoods. The following method measures how much a neighbourhood defined from the point of view of amenities coincides to the neighbourhood obtained from mobility patterns of users. It is only verified for the given partition of 380 neighbourhoods. The method to do so is directly taken from basic set theory and corresponds to the definition of identity of sets. In this context, the verification is on whether the amenities assigned to an single amenities cluster will be contained in a single neighbourhood found from mobility. Simply speaking, do the neighbourhoods found by clustering amenities correspond to neighbourhoods defined through mobility.

Two sets are defined to be equal if they are contained within each other, $N_1 = N_2$, if $N_1 \subset N_2$ and $N_2 \subset N_1$. An appropriate measure to illustrate this is the Dice similarity coefficient, discussed above, $DSC = \frac{|N_1 \bigcap N_2|}{|N_1|+|N_2|}$. The DSC takes values between 0 and 1, 0 being an absence of any similarity, and 1 being the identity of sets. Investigating this relation on all the existing pairs of overlap between neighbourhoods we find that there is at least a 0.7 similarity of both definitions of neighbourhoods in 42% of communities (162 out of 380). And there is at least 0.6 similarity in 69% of the cases (262 out of 380). This result seems to support the idea that a community, or neighbourhood is formed around a cluster of amenities, with the diversity being the local centre as defined in section 4. However, this result needs to be nuanced by the fact that the urban layout and the way people use it is dependent on a multitude of factors, amenities being only one of them. Further visual investigation leads to the observation that different communities can be found around a single cluster of amenities, indicating that the pure overlap of the two areas might present only a part of the actual picture of how neighbourhoods form in terms of communities and their access the the local amenities. The hierarchy of the communities is also an important factor to consider and further work could look deeper into this by comparing the various partitions of the hierarchy presented in section 5.

6.5 Limitations and improvements

The link between diversity of amenities and visits at a location is not an indicator of the visits to the amenities themselves. These results hold as long as the data is exact, further validating steps could be taken by, for example, applying the methodology to a data set on amenities from another source. Further research could be done to link the accessibility in the street network of a place and the diversity of amenities present. Moreover, the preferential choice of location for amenities at places with high through movement, often modelled with edge betweenness or closeness centrality, can also be used as explanatory variable in future works.

Possible future extensions of the entropy index, could shed light on more details of the local composition of the amenity set and the differences across the city's neighbourhoods. The methodology presented in section 5 could be further extrapolated to the city scale interactions between users. The aim of this work is to focus on the area around one's home, which has lead to a 3 km threshold on the distance from home at which the overlap of users is considered. At this scale, the density of local residents and their activity in terms of signal emission has implications on the overall results. The presence of hyper active users, with 5 to 10 times more points emitted over the covered period could bring in high noise in the definition of neighbourhoods with such outliers becoming a neighbourhood of their own. To counter that effect, the spatial resolution of the grid needs to be chosen as to include a sufficient amount of visitors per cell. Future improvements could be made in the investigation of locations of interest on the neighbourhood scale, both in terms of frequency of visit, and time spent. Additionally, a bigger data set of users could produce more significant results, and allow to reduce the size of grid cells. The relatively low, hourly, time resolution of the data set is compensated by the high spatial precision of the GPS coordinates. Even if emitted within the same hour, they can be assigned to different places and thus effectively be distinguished.

7 Conclusion

In order to ameliorate our life in cities, it is essential to understand the various scales of human mobility and the hierarchies of places that we visit. The two visions pre-

sented for the neighborhood in this work are not intended to be seen in opposition to each other, the one proving the other wrong, but rather as complementary and illustrating a slightly different vision of a similar place in our mental map of the city. Both visions, however, could have useful interpretations when in comes to urban design. The amenities vision proposing neighbourhoods as the areas of attraction towards the local maximum of diversity, it can be useful to look further into increasing the accessibility to the amenities within the areas defined this way. The finding of an exponential relation linking the number of visitors of places around home to the diversity of the amenities in the place supports this, and is in line with what the planning practice has been moving towards in the recent decades. On the other hand, the neighbourhoods that show up as communities by investigating the mobility around home could be further applied to understand segregation and other social variables such a deprivation. Both methods showed a capacity to build neighbourhoods that are meaningful in terms of size, and spatially coherent. The first one being a bit speculative and relying on the vision of the neighbourhood in urban design, the second being strongly anchored in the patterns of human mobility. At the same time, more recent research in the field of urban sciences has allowed to find various links between the urban environment and the well being of individuals, the communities definition of neighbourhoods presented here could allow to better develop this topic and investigate whether the communities are somehow related to the well being of individuals.

7.1 Reproducibility

The codes to reproduce the work are available at the following link.

Bibliography

Alessandretti, Laura, Ulf Aslak, and Sune Lehmann (Nov. 19, 2020). “The scales of human mobility”. In: *Nature* 587.7834, pp. 402–407. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-020-2909-1. URL: http://www.nature.com/articles/s41586-020-2909-1 (visited on 08/05/2021).

Alessandretti, Laura, Piotr Sapiezynski, et al. (July 2018). “Evidence for a conserved quantity in human mobility”. In: *Nature Human Behaviour* 2.7, pp. 485–491. ISSN: 2397-3374. DOI: 10.1038/s41562-018-0364-x. URL: http://www.nature.com/articles/s41562-018-0364-x (visited on 08/05/2021).

