UNIVERSITY OF LEEDS

THE TALE OF TWO CITIES: DIVERSITY IN LEEDS AND WARSAW

Working Paper 11/01

Aneta Piekut
Philip Rees
Gill Valentine
Marek Kupiszewski

Version 1.0

School of Geography, University of Leeds, Leeds LS2 9JT
April 2011

This Working Paper is an online publication and may be revised.

CONTACT DETAILS:

Aneta Piekut

Postdoctoral Research Fellow

School of Geography, University of Leeds, Leeds LS2 9JT, United Kingdom

Email: a.piekut@leeds.ac.uk

tel. +44 (0) 113 343 3327

web: http://www.geog.leeds.ac.uk/index.php?id=1277

Philip Rees

Emeritus Professor Philip Rees FRGS, FBA, CBE

School of Geography, University of Leeds, Leeds LS2 9JT, United Kingdom

email: p.h.rees@leeds.ac.uk

tel. +44 (0)113 343 3341

web: http://www.geog.leeds.ac.uk/people/p.rees/

Gill Valentine

Professor and Head of School

School of Geography, University of Leeds, Leeds LS2 9JT, United Kingdom

Email: g.valentine@leeds.ac.uk

tel. +44 (0) 113 343 3396

web: http://www.geog.leeds.ac.uk/people/g.valentine/

Marek Kupiszewski

Institute of Geography and Spatial Organization, Polish Academy of Sciences, Twarda 51/55, 00-818 Warsaw, Poland

Email: m.kupisz@cefmr.pan.pl

tel. +48 22 697 89 41

web: http://www.cefmr.pan.pl/staff/mk.html

TABLE OF CONTENTS:

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Communities across the world are becoming more diverse because of international migration, population ageing, residential mobility and life style choices. These processes result in communities differing in their diversity. In a European Research Council (ERC) project called *Living with Difference*, we seek to understand how people respond to living with others who are different from themselves. To provide a context for a survey and in-depth interviews, we prepared a description of residential community diversity in two cities, Leeds and Warsaw, using census data (UK 2001, Poland 2002). We used selected variables to represent in both cities the key social dimensions of difference: demographic, socio-economic, ethnic and disability. A standard cluster analysis using a k-means algorithm was implemented for each city separately and for the two cities combined. The paper presents the arguments for cluster selection, maps the results at community scale (Community Areas in Leeds, Urban Regions in Warsaw) and graphs the clusters in four dimensions. Different diversity clusters are identified in each city and the combined analysis shows how communities in one city compare with those in the other. We will use the diversity clusters with other more up to date population information to stratify the survey and to situate the interviews in their community context. The paper illustrates the advantages of collaboration between quantitative and qualitative social scientists.

ACKNOWLEDGEMENTS

The research reported here is supported by Advanced Investigator Award to Professor Gill Valentine from the European Research Council (ERC), grant agreement no 249658, 1 June 2010 – 31 May 2013. The research contributes to Project A "Mapping Social Diversity", within the *Live with Difference in Europe: making communities out of strangers in an era of super mobility and super diversity* programme. We would like to express our gratitude to Dr. Agata Górny and the Centre of Migration Research, University of Warsaw who supported us with 2002 Census database for Statistical Regions in Warsaw, and Assoc. Prof. Przemysław Śleszyński, Polish Academy of Sciences, Institute of Geography and Spatial Organization, who provided a map of Urban Regions. The Leeds Census data for 2001 were provided through the ESRC Census Programme's Census Data Unit at the University of Manchester. The Census data were produced by the Office for National Statistics and are Crown Copyright. Map boundaries for Community Areas derive from digital boundary data from ESRC Census Programme's UKBORDERS service at the University of Edinburg and were provided by Prof. John Stillwell of the School of Geography. Dr. Jo Sadgrove provided valuable advice on restructuring of the literature review. For more information on the "Living with Difference" project, see the following web pages: http://www.geog.leeds.ac.uk/projects/livedifference/.

CONTRIBUTIONS OF THE AUTHORS

Aneta Piekut carried out the analysis and described in the paper in collaboration with Phil Rees. Together they drafted the text of the paper.

Gill Valentine quality assured and accepted the review of the theory.

Marek Kupiszewski quality assured the representation of the social geography of Warsaw and proposed the combined cluster analysis.

1. INTRODUCTION

1.1. Globalization, migration and diversity

Today we live in a world that is becoming even more inter-connected. We are witnessing globalisation processes operating through increasing trade between countries, through the world wide flows of capital and money, and through the movement of workers, their dependants and refugees responding to disparities in development between countries, to the need for circulation of managers and scientists, to the stagnation of workforce populations in some countries and continued expansion in others, and to political and environmental crises which displace people.

It is estimated that the share of international migrants (as a stock[1]) in the world's population is about 3% or 214 million (UN 2009; SASI 2010). Minority ethnic groups[2] make up a much higher percentage of the populations of richer countries whose populations are growing slowly and much smaller percentages in the poorer countries whose populations are still growing fast. In the United Kingdom (a high income country) the share of minority ethnic group is about 10% (ONS 2010a[3]), while in Poland (a middle income country) it is less than 1% (Office for Foreigners 2010). Both countries are experiencing increases in their foreign population, in response to the globalisation processes outlined earlier. In both societies people are living with others who have different histories of migration, language, religion and culture.

However, Vertovec (2006) has commented that:

> *Observing ethnicity or country of origin (...) provides a misleading, one-dimensional appreciation of contemporary diversity. Over the past ten years, the nature of immigration to Britain has brought with it a transformative 'diversification of diversity' not just in terms of ethnicities and countries of origin, but also with respect to a variety of significant variables that affect where, how and with whom people live* (Vertovec 2006).

We infer from this quotation two things. The first is that the sources of immigrants are becoming even more diverse in the countries of destination. The second is that the characteristics of newcomer groups themselves are also extremely diverse. Within and between groups there is a tremendous range of occupational success and income rewards, and there is a huge variety of family and age structures. So for example, the Chinese community in Britain includes high achieving and talented research postgraduates at one extreme and unskilled and exploited shellfish gatherers at the other (Southern

[1] We use stock instead of flow in order to capture static population of people who do not live in the countries of their original citizenship or birth.
[2] In the text we use 'minority ethnic group' when we describe people of different ethnicity or nationality. However, in Poland an 'ethnicity' has a different meaning and the term 'minority ethnic' refers to groups who do not identify themselves with any national state. So for the Polish context we think about people of different nationality.
[3] http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=15147.

Standard 2010). The social dimensions of class and family produce communities which can be different both within and between groups differentiated on the basis of the international migration history. Immigrant groups are also differentiated in terms of religious belief, even when originating in the same country (such as India).

In the paper we draw from two approaches to study social diversity. On a one hand, we use qualitative data and mapping techniques and we refer to previous studies where residential patterns of population were analysed. These were mainly studies where diversity was perceived through race or ethnicity and social class differences. On the other hand, we incorporate broader understanding of the diversity from the qualitative studies embracing other important characteristics of 'difference'.

Diversity, indeed, extends beyond the race/ethnicity and social class dimensions to also encompass gender differences, differences in sexual orientation and differences in (dis)ability. Men and women, although bound together in the institution of the family, fare differently in the world of work. Although the acceptance for the same-sex couples is growing in both studied countries, persons identifying themselves as gay, lesbian or bisexual face considerable prejudice and discriminatory attitudes in a dominantly heterosexual society (for the UK see Cowan 2007; for Poland see Abramowicz 2007; CBOS 2010). People who are born disabled or who become disabled through injury or disease or through the process of biological ageing face difficulties in pursuing their lives. Their disability may affect their capacity for work and may define the need for care from family members or from social institutions.

This list of dimensions of difference forms the agenda of the Equality and Human Rights Commission (EHRC) established by Equality Act in 2006, and working since 2007, created through a merger of separate commissions on Racial Inequality, Gender Opportunities and Disability Rights. The Commissions and their EHRC successor are charged with monitoring disadvantages faced by different groups, establishing whether these disadvantages constitute unfair discrimination against members of a different group and then initiating negotiations with the body to remedy the discrimination. Prosecutions may be brought against offending parties under equal rights laws.

Equality laws and institutions are not so well established in Poland. The Equality Act came into force in 1st January 2011. According to the Act and other regulations, the Commissioner for Protection of Civil Rights (*Rzecznik Praw Obywatelskich* or Ombudsman) is the first port of call in case of discrimination. The Racism and Xenophobia Monitoring Group (*Zespół Monitorowania Rasizmu i Ksenofobii*) works under the auspices of Ministry of Internal Affairs and Administration. Additionally, there is the Government Commissioner for Equal Treatment who also monitors discrimination by age, gender, religious, nationality or sexual orientation, but the Commissioner has

limited legal powers (she analyses legal acts, cooperates with other institutions, indicates problems). The Commissioner's role was to prepare government policy in anticipation of the Equality Act.

The base for equal and non-discriminatory treatment is the Treaty on European Union and the Treaty on the Functioning of the European Union (Official Journal C 83 of 30.3.2010)[4]. There are two main directives that define a set of principles in this field: The *Racial Equality Directive* (2000/43/EC) against discrimination on grounds of race and ethnic origin and The *Employment Framework Directive* (2000/78/EC) against discrimination at work on grounds of religion or belief, disability, age or sexual orientation.

What is meant by diversity? By diversity we mean the degree to which communities are comprised of different populations. Diversity is a function of the number of groups that live in an area and the distribution of the population across these groups. We employ a broad understanding of difference to encompass differences in age, sexuality, ethnicity, religion, socio-economic status and in degrees of (dis)ability (see section 2 for a full discussion).

1.2. The nature of our two cities

We study two European cities, Leeds and Warsaw, centrally located in Western and Central Europe and linked in the last decade by two way migration. The two cities have characteristic which make them examples of both the highly ethnically diverse cities in Western Europe in case of Leeds and the more ethnically homogenous cities of Central Europe in case of Warsaw. Leeds is a city with a population recorded as 715 thousand in the 2001 Census (ONS 2003), estimated to reach 760 thousand in the mid-year estimates for 2009 (ONS 2010a) and 730 thousand in all populations ethnic projection analysis (Wohland et al. 2010). The main minority ethnic groups come from Asia and are Indian and Pakistani, and are followed by Other White and Black Caribbean ethnicities. Key indicators are age, occupation, ethnicity and disability for Leeds and the UK as a whole are set out in Table 1.1. These show that Leeds lies close to the national average in all dimensions.