Alonso, William (1964). *Location and land use : toward a general theory of land rent / William Alonso*. Cambridge, Mass.: Cambridge, Mass. : Harvard University Press. ISBN: 0-674-53700-9.

Aras, Rohan L., Nicholas T. Ouellette, and Rishee K. Jain (Jan. 14, 2021). “Automated identification of urban substructure for comparative analysis”. In: *PLOS ONE* 16.1. Ed. by Maxime Lenormand, e0245067. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0245067. URL: https://dx.plos.org/10.1371/journal.pone.0245067 (visited on 03/01/2021).

Arcaute, Elsa, Erez Hatna, et al. (Jan. 2015). “Constructing cities, deconstructing scaling laws”. In: *Journal of The Royal Society Interface* 12.102, p. 20140745. ISSN: 1742-5689, 1742-5662. DOI: 10.1098/rsif.2014.0745. URL: https://royalsocietypublishing.org/doi/10.1098/rsif.2014.0745 (visited on 06/22/2021).

Arcaute, Elsa, Carlos Molinero, et al. (Apr. 2016). “Cities and regions in Britain through hierarchical percolation”. In: *Royal Society Open Science* 3.4, p. 150691. ISSN: 2054-5703. DOI: 10.1098/rsos.150691. URL: https://royalsocietypublishing.org/doi/10.1098/rsos.150691 (visited on 07/23/2021).

Barbosa, Hugo et al. (Dec. 1, 2020). “Uncovering the socioeconomic facets of human mobility”. In: *arXiv:2012.00838 [physics, stat]*. arXiv: 2012.00838. URL: http://arxiv.org/abs/2012.00838 (visited on 05/02/2021).

Barthelemy, Marc (Mar. 2017). “From paths to blocks: New measures for street patterns”. In: *Environment and Planning B: Urban Analytics and City Science* 44.2,

pp. 256–271. ISSN: 2399-8083, 2399-8091. DOI: 10.1177/0265813515599982. URL: http://journals.sagepub.com/doi/10.1177/0265813515599982 (visited on 07/29/2021).

Barthelemy, Marc (June 2019). "The statistical physics of cities". In: *Nature Reviews Physics* 1.6, pp. 406–415. ISSN: 2522-5820. DOI: 10.1038/s42254-019-0054-2. URL: http://www.nature.com/articles/s42254-019-0054-2 (visited on 06/22/2021).

Barthélemy, Marc (Feb. 2011). "Spatial networks". In: *Physics Reports* 499.1, pp. 1–101. ISSN: 03701573. DOI: 10.1016/j.physrep.2010.11.002. URL: https://linkinghub.elsevier.com/retrieve/pii/S037015731000308X (visited on 07/02/2021).

Bassolas, Aleix et al. (Dec. 2019). "Hierarchical organization of urban mobility and its connection with city livability". In: *Nature Communications* 10.1, p. 4817. ISSN: 2041-1723. DOI: 10.1038/s41467-019-12809-y. URL: http://www.nature.com/articles/s41467-019-12809-y (visited on 02/25/2021).

Bazzani, Armando et al. (May 5, 2010). "Statistical laws in urban mobility from microscopic GPS data in the area of Florence". In: *Journal of Statistical Mechanics: Theory and Experiment* 2010.5, P05001. ISSN: 1742-5468. DOI: 10.1088/1742-5468/2010/05/P05001. URL: https://iopscience.iop.org/article/10.1088/1742-5468/2010/05/P05001 (visited on 08/17/2021).

Berry, Brian J. L. (1967). *Geography of market centers and retail distribution / Brian J.L. Berry.* Geography of market centres and retail distribution. Englewood Cliffs, N.J.: Englewood Cliffs, N.J.

Bettencourt, L. M. A. et al. (Apr. 24, 2007). "Growth, innovation, scaling, and the pace of life in cities". In: *Proceedings of the National Academy of Sciences* 104.17, pp. 7301–7306. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0610172104. URL: http://www.pnas.org/cgi/doi/10.1073/pnas.0610172104 (visited on 07/22/2021).

Cao, Wenpu et al. (May 2020). "Quantifying urban areas with multi-source data based on percolation theory". In: *Remote Sensing of Environment* 241, p. 111730. ISSN: 00344257. DOI: 10.1016/j.rse.2020.111730. URL: https://linkinghub.elsevier.com/retrieve/pii/S0034425720301000 (visited on 08/12/2021).

Chaskin, Robert J. (1997). “Perspectives on Neighborhood and Community: A Review of the Literature”. In: *Social Service Review* 71.4. Publisher: University of Chicago Press, pp. 521–547. ISSN: 00377961, 15375404. URL: http://www.jstor.org/stable/30012640 (visited on 07/22/2021).

Clauset, Aaron, Cosma Rohilla Shalizi, and M. E. J. Newman (Nov. 4, 2009). “Power-Law Distributions in Empirical Data”. In: *SIAM Review* 51.4, pp. 661–703. ISSN: 0036-1445, 1095-7200. DOI: 10.1137/070710111. URL: http://epubs.siam.org/doi/10.1137/070710111 (visited on 08/22/2021).

Cottineau, Clémentine and Elsa Arcaute (Dec. 2020). “The nested structure of urban business clusters”. In: *Applied Network Science* 5.1, p. 2. ISSN: 2364-8228. DOI: 10.1007/s41109-019-0246-9. URL: https://appliednetsci.springeropen.com/articles/10.1007/s41109-019-0246-9 (visited on 10/23/2020).