[4] Article 19: "Without prejudice to the other provisions of the Treaties and within the limits of the powers conferred by them upon the Union, the Council, acting unanimously in accordance with a special legislative procedure and after obtaining the consent of the European Parliament, may take appropriate action to combat discrimination based on sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation.". For a consolidated version of the treaty see:
http://eur-lex.europa.eu/JOHtml.do?uri=OJ:C:2010:083:SOM:EN:HTML.

Table 1.1. The social characteristics of Leeds and Warsaw, the UK and Poland at the start of the 21st century

Variable	Leeds 2001	UK 2001	Warsaw 2002	Poland 2002
% Aged 65+	15.5	15.9	16.5	12.7
% Manual occupations	47.5	47.1	32.0	52.5
% Not Host group	10.8	11.5	0.6	0.1
% Disabled	18.0	18.5	8.0[4]	11.6[4]/14.3[5]
Population	715,402	58,789,194	1,689,201	38,230,080
Area (km^2)	551.7	244,820	516.9	322,575
Population Density	1296.7	240.1	3267.9	118.5

Sources: 2001 Census of England & Wales, Scotland and Northern Ireland; 2002 Census of Poland.
Notes: 1. Not Host group in the UK = not White British Ethnicity, 2 Not Host group in Poland = not Polish citizens, 3. Disabled (Leeds only) = % Reporting Limiting Long-Term Illness, Health Problems and Disability, 4. Legally disabled only, 5. Disabled legally and biologically.

By contrast, Warsaw reveals its position as the national capital, having greater proportion of non-manual workers and a higher share than does Poland as a whole. Warsaw is approximately twice the size of Leeds in population[5], but occupies about the same land area, meaning that population densities are more than twice as high as in Leeds. The share of apartments in Warsaw's housing stock is about five times the share of flats in housing stock of Leeds (16.6 percent of the housing stock in Leeds in 2001 consisted of flats and in Warsaw in 2002 the percent was 88 percent). In general Polish towns and cities have much larger proportion of flats and apartments than English ones.

In comparison to Leeds, Warsaw is a much more ethnically homogenous city. However, when we change the perspective and compare Warsaw to other Polish cities, we can say that it is relatively multicultural and the most ethnically heterogeneous Polish city. It is estimated that every fourth foreigner living in Poland resides in Warsaw (Piekut 2010). In 2004 foreigners constituted 1.3% of Warsaw population, but only 0.2% of total population in Poland (Office for Foreigners 2004). The most numerous foreign immigrant groups living in Warsaw come from Asia, mostly from the Vietnam, but also from Armenia, Turkey and in recent years from China. Also important are citizens from the former Soviet Union (SU) countries, such as the Ukraine, Russia and Belarus. There are communities which originate in the European Union (France, Germany and Great Britain) and in the United States of America.

When we study cities and use local government administrative definitions of those cities, we need to be aware of how the administrative city fits within the functional city. There are two definitions of functional city – one defines the city as a continuously built up area (housing, business premises), while the other defines the city as a set of places with a high concentration of work places or service

[5] Bijak, Kicinger and Kupiszewski (2007) estimated that the population of Warsaw is undercounted by around a quarter of million persons.

facilities, linked by strong daily flows (journeys to work, to school/college/university, to shops or to other services).

Leeds Metropolitan District (the administrative definition we use) is an over-bounded area under the first city definition – its outer wards contain extensive tracts of farmland. However, it is an under-bounded area under the second definition. Workers, shoppers and night clubbers move into and out of Leeds MD each day from towns and villages in its wider urban hinterland (the Leeds City Region). Warsaw Municipality, on the other hand, is an under-bounded city under the first definition. The continuously built up area extends beyond the municipal boundaries into a ring of suburban settlements. The Warsaw urban region brings in daily commuters from a wider ring of satellite towns and villages.

These spatial considerations do impact on the range of diversity across the communities within each city that we would expect to observe. In Leeds we should observe the full range of residential communities from a dense inner city to the rural hinterland though we probably miss out some of the richest settlements in the outer ring of the urban region. In Warsaw we will observe less than the full range of urban communities missing out the varying characters of the suburbs and stand alone community settlements beyond the municipal boundary of Warsaw. When we compare our two cities we should bear in mind the geographical settings of our study areas.

1.3. What we mean by community

Before we proceed to review the literature on the social diversity of communities within cities, we should make clear how we define the term *spatial community*. Community is a form of social organisation based on commonality (Flint 2009). By spatial community we mean a group of people living in the same area who share the common space and are linked to each other by relations outside of the immediate household. These relationships may be social, such as talking to neighbours, shopping at the same places, belonging to associations that provide facilities for the common area, electing representatives to local councils and so on. The sense of spatial community identity is promoted by use of a common name for an area, which may have deep historical origins and which has persisted. In some cities names have been replaced by postal codes (e.g. the postal districts of Leeds such as LS16), but this is rare because postal codes are devices used by postal organisations, differ substantially between countries and are subject to operational change. Where we refer to community in the rest of the paper, we mean, unless otherwise specified, spatial community.

How big is a community within a city? These can range in size from people living in a few streets or blocks to quite large areas if population densities are low. In Leeds we chose to use 106 community

areas used in Unsworth and Stillwell (2004) to provide a common statistical description of the city over two censuses in 1991 and 2001, which were easily recognized by Leeds residents in a city council survey (Unsworth and Stillwell 2004: 19). These community areas contain about seven thousand people or about three thousand households. Each of these communities usually has an association and societies that bring together its residents. We could have chosen smaller areas (e.g. based on census enumeration or output areas) but these were felt to be too small to encompass the full range of social interactions; we could have chosen larger areas (e.g. electoral wards or districts) but these were felt not to describe the space of daily encounter which we felt should characterise the urban community. When we refer to neighbourhoods in the remainder of the paper we mean residential areas smaller than community areas, into which they can be grouped.

The ERC Living with Difference project, which has funded this study, has as its principal aim the discovery of people's attitudes in two European societies to people who are different from themselves. We will be testing hypotheses about the determinants of these views using the individual and household characteristics collected in a social survey. But the collective views of the communities in which people live can also influence views. In studies on election voting, the neighbourhood has been shown to have a strong effect (Johnston et al. 2000).

1.4. Research questions

We pose a set of research questions in this paper in order to understand social diversity in its spatial dimension. We need to define diversity dimensions and offer a broad understanding of diversity. Thus, the first question is *How is social diversity understood?* We review the literature on mapping social diversity and the literature on qualitative studies of diversity. We demonstrate that different communities have different levels of diversity and consequently people in these communities have different opportunities to encounter diversity. We propose to shift the diversity discussion away from one-dimensional, ethnicity dominated perspective. In doing so, we review previous studies on different aspects of diversity and highlight the diversity of diversity.

In Section 3 of the paper we discuss the comparability of data in British and Polish population censuses. In order to obtain as comparable an analysis as possible we wanted to select variables that were defined and measured in similar ways in both countries. Thus, the second question is *What are the common variables that should be used in the diversity analysis?* We compare the variables in the two countries that reflect diversity dimensions. We select study zones for spatial analysis of diversity, asking *Which study zones should be used to make of the diversity analysis comparable between the two cities?*

The paper includes some consideration of the methods used to measure diversity. We discuss the different indexes that have been employed to measure diversity level. Thus, the next question is *How can spatial diversity be measured?* The fourth part of the paper discusses analysis techniques. We wanted to pick the analysis procedure that would suit our multidimensional approach best. Then, we ask *How can we measure diversity to capture its multidimensionality?*

Much of the paper is given over to empirical analysis, where we map diversity in both cities. The mapping exercise is undertaken in three steps. First, we present one-dimensional diversity maps, i.e. the spatial distributions of people of different ages, occupations, ethnicities and disability in both cities. The research question is *What is the residential distribution of various social groups that that have different socio-demographic characteristics in Warsaw and Leeds?* Secondly, we wanted to produce typology of communities that would vary in terms of diversity level. We wanted to obtain community types that would be similar or different in the diversity dimensions that we have specified. The basic question is *Which communities are most/less diversified?* On the basis of the typology generated we want to determine which communities have low, medium or high levels of diversity. In doing so, we review the range of diversity indexes and select one for use.

Finally, the mapping exercise has one more specific objective. The statistical mapping analysis was the first phase of a larger study of diversity. The second phase will be a survey of respondent views on diversity. We need to use the knowledge of the social geography of diversity to ensure that the surveys in Leeds and Warsaw are as representative as possible. By producing diversity types of communities we provide strata for controlling the survey samples in both cities. Thus, another question is *What are the recommendations for interview survey?* The recommendations are based on the outcomes of the statistical and spatial analysis. The survey and other qualitative studies that are planned to be undertaken in the ERC "Living with Difference" project will continue to explore diversity, developing an in-depth understanding of diversity today, set in the context provided in this paper.

1.5. Aims and objectives

This paper has five broad aims, which follow on from the research questions. The first aim is to understand how social diversity has been previously studied in its spatial dimension. The second aim is to develop comparable indicators for studying diversity in two cities in different countries. The third aim is to select methods for describing that diversity. The fourth aim is to provide a systematic account of social geography of diversity in a comparison of two European cities. The fifth aim is to provide a framework for organising a representative sample of the inhabitants of the two cities for use in an interview survey. These aims and their associated objectives are set out in Table 1.2. By objectives we mean the tasks that need to be successfully executed to achieve the aims.