Cottineau, Clémentine, Olivier Finance, et al. (Nov. 2019). “Defining urban clusters to detect agglomeration economies”. In: *Environment and Planning B: Urban Analytics and City Science* 46.9, pp. 1611–1626. ISSN: 2399-8083, 2399-8091. DOI: 10.1177/2399808318755146. URL: http://journals.sagepub.com/doi/10.1177/2399808318755146 (visited on 10/19/2020).

Creatore, Maria I. et al. (May 24, 2016). “Association of Neighborhood Walkability With Change in Overweight, Obesity, and Diabetes”. In: *JAMA* 315.20, p. 2211. ISSN: 0098-7484. DOI: 10.1001/jama.2016.5898. URL: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2016.5898 (visited on 05/08/2021).

De Nadai, Marco et al. (Apr. 11, 2016). “The Death and Life of Great Italian Cities: A Mobile Phone Data Perspective”. In: *Proceedings of the 25th International Conference on World Wide Web*. WWW ’16: 25th International World Wide Web Conference. Montréal Québec Canada: International World Wide Web Conferences Steering Committee, pp. 413–423. ISBN: 978-1-4503-4143-1. DOI: 10.1145/2872427.2883084. URL: https://dl.acm.org/doi/10.1145/2872427.2883084 (visited on 06/24/2021).

Dong, Lei, Sicong Chen, et al. (Dec. 2017). “Measuring economic activity in China with mobile big data”. In: *EPJ Data Science* 6.1, p. 29. ISSN: 2193-1127. DOI: 10.

1140/epjds/s13688-017-0125-5. URL: http://epjdatascience.springeropen.com/articles/10.1140/epjds/s13688-017-0125-5 (visited on 06/01/2021).

Dong, Lei, Carlo Ratti, and Siqi Zheng (July 30, 2019). "Predicting neighborhoods' socioeconomic attributes using restaurant data". In: *Proceedings of the National Academy of Sciences* 116.31, pp. 15447–15452. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1903064116. URL: http://www.pnas.org/lookup/doi/10.1073/pnas.1903064116 (visited on 05/10/2021).

Ehrenhalt, Alan (2012). *The great inversion and the future of the American city*. 1st ed. New York: Knopf. 276 pp. ISBN: 978-0-307-27274-4.

Fleury, Vanessa et al. (Feb. 2021). "Geospatial analysis of individual-based Parkinson's disease data supports a link with air pollution: A case-control study". In: *Parkinsonism & Related Disorders* 83, pp. 41–48. ISSN: 13538020. DOI: 10.1016/j.parkreldis.2020.12.013. URL: https://linkinghub.elsevier.com/retrieve/pii/S135380202100002X (visited on 05/14/2021).

Florida, Richard L. (2012). *The rise of the creative class: revisited*. New York: Basic Books. 483 pp. ISBN: 978-0-465-02993-8 978-0-465-02995-2.

Gallotti, Riccardo, Giulia Bertagnolli, and Manlio De Domenico (Dec. 2021). "Unraveling the hidden organisation of urban systems and their mobility flows". In: *EPJ Data Science* 10.1, p. 3. ISSN: 2193-1127. DOI: 10.1140/epjds/s13688-020-00258-3. URL: https://epjdatascience.springeropen.com/articles/10.1140/epjds/s13688-020-00258-3 (visited on 03/01/2021).

Giannotti, Fosca et al. (Oct. 2011). "Unveiling the complexity of human mobility by querying and mining massive trajectory data". In: *The VLDB Journal* 20.5, pp. 695–719. ISSN: 1066-8888, 0949-877X. DOI: 10.1007/s00778-011-0244-8. URL: http://link.springer.com/10.1007/s00778-011-0244-8 (visited on 08/17/2021).

Glaeser, Edward L (2011). *Triumph of the city: how our greatest invention makes us richer, smarter, greener, healthier, and happier*. OCLC: 759841377. ISBN: 978-1-101-46733-6 978-1-101-47567-6. URL: http://search.ebscohost.com/login.

aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1117736 (visited on 05/19/2021).

Glazier, Richard H. et al. (Jan. 14, 2014). "Density, Destinations or Both? A Comparison of Measures of Walkability in Relation to Transportation Behaviors, Obesity and Diabetes in Toronto, Canada". In: *PLoS ONE* 9.1. Ed. by Reury F. P. Bacurau, e85295. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0085295. URL: https://dx.plos.org/10.1371/journal.pone.0085295 (visited on 05/08/2021).

González, Marta C., César A. Hidalgo, and Albert-László Barabási (June 5, 2008). "Understanding individual human mobility patterns". In: *Nature* 453.7196, pp. 779–782. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature06958. URL: http://www.nature.com/articles/nature06958 (visited on 08/05/2021).

Guest, Avery M. and Barrett A. Lee (1984). "How Urbanites Define Their Neighborhoods". In: *Population and Environment* 7.1. Publisher: Springer, pp. 32–56. ISSN: 01990039, 15737810. URL: http://www.jstor.org/stable/27503013 (visited on 08/06/2021).

Hidalgo, César A., Elisa Castañer, and Andres Sevtsuk (Dec. 2020). "The amenity mix of urban neighborhoods". In: *Habitat International* 106, p. 102205. ISSN: 01973975. DOI: 10.1016/j.habitatint.2020.102205. URL: https://linkinghub.elsevier.com/retrieve/pii/S0197397519311336 (visited on 05/10/2021).

Iravani, Hamid and Venkat Rao (Mar. 3, 2020). "The effects of New Urbanism on public health". In: *Journal of Urban Design* 25.2, pp. 218–235. ISSN: 1357-4809, 1469-9664. DOI: 10.1080/13574809.2018.1554997. URL: https://www.tandfonline.com/doi/full/10.1080/13574809.2018.1554997 (visited on 01/12/2021).

Jacobs, Jane (1962). *The death and life of great American cities / Jane Jacobs.* London: London : Jonathan Cape.

Jang, Myungjun and Chang-Deok Kang (Oct. 2015). "Retail accessibility and proximity effects on housing prices in Seoul, Korea: A retail type and housing submarket approach". In: *Habitat International* 49, pp. 516–528. ISSN: 01973975. DOI: 10.1016/j.habitatint.2015.07.004. URL: https://linkinghub.elsevier.com/retrieve/pii/S0197397515001393 (visited on 09/10/2021).

Jiang, Bin, Junjun Yin, and Sijian Zhao (Aug. 31, 2009). "Characterizing the human mobility pattern in a large street network". In: *Physical Review E* 80.2, p. 021136. ISSN: 1539-3755, 1550-2376. DOI: 10.1103/PhysRevE.80.021136. URL: https://link.aps.org/doi/10.1103/PhysRevE.80.021136 (visited on 08/17/2021).

Laurence, Peter L (2006). "The Death and Life of Urban Design: Jane Jacobs, The Rockefeller Foundation and the New Research in Urbanism, 1955-1965". In: *Journal of urban design* 11.2. Publisher: Taylor & Francis Group, pp. 145–172. ISSN: 1357-4809. DOI: 10.1080/13574800600644001.

Listerborn, Carina (Sept. 2007). "Who speaks? And who listens? The relationship between planners and women's participation in local planning in a multi-cultural urban environment". In: *GeoJournal* 70.1, pp. 61–74. ISSN: 0343-2521, 1572-9893. DOI: 10.1007/s10708-007-9114-8. URL: http://link.springer.com/10.1007/s10708-007-9114-8 (visited on 11/24/2020).

Louail, Thomas et al. (May 2015). "From mobile phone data to the spatial structure of cities". In: *Scientific Reports* 4.1, p. 5276. ISSN: 2045-2322. DOI: 10.1038/srep05276. URL: re (visited on 05/02/2021).

Louf, Rémi and Marc Barthelemy (Nov. 6, 2013). "Modeling the Polycentric Transition of Cities". In: *Physical Review Letters* 111.19, p. 198702. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.111.198702. URL: https://link.aps.org/doi/10.1103/PhysRevLett.111.198702 (visited on 01/26/2021).

— (Dec. 6, 2014). "A typology of street patterns". In: *Journal of The Royal Society Interface* 11.101, p. 20140924. ISSN: 1742-5689, 1742-5662. DOI: 10.1098/rsif.2014.0924. URL: https://royalsocietypublishing.org/doi/10.1098/rsif.2014.0924 (visited on 04/05/2021).

— (May 2015). "How congestion shapes cities: from mobility patterns to scaling". In: *Scientific Reports* 4.1, p. 5561. ISSN: 2045-2322. DOI: 10.1038/srep05561. URL: http://www.nature.com/articles/srep05561 (visited on 01/26/2021).

Lynch, Kevin (1960). *The image of the city / Kevin Lynch.* Cambridge, Mass.: Cambridge, Mass. ISBN: 0-262-12004-6.

Masucci, A. Paolo, Kiril Stanilov, and Michael Batty (Jan. 13, 2014). "Exploring the evolution of London's street network in the information space: A dual approach". In: *Physical Review E* 89.1, p. 012805. ISSN: 1539-3755, 1550-2376. DOI: 10.1103/PhysRevE.89.012805. URL: https://link.aps.org/doi/10.1103/PhysRevE.89.012805 (visited on 06/26/2021).

Moran, Daniel et al. (June 1, 2018). "Carbon footprints of 13 000 cities". In: *Environmental Research Letters* 13.6, p. 064041. ISSN: 1748-9326. DOI: 10.1088/1748-9326/aac72a. URL: https://iopscience.iop.org/article/10.1088/1748-9326/aac72a (visited on 07/24/2021).

Park, Robert Ezra (1936). "Human Ecology". In: *American Journal of Sociology* 42.1. Publisher: University of Chicago Press, pp. 1–15. ISSN: 00029602, 15375390. URL: http://www.jstor.org/stable/2768859 (visited on 07/23/2021).

Prieto Curiel, Rafael et al. (Feb. 24, 2021). "The heartbeat of the city". In: *PLOS ONE* 16.2. Ed. by Quan Yuan, e0246714. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0246714. URL: https://dx.plos.org/10.1371/journal.pone.0246714 (visited on 05/10/2021).