Table 1.2. Aims and objectives of the study

Aim	Description	*Objectives (sections)*
1	To understand how socio-spatial diversity has been studied in previous research	*To prepare a literature review on the diversity of social groups in cities (Sections 2.1, 2.2)* *To identify the main diversity dimensions in cities (Section 2.3)*
2	To compare social group variables used in UK and Polish censuses	*To select comparable study zones within each city (Section 3.1)* *To describe census variable definitions in both censuses and select comparable variables (Section 3.2, Appendix A1)*
3	To analyse diversity in its various dimensions	*To prepare a survey of spatial diversity measures (Section 4, Appendix A2)* *To determine the best way to measure spatial diversity (Section 4)* *To plot maps of social groups that have different socio-demographic characteristics (Sections 5 and 6, Appendix A4)*
4	To produce a typology of spatial communities in Leeds and Warsaw by degree and nature of diversity	*To generate through cluster analysis a set of typologies from which a best typology for each city is selected (Sections 6.1, 6.2, 6.3, 6.4)* *To describe and interpret the cluster types through statistical profiling and mapping (Sections 6.1, 6.2, 6.3, 6.4)* *To combine data from the two cities to show where communities in each of them are positioned in social space (Section 6.5)*
5	To use the results of our analysis to control and check the planned social survey	*To employ cluster types as sampling strata and their populations as sample weights (Section 7)*

1.6. Outline of the paper

The paper is organized as follows. In this first section we have set out the meanings of diversity and community and outlined our research questions, aims and objectives. In Section 2 we review relevant literature on diversity and the associated socio-geographic processes. In Section 3 we describe the census data in the UK and Poland which we will use to characterize the diversity of communities in Leeds and Warsaw. In Section 4 we discuss methods for indexing diversity using the census data and for creating the typology of communities with similar diversity profiles. In Section 5 we map indexes for each diversity dimension. In the Section 6 we report on the results of a multidimensional cluster analysis identifying the different diversity types of communities in the two cities. In Section 7 we show how the community diversity types can be used to stratify the planned social survey. The paper concludes with an executive summary in Section 8 and discussion of the main findings in Section 9.

2. REVIEW OF RELEVANT LITERATURE

Opportunities to encounter diversity will vary from community to community. The basic question we ask is: what might happen when different groups, with different values and needs, live together in the same community? In this section we discuss previous studies on mapping social diversity and residential patterns of various social groups. Before doing that we briefly discuss different approaches used to study and understand the social diversity[6]. Then we review literature on different residential patterns of various social groups. Finally, on the basis of the reviews we divide diversity into several dimensions that according to us are crucial.

2.1. Diversity – a new or an old concept?

The term diversity is differently used by scholars. 'Diversity' discussions in the past inevitably led to emphasising the social class and ethnic dimension of the phenomenon, with the latter perspective emerging in recent decades. Undeniably, the 20th century witnessed an increase of international mobility and the diversification of different forms of mobility in relation to the aim of migration, period of stay or frequency of movements (Castles, Miller 1998). Political transformation in the Central Eastern Europe has brought emergence a new geography of mobility in Europe too (Koser, Lutz 1998).

International migration, the recognition of rights of different minority groups (e.g. women, ethnic minorities, sexual minorities, and disabled people) and the development of internationally agreed human rights have led to an increase in the perception that contemporary societies are socially and culturally diverse. On the one hand, thanks to more intensified international migration, people with lots of different backgrounds more often than in the past live in one residential space, creating opportunities for inter-group encounters, socializing or conflict. On the other hand, *people are much more aware that other people are different*, not only in terms of social class or ethnicity.

Diversity studies have been criticised by some scholars for paying too much attention to ethnicity and country of origin that provide a misleading, one dimensional appreciation of contemporary diversity (Vertovec 2006; 2009). For example, Vertovec recognised that any one ethnic group is itself diverse in terms of gender and age profile, migration status, labour market experiences or spatial distribution. The interplay of these factors he termed 'super-diversity' (Vertovec 2006). For example, Indians in the UK may originate from the Punjab, Gujarat or South India; they may have migrated via East Africa. Indians may also be distinguished in terms of religious belief, as Hindus, Sikhs, Muslims, or

[6] A comprehensive review of the literature on social diversity will be prepared in another working paper by researchers of "Living with Difference" project.

Christians. The Vietnamese in Poland may have come from Hanoi or the southern regions of Vietnam. They may be Buddhists or Christians.

Putnam (2007) conceives of diversity as a social identity that can be de-constructed and re-constructed. However, he explores the ethnic dimension too. According to him diversity is not positive in a short run and works against social cohesion, fosters social isolation and reduces social solidarity. In the long run however, societies can produce better, cross-cutting forms of social solidarity and more encompassing identities (Putnam 2007). He concludes that by creating a new, broader sense of 'we' and diminishing social distance, a society can deal with diversity.

Moving away from the ethnic perspective, diversity is understood by some scholars in a broader way. The term 'difference' is used to encompass various minorities, not only ethnic, but also gender, sexual and dis(ability) minorities (Jazeel 2009). Another perspective is that societies have always been diverse, because *people have complex social identities and they belong to different social groups*. The social identities and social categories or groups that people belong to are: gender, age, sexuality, ability, race, ethnicity and religion. The levels of diversity thus can vary at the lower levels, like community or neighbourhood level (Talen 2010).

In the paper we would like to shift the discussion on diversity from exclusive, minority thinking to inclusive thinking about the whole society. In the next section we will discuss the literature on mapping diversity in the broader meaning.

2.2. Mapping social diversity

People in different groups live in different areas which reflect their choices and their constraints. The differences between people lead to them making different choices about their residential location, subject to the constraints imposed by the resources available and the operation of markets and legal systems. *Different residential patterns generate different opportunities to experience social difference through contact with relation to the dimensions of age, ethnicity, religion/belief, disability and socio-economic status*.

Here we discuss some of the debates about the social geography of cities and factors influencing where different people live in cities. We start with the studies undertaken in the late 19th and early 20th centuries (mostly the Chicago School and its legacy) when the social class perspective prevailed. Then we review literature on ethnic diversity that has developed with increase of international migration in the 20th century. Finally, we discuss the literature on multidimensional understanding of social diversity to include broader thinking on diversity. While the two first sections draw mainly from the quantitative studies, the third one incorporates more qualitative perspectives.

2.2.1. Social class perspective – the beginnings in studying residential patterns

The analysis of the geography of social diversity in cities has a long pedigree. Archaeologists, social historians, and novelists have provided descriptions of urban social life across the ages, ranging from accounts of ancient Rome (e.g. McCullough 1991), Tudor London (e.g. Sansom 2003), Victorian Manchester (Engels 1845, 1887)[7] and the turn of 19th/20th century working class Warsaw (Żarnowska 1985). At the end of the 19th Century the Rowntree family of Quaker businessmen and philanthropists set the standard for the detailed measurement of poverty in British studies, which continues to the present in the work of the Joseph Rowntree Foundation (e.g. Parekh et al. 2010)[8]. Charles Booth, a successful businessman and philanthropist, produced a series of studies on Life and Labour in London (Booth 1889), in which he mapped the distribution of seven social classes from "Lowest class, vicious, semi-criminal" to "Upper-middle and upper classes, wealthy". Orford et al. (2002) replicated Booth's analysis for inner London using 1991 census data and showed incredible persistence of the patterns of poverty and riches. Buck et al. (2002) use Booth's themes in their contemporary analysis of a much bigger London metropolitan region.

One important proposition that came out of work on the Chicago School of urban sociology was that social communities were organized in concentric rings around a central city core where work was concentrated (Burgess 1925). There was a gradient from *poverty* (inner ring) to *riches* (outer ring). This spatial model of the pattern of social class in the city was challenged by many authors. Hoyt (1939) carried out a major study of housing rents and owned values in Chicago neighbourhoods, showing that their variation was aligned to radial sectors in the city reflecting in part *access to valued external environments* (proximity to lakes and parks, distance from industrial sites which were highly polluting). Rees (1970, 1979) demonstrated that both spatial structures applied simultaneously to the 1960 Chicago metropolitan area and that there was no universal spatial structure. The spatial structure of social groups in any city depended on a *complex history of market and government decisions* reacting to the location rent gradient and amenity values of the urban environment in cities.

[7] Engels' account of poverty and disease in 1840s Manchester was a powerful influence on the birth of socialism. In his preface to the 1892 edition Engels recognizes the contribution made by sanitation engineering and public health to reducing the urban burden of disease observed in his 1845 study.

[8] A first in a series of three social surveys of the poor in York by Benjamin Seebohm Rowntree (Seebohm Rowntree 1901) measured the degree of poverty in poor communities in the city of York and argued for absolute thresholds of income needed to maintain nutrition intake. Since then the definition of poverty has become more sophisticated (e.g. British people classified officially as poor if their household income, corrected for the number and age of members, is less than 60% of the median household income for the country), reflecting the efficacy of welfare state benefits in eliminating Rowntree's absolute poverty.

Researchers in west coast America were sceptical of the application of this Chicago model outside the cities in which it had been tested (industrial metropolises in the eastern and mid-western states[9]). Based on studies of San Francisco and Los Angeles where the *spatial mosaic of social class* was much more complex, Eshref Shevky and colleagues (Shevky, Lewin 1949; Shevky, Williams 1949; Shevky, Bell 1955) suggested that what was important was to measure the social character of neighbourhoods using selected and *theoretically derived indexes* and to plot these indexes on aspatial graphs. These indexes were called Social Rank, Urbanization and Segregation. Critics of the social area theory and indexes then spent effort to establish the dimensions of socio-spatial variation in cities through empirical analysis using *principal components analysis* (aka factor analysis; see Robson 1969; Timms 1970; Rees 1979 for reviews of the field of 'factorial ecology')[10].

The studies carried in the last century demonstrated that people belonging to different social classes live in different areas in cities. Their decisions are shaped by housing policy and housing market developments which both influence the price of the land and housing, so the spatial patterns depend on the *wealth* and *income*. Another factor that proved to shape people's residential preferences is an *occupation* – some people live in proximity to their work places. The nature of today's society is different (more fragmented and fluid), thus social class is understood differently; class inequalities are based not only on differences in wealth and income, but also reflect the access to other resources, e.g. cultural capital (Duncan 2009). Thus, the *education level* is more often used to explain spatial patterns of urban populations (Freeman 2009).

2.2.2. Ethnic diversity perspective

There is a huge literature on the spatial distribution of minority ethnic groups and the ways in which those spatial distributions change over time. The traditional model was formulated in an important series of studies of foreign-born, international immigrants and of native-born internal in-migrants in Chicago in the period 1910-1970 (Park et al. 1925; Wirth 1938; Duncan, Duncan 1957). This model saw each new wave of immigrants establishing settlements in 'port of entry' neighbourhoods with cheap housing and access to industrial and service workplaces. The immigrant communities grew through high natural increase and further immigration fuelled by family and friendship connections in their communities of origin. Because of the demographic potential of such migrant communities, it was necessary for the ethnic community to expand physically and for the group members to become more dispersed. After several generations the first groups are replaced by new waves of migrants with

[9] They saw a more variegated and fluid organization of space in west coast cities, consequent on the adoption of the automobile as the principal means of intra-urban transportation.