Riccardo, Gallotti, Bazzani Armando, and Rambaldi Sandro (Sept. 2012). "TOWARDS A STATISTICAL PHYSICS OF HUMAN MOBILITY". In: *International Journal of Modern Physics C* 23.9, p. 1250061. ISSN: 0129-1831, 1793-6586. DOI: 10.1142/S0129183112500611. URL: https://www.worldscientific.com/doi/abs/10.1142/S0129183112500611 (visited on 08/17/2021).

Roth, Camille et al. (Jan. 7, 2011). "Structure of Urban Movements: Polycentric Activity and Entangled Hierarchical Flows". In: *PLoS ONE* 6.1. Ed. by Matjaz Perc, e15923. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0015923. URL: https://dx.plos.org/10.1371/journal.pone.0015923 (visited on 09/05/2021).

Salesses, Philip, Katja Schechtner, and César A. Hidalgo (July 24, 2013). "The Collaborative Image of The City: Mapping the Inequality of Urban Perception". In: *PLoS ONE* 8.7. Ed. by Alain Barrat, e68400. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0068400. URL: https://dx.plos.org/10.1371/journal.pone.0068400 (visited on 06/24/2021).

Schläpfer, Markus et al. (May 27, 2021). “The universal visitation law of human mobility”. In: *Nature* 593.7860, pp. 522–527. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-021-03480-9. URL: http://www.nature.com/articles/s41586-021-03480-9 (visited on 06/01/2021).

Sevtsuk, Andres (Dec. 2014). “Location and Agglomeration: The Distribution of Retail and Food Businesses in Dense Urban Environments”. In: *Journal of Planning Education and Research* 34.4, pp. 374–393. ISSN: 0739-456X, 1552-6577. DOI: 10.1177/0739456X14550401. URL: http://journals.sagepub.com/doi/10.1177/0739456X14550401 (visited on 05/10/2021).

Shannon, Claude Elwood (1949). *The mathematical theory of communication / by Claude E. Shannon and Warren Weaver*. Red. by Warren Weaver. Urbana: Urbana : University of Illinois Press.

Strano, Emanuele et al. (Dec. 2012). “Elementary processes governing the evolution of road networks”. In: *Scientific Reports* 2.1, p. 296. ISSN: 2045-2322. DOI: 10.1038/srep00296. URL: http://www.nature.com/articles/srep00296 (visited on 04/05/2021).

Sulis, Patrizia et al. (June 23, 2018). “Using mobility data as proxy for measuring urban vitality”. In: *Journal of Spatial Information Science* 16, pp. 137–162. ISSN: 1948-660X. DOI: 10.5311/JOSIS.2018.16.384. URL: http://josis.org/index.php/josis/article/view/384 (visited on 06/24/2021).

Sung, Hyungun, Sugie Lee, and SangHyun Cheon (June 2015). “Operationalizing Jane Jacobs’s Urban Design Theory: Empirical Verification from the Great City of Seoul, Korea”. In: *Journal of Planning Education and Research* 35.2, pp. 117–130. ISSN: 0739-456X, 1552-6577. DOI: 10.1177/0739456X14568021. URL: http://journals.sagepub.com/doi/10.1177/0739456X14568021 (visited on 06/24/2021).

Tuan, Yi-Fu (2011). *Space and place: the perspective of experience*. 7. print. Minneapolis, Minn.: Univ. of Minnesota Press. 235 pp. ISBN: 978-0-8166-3877-2.

Weiss, Linda et al. (June 2007). “Defining Neighborhood Boundaries for Urban Health Research”. In: *American Journal of Preventive Medicine* 32.6, S154–S159. ISSN:

07493797. DOI: 10.1016/j.amepre.2007.02.034. URL: https://linkinghub.elsevier.com/retrieve/pii/S0749379707001390 (visited on 07/22/2021).

Xu, Yanyan et al. (Sept. 2020). “Deconstructing laws of accessibility and facility distribution in cities”. In: *Science Advances* 6.37, eabb4112. ISSN: 2375-2548. DOI: 10.1126/sciadv.abb4112. URL: https://advances.sciencemag.org/lookup/doi/10.1126/sciadv.abb4112 (visited on 02/25/2021).

Zhong, Chen et al. (Feb. 2017). “Revealing centrality in the spatial structure of cities from human activity patterns”. In: *Urban Studies* 54.2, pp. 437–455. ISSN: 0042-0980, 1360-063X. DOI: 10.1177/0042098015601599. URL: http://journals.sagepub.com/doi/10.1177/0042098015601599 (visited on 08/12/2021).