[10] The most important factor in those studies was related to socio-economic status; the second most important dimension sorted neighbourhoods by the types of families/households that lived there; the third most important dimension distinguished neighbourhoods with high concentrations of minority groups from those dominated by the majority group

different origins[11]. If migrant group members were successful in the economic system of the host city, the migrant households would seek to improve their housing by moving outwards from the original communities. This spatial spread may involve continued segregation from the host population or spatial mixing as a result of integration or through assimilation through inter-marriage and host language acquisition[12].

However, this model needs to be modified by considering the socioeconomic, cultural and geographic background of immigrants. The Chicago model assumes that migrants come as a poor population. But many immigrant groups come with substantial human capital and business skills which enable them either to locate initially in wealthier neighbourhoods or to move to those areas once established[13]. Another major departure from the Chicago model is the circulation of international elite migrants, sought after by employers because of their scientific, technical or business skills. These immigrants have high incomes and can choose their preferred residential area. These foreign students will cluster in different residential areas, namely these close to their places of study or in the neighbourhoods with cheap, for rent housing[14].

This diversity of migration histories and backgrounds means diversity in spatial outcomes[15]. Studies on ethnic minorities in urban space very often include other explanatory variables and can be divided into three subgroups: those focused on socioeconomic differences between ethnic and racial groups (the 'social class' model), on non-socioeconomic differences, e.g. race, ethnicity, discrimination and their impact on spatial behaviour (the 'ethnic' model) (Darroch, Marston 1971), but also research employing institutional perspective, e.g. city development and transformations, housing market, which could be called the 'structural' model.

Several researchers have examined whether the socio-economic profile of a minority ethnic group explains its degree of segregation. Peach (2009) reviews this work. Taueber and Taueber (1964)

[11] This process was seen as analogous to the 'invasion and succession' of plant species colonizing new territories (e.g. laid bare by the retreat of the ice sheets at the end of periods of glaciations). It was recognized that human communities are more complex and many alternative futures for a migrant group were possible.
[12] This model provides a partial account of the immigration of successive groups in Leeds (the Irish in the mid-19th century, the Jews from Eastern Europe in the 1890s and 1900s and West Indians, Indians, Pakistanis and Bangladeshis in the 1950s to 1970s) and in Warsaw (spatial assimilation of Ukrainian immigrants in Warsaw stemming from their cultural proximity and mixed marriages with Poles).
[13] Examples in Leeds of such a group were Indians of East African origin (forced out of East Africa in the 1970s.) and the Vietnamese or Chinese in Warsaw that set up their own businesses.
[14] The number of international immigrants to the UK coming for formal study reasons grew from 56 thousand in 1991 to 211 thousand in 2009, and in Poland number of foreign students grew from 4 thousand in 1991 to 17 thousand in 2009.
[15] Peach (2009) has reviewed a large number of spatial segregation studies. He concludes that in the UK the spatial segregation of ethnic minorities decreased between 1991 and 2001 and that although some cities had ethnic enclaves with large minority ethnic groups' populations it was misleading to label these areas 'ghettos' of exclusion. In Warsaw, due to a small number of residents with non-Polish ethnic background, the segregation processes are invisible.

demonstrated that only 12 percent of Black/White segregation in Chicago in 1960 could be accounted for by income differences. Rees (1979) extended this analysis to 13 American cities. He showed that the socio-economic model for the spatial distribution of race groups in 13 American cities was completely inadequate. Other forces were operating, including a preference for living in areas with families belonging to the same racial and ethnic group and from discrimination mechanisms in the housing market[16].

The 'ethnic' model was more popular among European researchers. Although socioeconomic variables play important role in defining ethnic minorities residential patterns in urban space, cultural factors play also a major causal role, e.g. country of origin (Blom 1999), type of social networks and existence of mutual support or availability of religious or national services, e.g. restaurants, churches (Daley 1998; Peach 1996a). Other authors studied the housing careers of minority ethnic groups (Abramsson et al. 2002; Bolt, Van Kempen 2002; Bowes et al. 2002; Magnusson, Ozuekren 2002; Murdie 2002; Musterd, Deurloo 2002), dwelling choices by price and type (Blom 1999; Deurloo, Musterd 2001; Friedrichs 1998; Van Kempen, Van Weesep 1998), control over institutional resources of local population and regulation over rent control (Phillips 1998; Giffinger 1998).

The self-segregation concept brings a new perspective on the spatial practices of minority ethnic groups. It implies that ethnic minority groups decide to segregate out of choice not as a result of social or economic constraints. Self-segregation gives them easier access to some ethnic resources or in order to support co-ethnics (Poulsen 2009). the self-segregation concept was challenged in a study of British Muslims in Bradford (Phillips 2006). It has shown that minority ethnic communities (Pakistanis and Bangladeshis) did not self-segregate on purpose and would like to live in more ethnically mixed areas.

Finally, it should be added that beside residential segregation of different ethnic groups, which attracted most attention, there are studies of segregation in different urban places, e.g. the school or workplace. School segregation may be higher than residential one (Burgess et al. 2005). The reason for this difference was demographic: immigrant minorities have much higher child populations than the White majority. Although there is considerable literature on ethnic segregation across occupations, little attention has been paid to the extent of ethnic segregation in workplaces.

[16] Realtors guided prospective owners and renters to segregated neighbourhoods and exploited the price differences at racial community boundaries, buying cheaply from the leaving ethnic group and selling dear to the entering ethnic group (Rees 1979).

2.2.3. Broader thinking about social diversity – multidimensional socio-spatial diversity

Apart from the social class and ethnic dimensions there are other characteristics through which social relations are constructed and social attitudes can be produced in the urban space. The other most common axes of social difference are age, gender, sexuality, dis(ability) and religion. We discuss each of them in the urban space briefly below and we also incorporate some insights from qualitative literature in order to better understand why people of different characteristics live in different communities.

Age segregation may be observed in different types of places: home and household, workplace, neighbourhood, public institutions and other activity spaces. Many of the patterns of age segregation are comparatively recent phenomena in contemporary society and stem from economic and political changes, i.e. shifting away from home based system of production when people belonging to different generations shared the same household space and were working at homes (Vanderbeck 2007). Today, most families in the United Kingdom and in Poland live in two-generation households – parents and their offspring. But also delay in the age of marriage, the prolonging of the empty nest stage in the life course and the ageing of the population creating more lone person widow or widower households have increased the number of one person households (Dorling and Rees 2003).

Residential patterns influence opportunities to encounter other people and develop social networks with people of different age groups. Some research show that spatial isolation of elderly is followed by their social isolation (Becker 2003); others highlight the role of the neighbourhood as a socialisation spaces for older residents (Fadda et al. 2010; Scharf et al. 2003). Age segregation, lack of meaningful contact and understanding of people in different age groups may lead to ageism, i.e. prejudice based on age (Maxey 2009).

Men and women use, experience and understand urban space differently too. *Gender* divisions of space are common in cities, since urban space is socially constructed and reflects social relations. Spatial segregation begins in the household space and patriarchal relations in family and the gender division of labour is replicated outside home space. Women's spaces are meant to be private and devoted to reproduction, while men's tend to be public and devoted to production (McDowell 1999).

Gender relations are often portrayed along with other dimensions of diversity and depend upon *race*, *sexual orientation*, *ethnicity*, *social class* and *socio-economic status* (Miranne, Young 2000; Peake 1993). Examples of this work includes studies of occupational gender segregation and its influence on the neighbourhood relations (Pratt, Hanson 1988), studies of 'pink collar' jobs (clerical and secretarial

jobs) by women in the working-class suburban areas (Mez, Bühler 1998), unequal access to urban goods and services of women of different socio-economic statuses, e.g. childcare facilities, housing, transportation (McDowell 1993; McDowell et al. 2005).

Sexuality is another type of social identity that plays decisive role in spatial relations of individuals. Decades ago, Levine (1979) showed that major American cities contained 'gay ghettos' (called also 'gay villages', and more recently 'queer spaces', e.g. Rushbrook 2002) – neighbourhoods where lesbian, gay, bisexual and transgender people (LGBT) tend to live and places they gather and spend spare time. More up-to-date studies have revised the concept and shown how geographies of sexual difference are produced in specific sites and through specific spatial practices (Chauncey 1995).

Urban space in Western societies is organised and appropriated by heterosexuals and does not reflect gay and lesbian lifestyles (Valentine 1993; Valentine, Skelton 2003). Thus, LGBT people's residential patterns may be different from heterosexual people's residential preferences. There are some studies linking gay men with gentrification process (Miller 2009). On the other hand, the 'urban gay enclave' concept can be criticized as it presents gay communities as uniform and denies differences among LGBT people in relation to their class, ethnicity or gender.

One of the main topics in the literature on *(dis)ability* is the body in public space and exclusion of some people by environmental barriers in urban space (Butler, Bowly 1997; Dear et al. 1997; Matthews, Vujakovic 1995). Other studies examined influence of physical barriers on attitudes towards presence of disabled people in urban public spaces (Dear et al. 1997; Takahashi 1997).

Disabled people are very often multiply disadvantaged. As Chouinard (1997) points out: "being *out of place* also finds expression in economic and cultural marginalization". Disabled people have more difficult access to everyday life space, but also they are excluded from education and labour markets, e.g. by lack of specific accommodation in the workplace (Wilton 2004). It is worth adding that there are *gender* differences among disabled too. Disabled woman face more barriers and are less visible in public spaces (Chouinard 1997). Disability rates are strongly linked with increasing *age* and they correlate with *poverty*, so it is possible to find spatial concentrations of the disabled in the residential space of Leeds or Warsaw.

Religion is another aspect of diversity that can serve as a criterion for spatial differentiation. Many studies on religious segregation focus on the two distinctive groups, namely Jewish (e.g. Valins 2003), and Muslim (Phillips 2006; Varady 2008), or both groups (Flint 2010). Access to a place of worship does play a role in selecting residential neighbourhood for believers. Some studies are devoted to local conflicts surrounding building of places of worship, e.g. mosques in France (Cesari

2005) or mosques in Greece (Traindafyllidou, Gropas 2009). Indeed, 'Muslim spaces', marked in the urban spaces by mosques, constitute spaces of social unrest in secular societies. Phillips (2006) notes that similar sentiments were expressed about the synagogues in Jewish enclaves in East End of London in the 19th century.