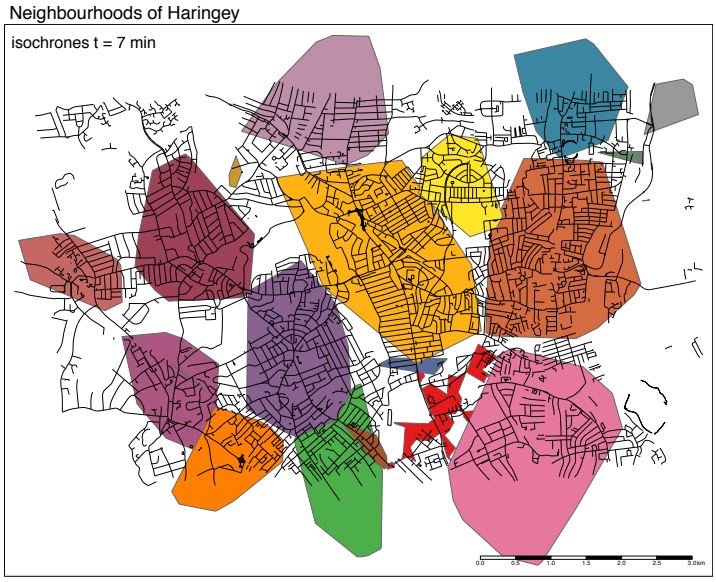
A Appendix

A.1 Selection of amenities

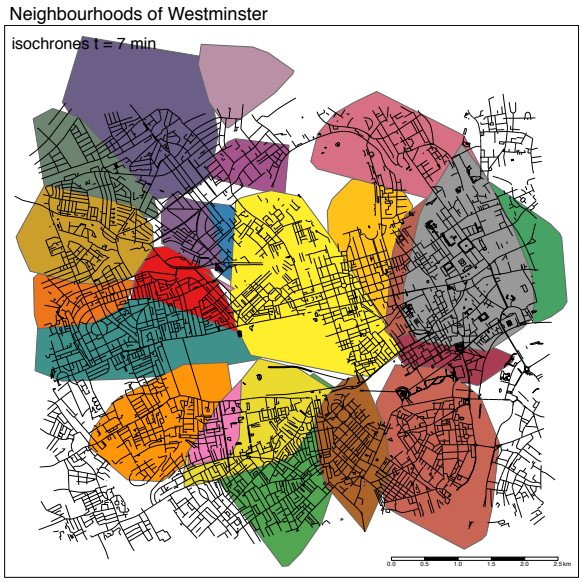
	amenities_1	n_1	amenities_2	n_2	amenities_3	n_3
1	restaurant	7687	clinic	357	musical_instrument	80
2	cafe	5502	art	329	perfumery	80
3	fast_food	4785	childcare	313	nursing_home	73
4	convenience	4131	books	309	internet_cafe	58
5	pub	3780	bicycle	290	townhall	58
6	school	3682	taxi	288	events_venue	57
7	place_of_worship	3431	deli	261	courthouse	52
8	hairdresser	3305	theatre	257	paint	50
9	clothes	3041	greengrocer	232	hifi	43
10	beauty	1613	gallery	221	coffee	40
11	hotel	1355	sports	218	chocolate	39
12	supermarket	1348	museum	215	public_bookcase	38
13	pharmacy	1322	college	210	pastry	37
14	recycling	1106	veterinary	199	appliance	32
15	bank	1090	hospital	191	tea	32
16	artwork	1009	computer	175	tobacco	32
17	community_centre	924	chemist	174	cheese	24
18	bookmaker	834	police	172	boutique	23
19	dentist	779	university	154	conference_centre	20
20	post_office	700	department_store	151	general	20
21	doctors	634	nightclub	151	language_school	18
22	jewelry	606	pawnbroker	149	crematorium	17
23	bakery	565	health_food	140	food_court	16
24	optician	538	marketplace	135	beverages	15
25	mobile_phone	520	fire_station	133	camera	15
26	kindergarten	515	cinema	126	music_school	15
27	shoes	505	studio	122	embassy	9
28	library	455	photo	104	monastery	7
29	alcohol	403	public_building	92	model	6
30	butcher	394	toys	91	organic	5
31	doityourself	373	ice_cream	90	planetarium	1
32	florist	363	seafood	89	spices	1

Table A.1: Summary table of amenities of London according to OSM map feature tags.