Especially interesting are studies that identify the intersections of diversity dimensions, for example Muslim concentration and its relation to *housing discrimination* and *socioeconomic status* (Varady 2008), *class* structure and religion background within one *ethnic* group, e.g. spatial distribution of three faith groups – Hindu, Sikh and Muslim – among Indian ethnic population (Munoz 2010). Some studies recognise wider array of factors, for example, that Muslims neighbourhood aspirations depend on multiple identities, i.e. their *class*, *gender* and *age* (Phillips 2006).

2.2.4. Social diversity revisited

The residential segregation of different groups can be a result of social relations developed at the community level and processes like assimilation, integration and social mixing. While previous studies on mapping diversity in the 20th century highlighted social class differences or ethnic backgrounds of various social groups, other *social studies acknowledge the multidimensionality of social diversity and take into account multiple social group affiliations of people and their multiple social identities that produce differences in the residential patterns in the city*. In the exercise of mapping the social diversity we want to capture these various characteristics that compose social diversity.

We have incorporated the broad understanding of diversity drawing on different quantitative and qualitative social studies. In it worth to note that diversity is differently verbalised by scholars: some speak about different characteristics, some about different group affiliations while other emphasise different social identities. Because the rest of the paper is based on the analysis of the census data for residential communities, *we will use the statistical categories to describe the diversity*. However, we know that the data have limitations and they do not fully reflect social reality which is much more complex.

2.3. Dimensions of difference

We now sum up the conclusions of this extensive though still partial review of the social geography of diversity in Western cities. People and therefore the communities they live differ in several important ways, which we call dimensions of diversity. A basic sociological proposition says that people can be ascribed to different social categories (i.e. people who do not interact with each other, but share the same characteristics) and social groups (i.e. people who interact with each other,

recognise themselves as members of the same group and share similar identity). This constitutes a good starting point in dividing diversity into certain dimensions.

2.3.1. Family/Demographic status

Each individual is a part of a primary group, i.e. people with whom he/she has close, personal or kinship-based relationships. The relations depend on our age, gender, sexual orientation and family type we are involved with. A widely employed perspective for analysis of family/demographic status is the life course. Family status is age dependent. We are all born, age and then die. During the course of life we travel through a variety of stages in which we live with others for shorter or longer periods in families, in households or alone. We are born into a family and are looked after by one or more usually two parents, for periods of 15 to 30 years to a greater or lesser extent. From age 16 onwards we tend to move from our family of origin into independent living, either in new shared households, or as new cohabiting or married couples. The age and speed of this transition depends on the structure of the education system (e.g. in England, middle class students move away from home to attend university, in Poland students will tend to live with their parents while studying at their local university). Things happen to change your life course status whether you like it or not (e.g. you age, your spouse/partner dies) or as result of choice (e.g. you marry or not, you participate in a partnership relation with people with the same or opposite gender, you have children or not), but the range of choice changes with age.

2.3.2. Socio-economic status

Individuals belong also to many secondary groups (more formalised and institutionalised, e.g. university, working place and some interest groups). The most important are reference secondary groups – people we compare with and refer to when we evaluate our achievements, aspirations and ambitions. Level of education or qualifications, occupation and income constitute basic characteristics that help people compare themselves with others. It is useful to broaden the life course into the socio-economic life course. Socio-economic status is age dependent as well. In the first years of our lives we are dependent on our parents who support us financially. Later on, we enter labour market and throughout our professional career we change our occupation and our incomes rise or fall. Our socio-economic status is reflected in our consumption patterns, e.g. the housing we consume. Socio-economic status differs from family/demographic status in allowing people to move into and out of different strata. So some persons can be upwardly mobile in their work career as they gain promotion in a successful business or public body, but other persons become downwardly mobile when they are made redundant and have to accept much lower waged work.

2.3.3. Ethnicity/Affinity

Both primary and secondary groups are situated in affinity groups, i.e. larger social groups that individuals feel attached to although they do not know each other in person, e.g. national, ethnic or religious groups. We can again broaden the life course into the ethnic life course. Ethnicity is generally rather a 'sticky'/unchangeable characteristic, i.e. people tend to stay in the same group for their whole life. However, there are some groups of people that are more prone to change their ethnic affiliation, e.g. international migrants or people with dual (or multi) ethnic background.

2.3.4. Disability

Disability is a different dimension of diversity. It is a broad category, as it includes mental and physical disability. Some chronic or serious illnesses, like cancer or HIV/AIDS are sometimes included as disabilities. As treatments for various cancers or HIV/AIDS improve, the illnesses change from ones that kill people to ones that can be lived with, but with a degree of disability (e.g. side effects of drug treatments). It is a relative characteristic: being disabled depends on society's definition of who is disabled. It is socially constructed then, especially in the case of mental ill-health, and also depends on people's perception of their limitations. We can combine this diversity dimension with a life course and consider the disability life course. Disability can be a constant characteristic or change over time. People are born with some disabilities (i.e. congenital disorders), but also there are types of disabilities that we acquire with age, e.g. vision, hearing or mobility impairments. Other types of disability are totally accidental, e.g. as a result of a car crash or another accident, from which we may fully or partially recover.

2.3.5. Dimensions of diversity

Bearing in mind these dimensions, diversity has also two important traits that should be underlined. These are

- *Interdependence*. Diversity dimensions are interdependent. Changes in one dimension impact other dimensions. For example, when you get older, you improve your education or change occupation. Similarly, the probability of disability is higher for older people or for people that are in a higher risk occupation, such as manual workers.
- *Complexity*. All specified diversity dimensions can be studied in a cross-sectional manner at the different spatial levels. We can consider how diverse families or households are, how diverse small neighbourhoods are, how diverse larger communities within cities are and how diverse the whole city population is. Diversity will vary depending on the scale at which it is studied, though there is linkage between spatial units at different scales through aggregation. Spatial units have different level of diversity in the specified dimensions.

The aim of the paper is to understand diversity and its dynamics and investigate what is the relation between diversity and urban space. We argue that decisive role in diversity construction in urban space is played by interplay between the diversity dimensions. Because diversity is socially and spatially constructed people encounter different diversity levels in different urban spaces, namely communities. Consequently, we argue that changes in characteristics of the four dimensions of diversity are of prime importance for people's encounter of diversity.

2.3.6. Profiling communities using the dimensions of diversity

Assuming we have measured the dimensions of diversity in residential communities properly, how do we combine the information from several dimensions to obtain profiles of community areas. The basic idea is to use the community area variables as inputs to a classification exercise in which community areas are grouped into a set of types. This grouping is implemented using a technique called *cluster analysis* (Everitt et al. 2001). Cluster analysis is employed in *geo-demographics* (Webber, Craig 1978; Harris et al. 2005) to identify types of residential neighbourhood populations. What we will do is use geo-demographic techniques to identify clusters of community areas in Leeds and of urban regions in Warsaw that have similar diversity profiles.

2.4. Conclusions for the diversity case studies in Leeds and Warsaw

The literature review above allows us to draw the conclusion that many complex factors determine how diverse residential communities are and where different social groups concentrate or are under-represented. These factors affect all our dimensions of diversity: age and family status, gender, socio-economic status, disability, ethnicity, religion or sexual orientation. We do not argue that the list is comprehensive; nevertheless these aspects represent dimensions of diversity that we distinguished in the beginning of the chapter.

The aim of the paper is to demonstrate social diversity and its spatial dimension in two European cities. Can the residential variation of diversity be explained by one decisive factor? Indeed, many studies concentrate on one-dimensional diversity and focus on groups that differ either in terms of gender, age, sexual orientation or ethnicity. But there are some intersectional studies and two or three dimensions are recognised. However, multidimensional diversity, i.e. diversity undertaken from the perspective of many different dimensions, is not so commonly analysed in the geographic studies.

In this paper we argue that the nature of diversity is the result of interplay between many interconnected factors. More specifically we argue that decisive role in diversity construction in urban space is played by relationship between four diversity dimensions: demographic, socio-economic, ethno-cultural and disability. All four dimensions have different dynamics in the life course:

demographic diversity depends on gender, sexual orientation and family type; socio-economic diversity – on education, occupation and income; ethnic diversity – nationality, ethnicity and religion and disability is a function of health, social definition and individual perspective.

In rest of the paper we will analyse social diversity in Leeds and Warsaw from the perspective of these dimensions. We use statistical data from UK and Polish censuses. Analysis of multidimensional diversity demands an original theoretical framework and analysis techniques. We believe that this can be achieved by employing cluster analysis which allows us to include many factors. We will present diversity maps of both cities with analysis of diversity variations at the community level.

3. DATA DESCRIPTION

In this section we provide descriptions of the censuses that are organised in the UK and Poland which result in statistical datasets that can be used in international comparisons. We select and present the geographical areas used in the analysis – Community Areas for Leeds and Urban Regions for Warsaw. Next, we go through variable selection and detailed comparison between the two cities, which enables us to carry out diversity analysis in both urban spaces.

3.1. Datasets

The first step before conducting statistical analysis of diversity in Leeds and Warsaw was to prepare a comparison of UK and Polish census data. In this sub-section we provide short descriptions of the censuses: their organisation, population, geography and availability of census data.

3.1.1. Census organisation

The first census in the United Kingdom took place in 1801. A census was taken every 10 years except in 1921 in Ireland (when there was a civil war) and in 1941 in the whole UK (when there was a World War). In the first census people were asked five questions, but the questionnaire was progressively expanded and the number of questions reached 50 in the 2001 census.

The administration of the 2001 Census in the UK was organised in six phases (Rees et al. 2002). In the first phase, three weeks before the Census Day (29th April 2001) self-completion forms were delivered by enumerators to households and managers in charge of communal establishments. In the second phase, people completed the forms on the Census Day 29 April 2010 and posted back the forms to the Census Agencies. In the third phase (9-18 May) enumerators visited non-respondent households and helped them complete the form, when contact was made. In the fourth phase, the data in the forms was captured digitally through automated scanning and edit processing to correct inconsistencies and impute missing items. In the fifth phase, a census coverage survey was undertaken to estimate how many households and individuals in households had been missed. The sixth phase was to model (i.e. impute) missing households and missing persons within households using sophisticated versions of the capture/re-capture technique (Diamond et al. 2002).

In Poland, the first census was organised in 1789 in Poland and then 1808 in The Duchy of Warsaw and then repeated irregularly every several years. After partition of Poland the Polish people were included in censuses in Russia, Austria or Prussia. The first Polish census took place after restoration of Poland's independence in 1921 and then in 1931. Next censuses took place after the World War II in 1950, and then every ten years (1960, 1970, 1980, 1988, 2002). There were also carried one

summarising census in 1946 and three representative studies, called micro-censuses in 1974, 1984 and 1995 (GUS 2003a).

The field work for the 2002 census in Poland lasted from 21 May to 8 June. The census data were collected by enumerators during face-to-face interviews with an adult representative of a household. The Census Day was 20 May 2002. About 2 percent of population (over 730,000) was not captured in the census and their basic personal data, namely names, addresses, years of birth and gender, was copied from registers. More sophisticated data of missing persons was not modelled, but the answers were marked in the statistical tables as 'no answer' (GUS 2003b). The fieldwork was followed by a census coverage survey covering 64,000 households and 192,000 people to evaluate quality of the data (GUS 2003b; GUS 2003c). Similarly as in the UK, the data in the forms was automatically scanned and edited in order to correct inconsistencies and impute missing items.

The next census takes place in both countries in 2011.

3.1.2. Census population

In the United Kingdom the census covered *all people usually resident* in the area. A *usual resident* is generally defined as someone who spends the majority of their time residing at that address. It includes:

- people who usually live at that address but who are temporarily away from home (on holiday, visiting friends or relatives, or temporarily in a hospital or similar establishment) on Census Day;
- people who work away from home for part of the time, or who are members of the Armed Forces;
- a baby born before 30th April 2001 even if it was still in hospital;
- people present on Census Day, even if temporarily, who have no other usual address.

The usual resident population did not include people present at an address on Census Day whose usual address was elsewhere (ONS 2004)[17].

In Poland the census covered 1) *permanent population* living in Poland, that is all people that were registered for permanent stay in the register of residents [18], both those who were present in dwellings

[17] The usual resident population did not include people away from their home address who had been living, or intended to live, in a special establishment such as a residential home, nursing home or hospital for six months or more (they were enumerated as usually resident at the special establishment).

[18] In the light of Polish law people living in Poland should register their current place of residence in a local administration unit and repeat the registration act each time they change their place of living for longer than three months (two months in 2002). It should be mentioned that it is a kind of 'dead law', since many people do not follow this requirement.

or other buildings and those who were absent during the census, and 2) *temporary population* (not registered for permanent stay in Poland) present in dwellings or other buildings in Poland on a Census Day.

The 2002 census produced three categories of population:

- *Permanent population* – people that were registered for permanent stay, present or absent on a Census Day;
- *Population de facto living in Poland* – includes 1) permanent population that was present on a Census Day or absent, but for less than two months and 2) temporary population, but living in a given place for longer than two months[19];
- *Resident population* (new census category) – includes 1) permanent population without people absent for longer than 12 months and 2) temporary population living in a given place on a Census Day for longer than 12 months (including foreigners). This definition is in agreement with the definition of resident population arising from the UN (1998) recommendations on definition of migration.

Most of the census statistics use as a base the population *de facto* living in Poland, e.g. 38,230 thousand people. Residents constitute 37,620 thousand (GUS 2003b).

3.1.3. Census geography

In the United Kingdom, the basic units for collecting census data are enumeration districts (EDs) that are drawn along obvious topographic features, e.g. roads, rivers, railway lines. Until 1991 census EDs were used for both data collection and output. In 2001 enumeration and output geographies were constructed separately. Output areas (OAs) were built from contiguous unit postcodes using a zone design algorithm (Martin 2002) that ensured they all had above confidentiality threshold numbers of people (100) and households (40) and relatively uniform population (125 households and 300 people), nesting inside electoral wards and local government districts. OAs were designed also to be socially homogenous for a few selected census variables and to avoid convoluted shapes. The zone design algorithm was developed by Martin (2002) based on earlier work by Openshaw (Alvanides et al. 2002). There are almost 2.5 thousand Output Areas in Leeds with an average of 293 residents and 124 households (Table 3.1).

[19] Population *de facto* living in Poland excludes people that arrived from abroad for a temporary stay.

Table 3.1. Census geography in Leeds

Level	Number	Size (km²) (average)	Population (average)	Households (average)
Output Areas (OAs)	2 439	0.22	293	124
Community Areas (CAs)*	106	5.2	6 749	2 845
Wards	33	16.7	21 679	9 140
Leeds (district level)	1	551.7	715 402	301 614

Source: ONS.
Notes: * - the special scale selected for analysis (see sub-section 3.1.5).
Output Areas (OAs) – these are the smallest units for statistical data from the 2001 Census are published.
Community Areas (CAs) – these are groupings of OAs in Leeds, recognized as communities by the residents.
Wards – these are the electoral areas used in district elections in Leeds.
Districts – this is the term used in the UK to refer to the lowest level of administrative unit.

In Poland the basic units for collecting census data are also enumeration districts. However there is no separate output geography. Nine enumeration areas (EAs) comprise one statistical region (SRs). Enumeration areas contain up to 500 people and 200 dwellings and statistical regions contain up to 2700 people and 999 dwellings. Statistical regions fit into administrative boundaries (e.g. of Warsaw municipality). Warsaw was covered by 7,930 thousand enumeration areas and 1,442 statistical regions in 2002. The average population living in an enumeration area is 213 (but no more than 500 people and 100 households) and 1,171 live in a statistical region (but no more than 2,700 and 500 households). Table 3.2 presents the statistics for EAs, SRs, Urban Regions (URs), Districts and Warsaw Municipality.

Table 3.2. Census geography in Warsaw

Level	Number	Size (km²) (average)	Population (average)	Households (average)
Enumeration Areas (EAs)	7 930	0.01	213	96
Statistical Regions (SRs)	1 442	0.36	1 171	525
Urban Regions (URs)*	92	5.6	18 361	8 235
Districts	18	28.7	93 845	42 088
Warsaw (county level)	1	516.9	1 689 201	757 578

Source: GUS.
Notes: * - the special scale selected for analysis (see sub-section 3.1.5).
Enumeration Areas (EAs) – these are the units for which statistical data in the 2002 Census were gathered.
Statistical Regions (SRs) – these are the smallest units for statistical data from the 2002 Census are published.
Urban Regions (URs) – these are groupings of SRs in Warsaw. They were defined by City Council in the 1967.
Districts – these are basic units that Warsaw is divided into. Districts have their own administrative bodies.
County – this is the term used in Poland to refer to the medium level of administrative unit[20].

[20] Territorial structure of administration in Poland consists of three levels. The highest level of administrative division is voivodship (which is called *województwo*); it corresponds with the NUTS 2 level. There are 16 voivodships in Poland. County, called *powiat*, is a lower-level unit than region; there are around 379 counties in Poland, it is equal to LAU 1 level. The lowest level of administrative division comprises unit called *gmina* – commune; there are almost 2,500 communes in Poland. Commune is equal to LAU 2 level.

3.1.4. Availability of census data

The availability of census data differs between countries. In the United Kingdom academic users are given free online access to 1971-2001 census data resources. The webpage Census.ac.uk offers access to aggregate data, look-up tables, interaction data, digital boundary data and microdata, and provides metadata for longitudinal studies.

Meanwhile in Poland only basic data from 2002 census are available. Moreover these data are available as look-up tables for administrative units, i.e. at the level of regions, counties, cities and some for communes, but not for basic census geography units. What this means is that data at the level of statistical regions or enumeration areas cannot be browsed online. Interaction data, data on the lower levels of aggregation, microdata and boundaries maps are available by purchase. Census data are available for the population *de facto* living in Poland which differs from the actual residents of the city (for definitions see earlier).

In the analysis we used data for Output Areas and Statistical Regions which were aggregated for the Community Areas in Leeds and Urban Regions in Warsaw. The final study regions are described in the next section.

3.1.5. Study Zones

The basic statistical units used in the UK and Polish census (output areas in Leeds and enumeration areas in Warsaw) are too small to constitute satisfactory geographical communities for diversity analysis. There was a need to find bigger regions. To select the final study zones we reviewed different levels of spatial organization in both cities. Leeds is divided into 33 wards. Warsaw is divided into 18 districts. The maps with this basic administration division of both cities are provided below drawn to the same scale (Figures 3.1 and 3.2).

Both Leeds wards and Warsaw districts are inappropriate areas for socio-spatial analysis: they significantly differ in terms of area size and population size and every ward/district houses populations that are too heterogeneous. Socio-spatial analysis that will reveal important differences regarding different dimensions of diversity should be developed on the basis of lower level of data aggregation.

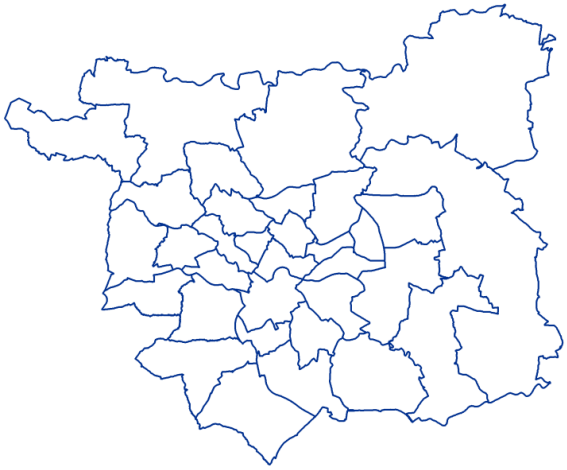


Figure 3.1. Wards in Leeds (33)