A.2 Clustering



(a) Haringey



(b) Westminster

Figure A.1: Clusters of amenities in selected boroughs

A.3 Entropy

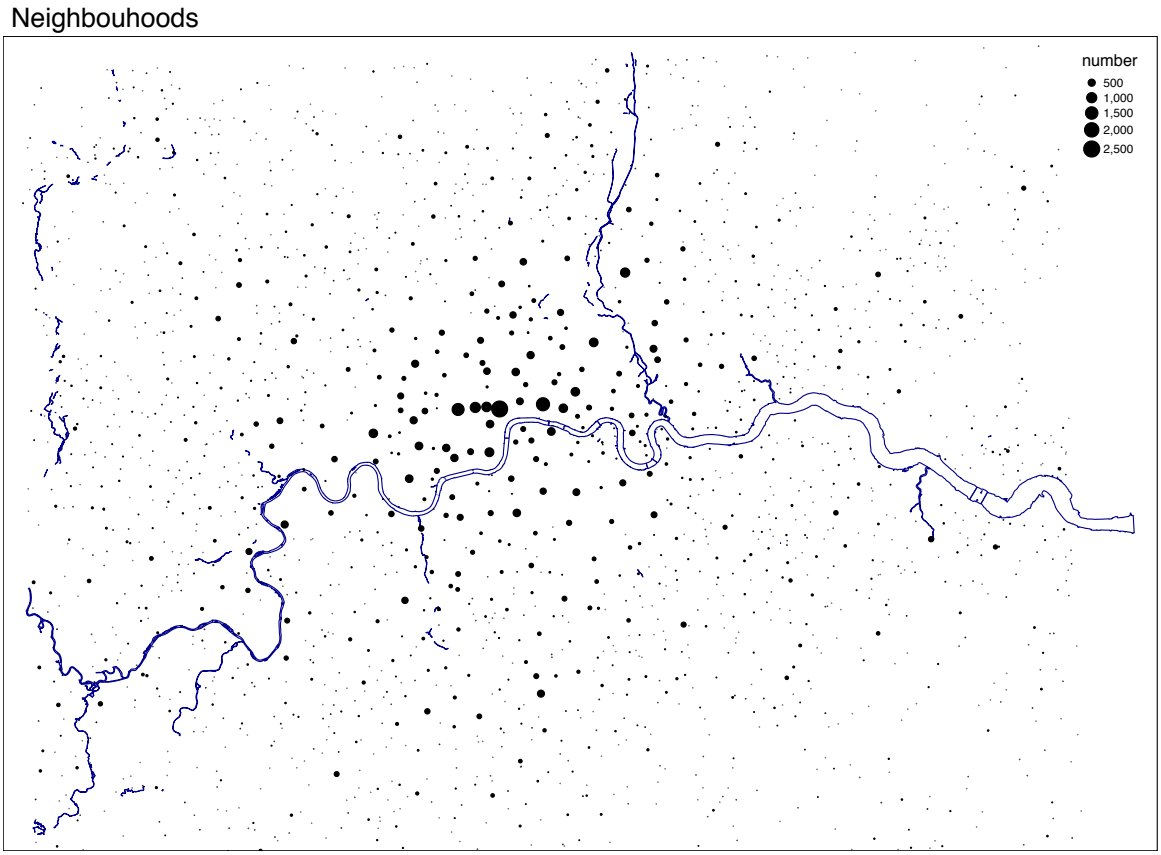
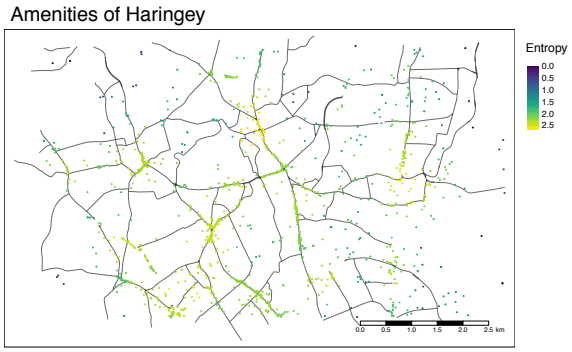


Figure A.2: Alternative visualisation of the neighbourhoods from amenities by size of each cluster.

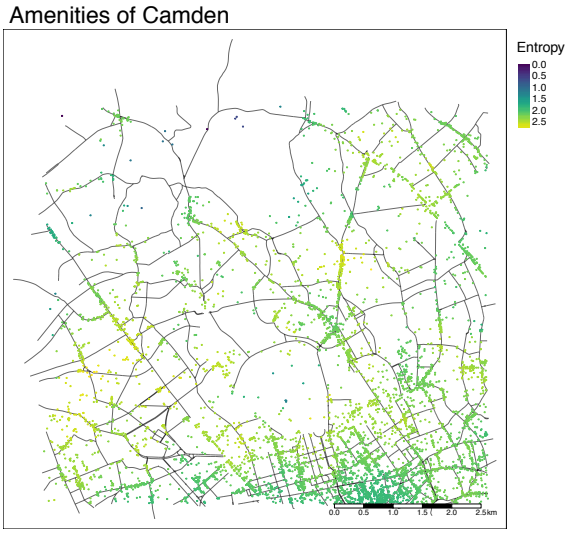
A.4 Online supplementary material

An online version of the map of the neighbourhoods from mobility data can be found at: https://rpubs.com/ischlo/london_neighbourhoods

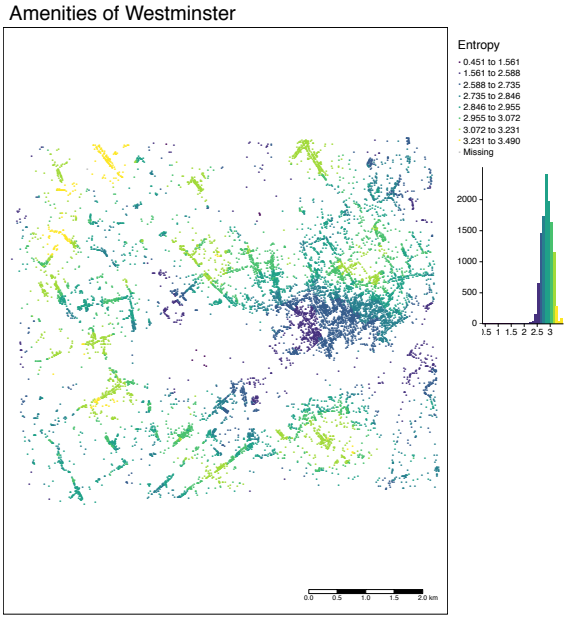
Further online material can be found here:https://rpubs.com/ischlo/online_supp.



(a) Entropy index of amenities in the borough of Haringey.



(b) Entropy index of amenities in the borough of Camden.