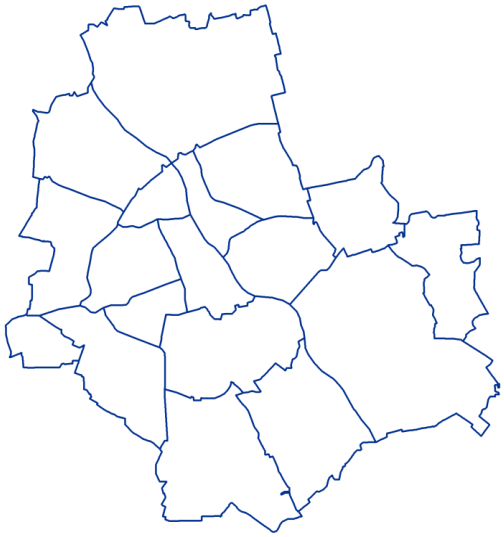
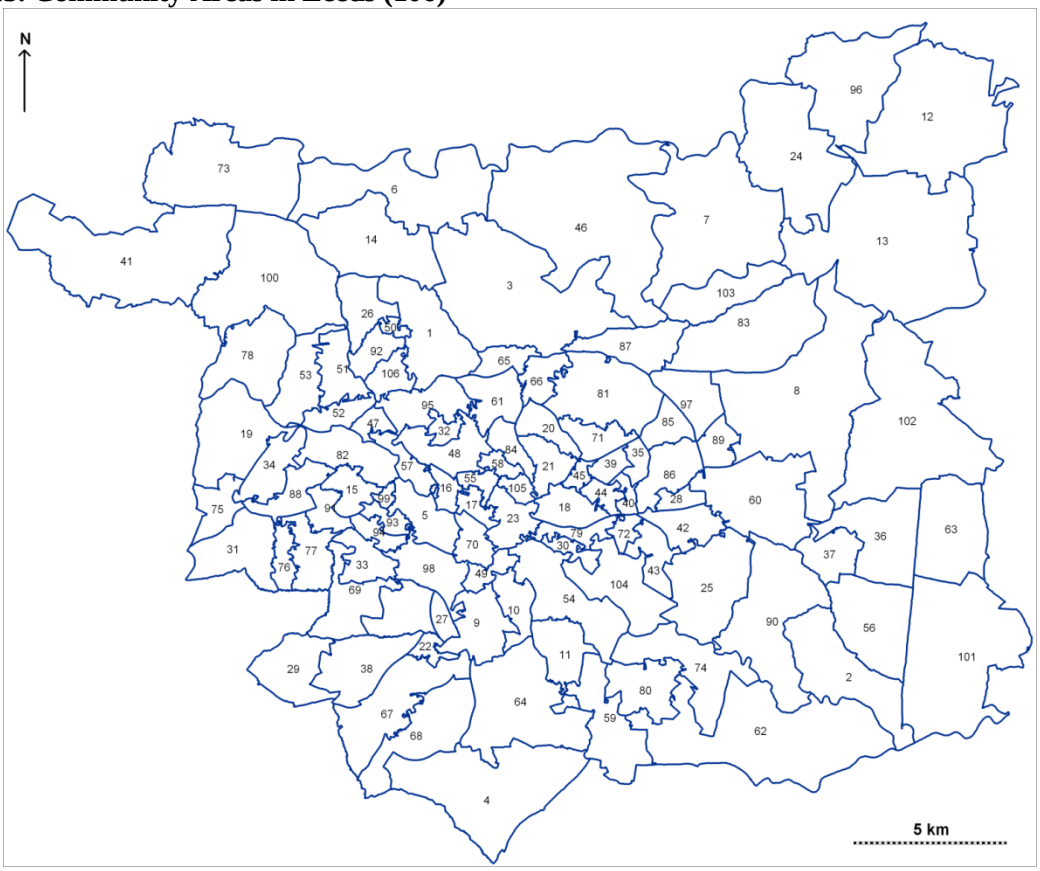


Figure 3.2. Districts in Warsaw (18)

The smaller areas should be known by name to city residents and recognised as neighbourhoods by people. In Leeds Stillwell and Shepherd (Unsworth, Stillwell 2004) prepared aggregations of census output areas which they called Community Areas (see Figure 3.3.). The areas could be matched easily with sets of enumeration districts from the 1991 Census so they could be useful for detailed spatial comparisons between 1991 and 2001. They were also designed to match neighbourhoods identified as meaningful in Leeds City survey of residents. Consultations among staff at the University of Leeds confirmed that they were named zones familiar to people who worked and lived in the city.

Figure 3.3. Community Areas in Leeds (106)



Source: Unsworth, Stillwell 2004: 20.

1	Adel	24	Collingham & Linton	52	Horsforth Newlaithes & Woodsid	80	Rothwell
2	Allerton Bywater & Gt and Lt Preston	25	Colton	53	Horsforth West End	81	Roundhay
3	Alwoodley & Wigton Moor	26	Cookridge	54	Hunslet / Stourton	82	Sandfords, Ganners & Moorside
4	Ardsley East / West	27	Cottingley	55	Hyde Park	83	Thorner
5	Armley	28	Crossgates	56	Kippax	84	Scott Hall & Miles Hill
6	Arthington & Pool	29	Drighlington	57	Kirkstall	85	Seacroft North
7	Bardsey & East Keswick	30	East Bank	58	Woodhouse	86	Seacroft South
8	Barwick & Scholes	31	Fairbank	59	Lofthouse & Robin Hood	87	Shadwell
9	Beeston	32	Far Headingley	60	Manston	88	Stanningley
10	Beeston Hill	33	Farnley	61	Meanwood	89	Swarcliffe
11	Belle Isle	34	Farsley & Rodley	62	Methley	90	Swillington
12	Boston Spa	35	Fearnville	63	Micklefield	91	Swinnow & Fairfields
13	Bramham	36	Garforth East	64	Middleton	92	Tinshill
14	Bramhope	37	Garforth West	65	Moor Allerton	93	Upper Armley
15	Bramley	38	Gildersome	66	Moortown	94	Upper Wortley
16	Burley	39	Gipton North	67	Morley North	95	West Park
17	Burley Lodge & Little Woodhous	40	Gipton South	68	Morley South	96	Wetherby
18	Burmantofts, Lincoln Green & Ebor Gardens	41	Guiseley	69	New Farnley	97	Whinmoor
19	Calverley	42	Halton / Whitkirk	70	New Wortley	98	Wortley
20	Chapel Allerton	43	Halton Moor	71	Oakwood	99	Wythers
21	Chapeltown	44	Harehills	72	Osmondthorpe	100	Yeadon
22	Churwell	45	Harehills Triangle	73	Otley	101	Ledston & Ledsham
23	City Centre	46	Harewood and District	74	Oulton & Woodlesford	102	Aberford
		47	Hawksworth	75	Priesthorpe	103	Scarcroft
		48	Headingley	76	Pudsey	104	Cross Green
		49	Holbeck	77	Pudsey Lowtown	105	Little London
		50	Holt Park	78	Rawdon	106	Ireland Wood
		51	Horsforth	79	Richmond Hill		

Warsaw districts are divided into lower level auxiliary/supporting areas, called residential districts/estates (over 200 in number), 92 Urban Regions (URs), 149 Metropolitan Informational System (*Miejski System Informacji*) or electoral districts. For example, Praga Południe district is divided into 8 residential districts, 6 metropolitan information system areas and 12 electoral districts. The differences in the boundaries of these spatial divisions are visible in Figures 3.4 and 3.5.

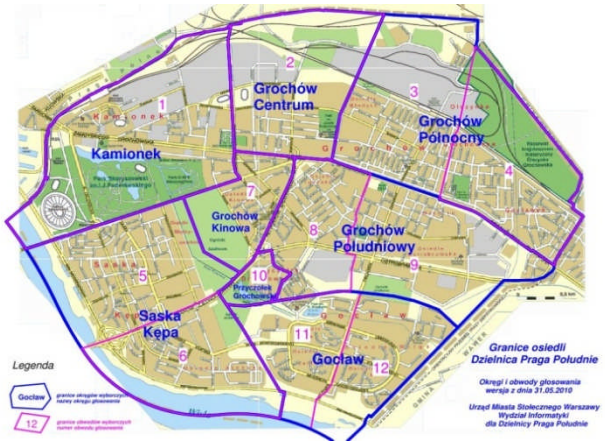


Figure 3.4. Residential districts (blue) and electoral districts (pink) in the Praga Południe district
Source: http://www.pragapld.waw.pl/assets/images/zdjecia_aktualnosci/2010/czerwiec2/mapa_osiedli.jpg.

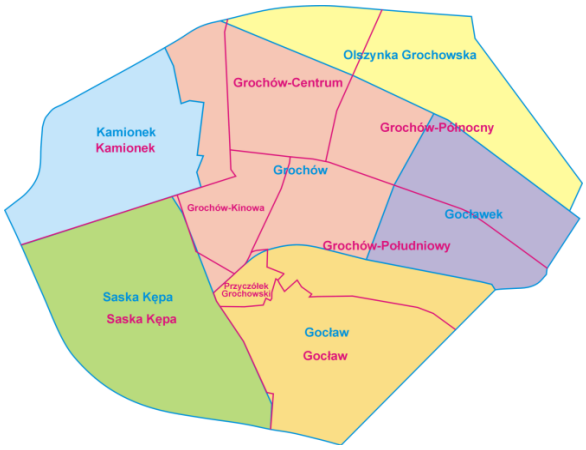
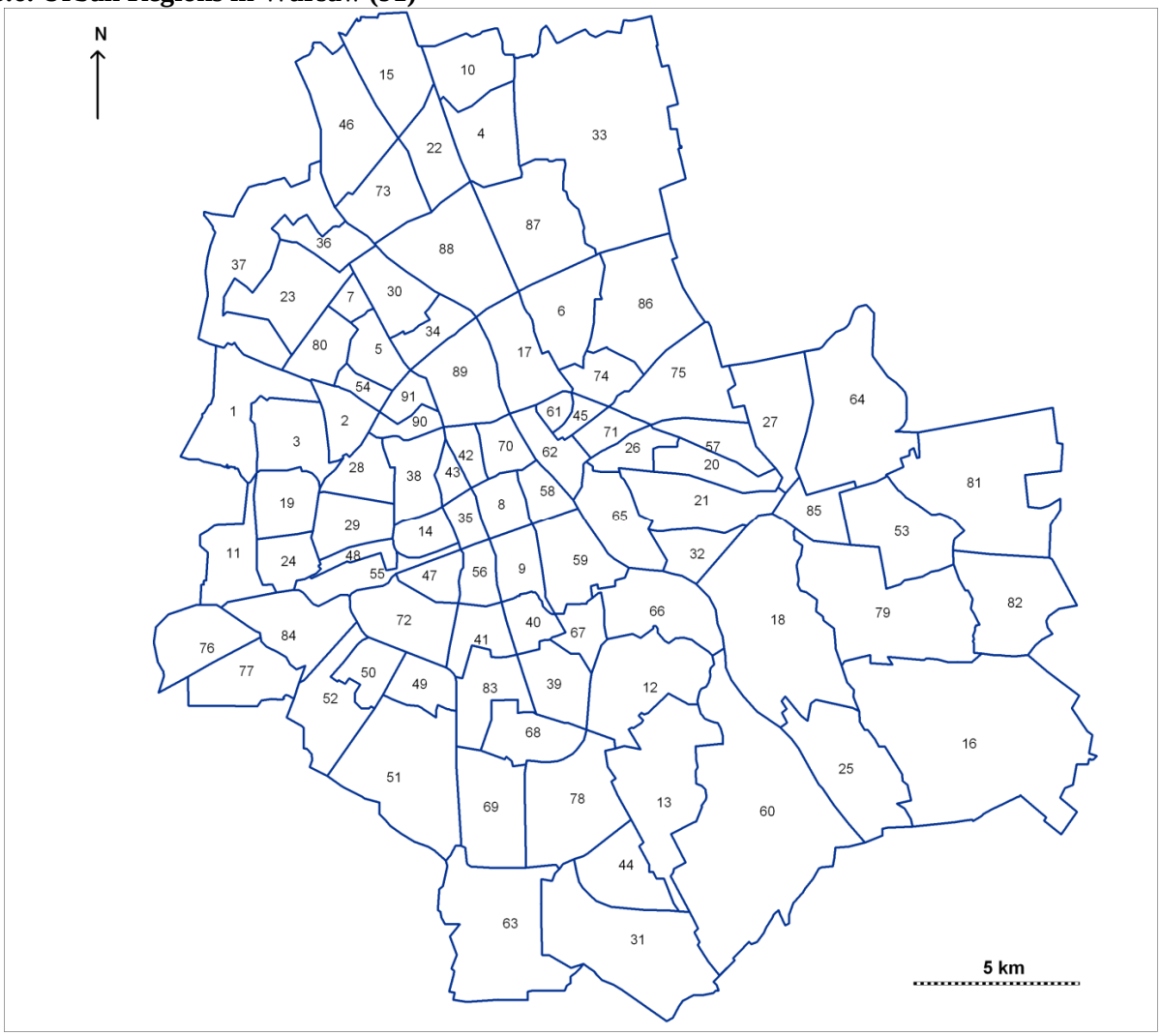