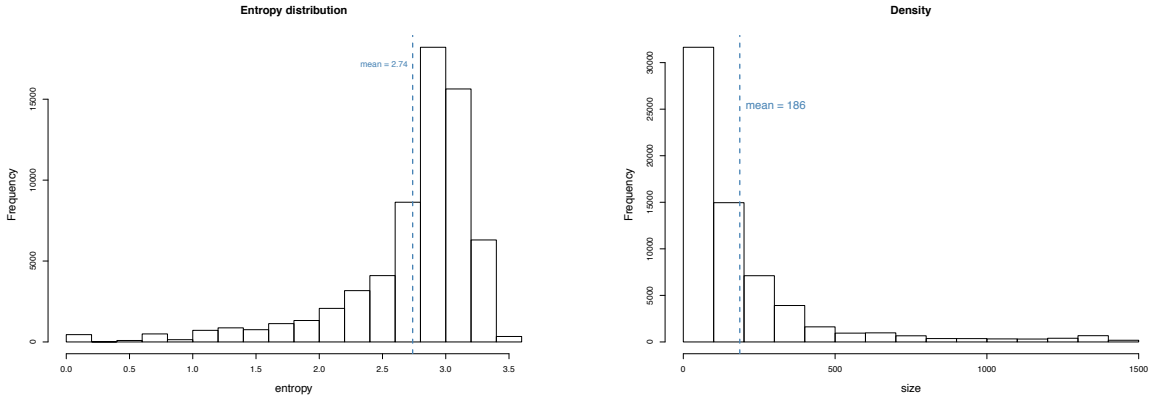


(c) Entropy index of amenities in the borough of Westminster.

Figure A.3: Entropy index of amenities for selected boroughs.

A.5 The grid

The grid looks the following way when zoomed at :



(a) (b)

Figure A.4

Figure A.5: Illustration of the hexagonal grid with non residential cells merged to the nearest residential one.

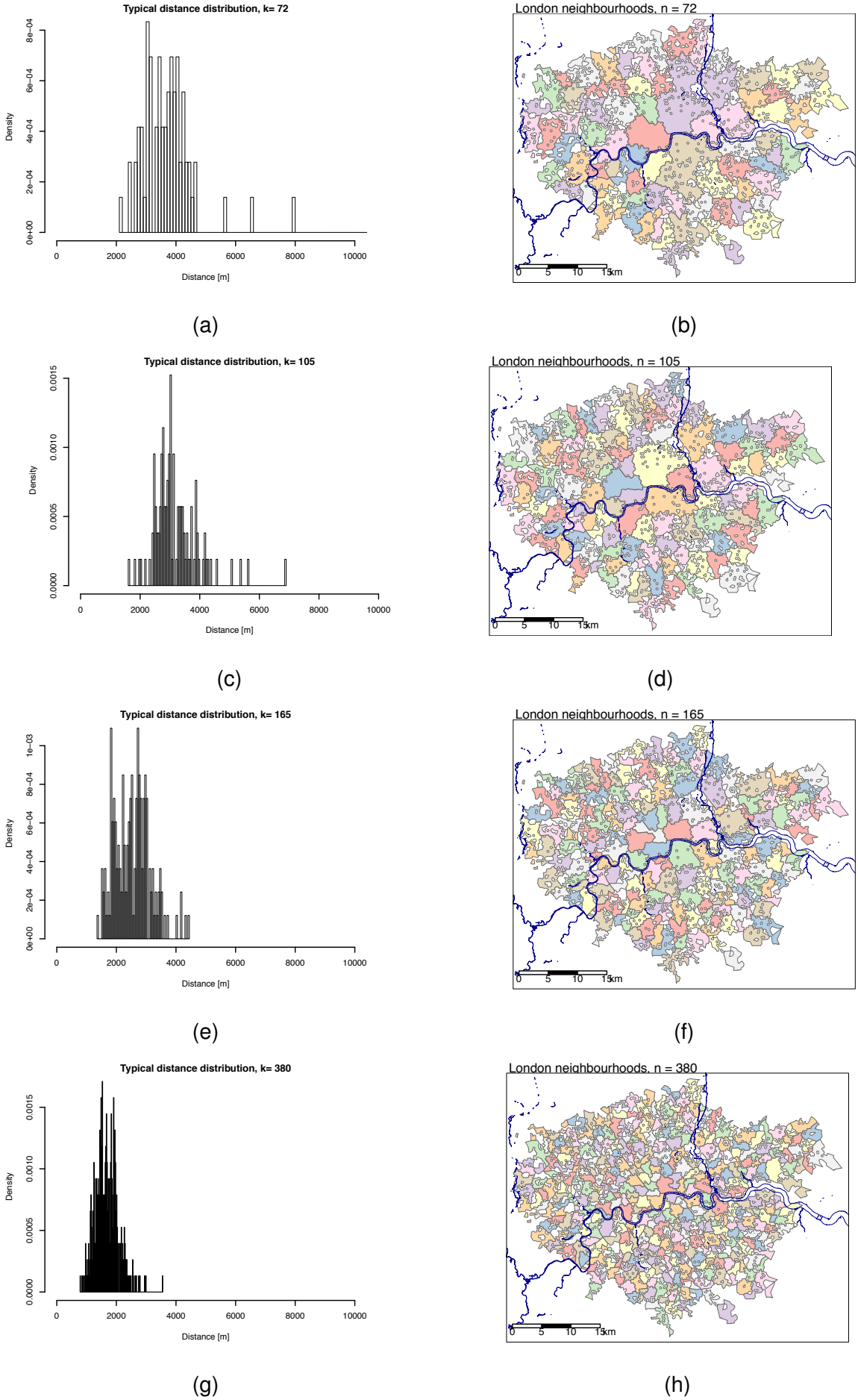


Figure A.6: Percolation on mobility