Figure 3.5. Residential districts (pink) and MSI areas (blue) in the Praga Południe district
Source: http://pl.wikipedia.org/wiki/Praga_Po%C5%82udnie.

Estates are geographical units recognised by people on the ground, but they are not used in the official statistics and their maps are unavailable. Urban Regions boundaries are the closest to the estates. Their names are more meaningful for the people of Warsaw than MSI units which were created for local communication and address boards. They are used by the city council, Polish geographers and research companies. Data for urban regions can be aggregated from statistical regions data, as the SRs nest into URs (except for two urban regions – Młociny-Las and Wólka Radiowo – which were merged, because they were covered by one statistical region). Figure 3.6 is a map of Warsaw's Urban Regions.

We believe CAs in Leeds and URs in Warsaw are reasonably comparable. The numbers of zones are similar – 106 and 91, and they have nearly the same areal extent, 5.2 and 5.6 square kilometres respectively. Urban Regions in Warsaw do have larger average populations than Community Areas in Leeds and so are 2.5 times as dense, reflecting the dominance of medium size apartment buildings.

Figure 3.6. Urban Regions in Warsaw (91)



Source: City Council; map prepared by Polish Academy of Sciences, Institute of Geography and Spatial Organization.

1	Babice	26	Kamionek	50	Okęcie Przemysłowe	75	Targówek Przemysłowy
2	Bemowo-Wschód	27	Kawęczyn	51	Okęcie-Lotnisko	76	Ursus-Gołąbki
3	Bemowo-Zachód	28	Koło-Północ	52	Okęcie-Opacz	77	Ursus-Skorosze
4	Białołęka Dworska	29	Koło-Południe	53	Park Sobieskiego	78	Ursynów
5	Bielany	30	Las Bielański	54	Piaski	79	Wawer
6	Bródno	31	Las Kabacki	55	PKP-Wola	80	Wawrzyszew
7	Brzeziny	32	Lotnisko	56	Pole Mokotowskie	81	Wesoła Północ
8	Centrum-Północ	33	Mańki-Brzeziny	57	Postojowa	82	Wesoła Południe
9	Centrum-Południe	34	Marymont	58	Powiśle-Północ	83	Wierzbno
10	Choszczówka	35	Mirów (Środm. Zach.)	59	Powiśle-Południe	84	Włochy
11	Chrzanów	36	Młociny	60	Powsin	85	Wygoda
12	Czerniaków	37	Młociny-Las & Wólka Radiowo [merged]	61	Praga II i III	86	Zacisze
13	Czerniaków Wilanów			62	Praga-Centrum	87	Żerań Wschodni
14	Czyste	38	Młynów	63	Pyry	88	Żerań Zachodni
15	Dąbrówka	39	Mokotów Centrum	64	Rembertów	89	Żoliborz Centralny
16	Falenica	40	Mokotów Stary	65	Saska Kepa	90	Żoliborz Przemysłowy
17	FSO	41	Mokotów Wschód	66	Siekierki	91	Żoliborz Zachodni
18	Gocław	42	Muranów Wschodni	67	Sielce		
19	Górce	43	Muranów Zachodni	68	Służew		
20	Grochów-Północ	44	Natolin	69	Służewiec Południowy		
21	Grochów-Południe	45	Nowa Praga	70	Stare i Nowe Miasto		
22	Henryków	46	Nowodwory	71	Szmulowizna		
23	Huta Warszawa	47	Ochota-Centrum	72	Szosa Krakowska		
24	Jelonki	48	Odolany	73	Tarchomin		
25	Julianów	49	Okęcie Nowe	74	Targówek Mieszkaniowy		

3.2. Variables

In this section the main census variables of UK and Polish censuses that will be used to measure diversity are presented and compared. By making the variables as comparable as possible, we can develop a joint analysis of diversity as well as separate analyses for each city.

3.2.1. Census data for demographic diversity

Demographic variables are available in censuses for individuals by age, gender, marital status and household representative status; for households by type (family/not-family), composition and generation (by age); and for families by type (by marital status), composition and generation (by age). Data from both censuses are comparable in terms of age, generation, household size, marital status, family composition, and partially comparable in case of disability (Table 3.3). More detailed tables with variable comparisons and definitions are provided in Appendix 1, Table A.1.1.

Table 3.3. Comparison of demographic variables in the British and Polish censuses

Variable	Availability		Variable attributes
	UK	Poland	
Age	✓	✓	Numerical (integer)
Generation	✓	✓	Young people (≤29) Middle-aged people (30-59) Older people (≥60)
Household composition	✓	✓	One person household Family household Non-family household
Household composition by generation	x	✓	No variable in the UK census
Household size	✓	✓	Numerical (integer)
Marital status	✓	✓	Single Married Separated Divorced Widowed
Family definition	?	?	Slightly different definitions
Family status	?	?	Different classifications
Family type	✓	✓	Married couples with dependent children Cohabitating couples with dependent children Lone parent with dependent children Families without dependent children
Generation in family	✓	x	No variable in Polish census
Living arrangements	✓	x	No variable in Polish census

Notes: ✓ - variable present and comparable, x – variable absent, ? - variable present, but not comparable
UK = United Kingdom, PL = Poland.

3.2.1. Census data for socioeconomic diversity

Socioeconomic variables available in censuses include level of education/qualifications, economic activity, occupation and industry of main job, income/source of livelihood, tenure, accommodation quality and building age (Table 3.4 and Table A1.2 in Appendix 1). Many data on socio-economic status were collected in a similar way in both censuses, e.g. data on economic activity, including employment, unemployment, main job, occupation and industry of employment. Household space and housing tenure types could be compared (after merging some variables) as well.

Table 3.4. Comparison of socioeconomic variables in the British and Polish censuses

Variable	Availability		Variable attributes
	UK	Poland	
Economically active definition	✓	✓	The same definitions
Employed definition	✓	✓	The same definitions
Unemployed definition	✓	✓	The same definitions
Main job definition	✓	✓	The same definitions
Occupation classification	✓	✓	Classification (ISCO-88)
Employment type	✓	✓	Employee Employers / Self-employed with workers Self-employed without workers Others
Income	x	x	No question in both censuses
Source of livelihood	x	✓	No question in UK census
Socio-economic position	?	?	UK – based on individual's occupation Poland – based on household's source of livelihood
Industry of employment	✓	✓	classification
Level of qualification / education	✓	✓	UK – level of qualifications PL – level of education
Household space	✓	✓	Households with residents Households with no residents
Accommodation type	✓	x	No question in Polish census
Building age	x	✓	No question in UK census
Tenure	✓	✓	Privately owned (owner occupied) Rented Other

Notes: ✓ - variable present and comparable, x – variable absent, ? - variable present, but not comparable
UK = United Kingdom, PL = Poland.

3.2.2. Census data for ethnic/affinity diversity

Ethnic/Affinity variables available in the two censuses include nationality (citizenship, country of birth), ethnicity (declaration), religion (declaration) and language (used at home). The comparison is provided in Table 3.5 and more detailed information in Table A1.3 in Appendix 1.

Cultural diversity in both censuses – due to different experiences with other cultures and migration of the native population – was measured differently. No variables are directly comparable. Some could be used for comparisons, but they will still differ in the type of data that are used to compute them.

Table 3.5. Comparison of ethnic/affinity variables in British and Polish censuses

Variable	Availability		Variable attributes
	UK	Poland	
Immigrant	?	?	Person with migrant origin outside UK = Long-term foreign immigrants
Foreigner	?	?	UK – place of birth PL – citizenship
Ethnic / national group	?	?	UK – 16 categories PL – 99 groups
Multiple ethnic identifier (household)	✓	x	No variable in Polish census
Citizenship	x	✓	No question in UK census
Religion	✓	x	No question in Polish census
Language used at home	x	✓	No question in UK census

Notes: ✓ - variable present and comparable, x – variable absent, ? - variable present, but not comparable
UK = United Kingdom, PL = Poland.

In this study, we straddle a major intellectual divide in thinking about people who belong to different ethnic groups, i.e. groups who share common cultural characteristics, like tradition, values, beliefs, religious faith, language, homeland and migratory status (Li, Skop 2009). Ethnic groups are differently captured in the statistical data in both countries. In the UK we have adopted the concepts of ethnic origin and self-identification to define the ethnic group. These are people with common racial and national characteristics who chose themselves which group to identify with when completing census or survey questions. Self-identification of ethnic group membership, used in this study, has no necessary connection to migration history or descendant generation. It means that socio-economic advantage/disadvantage and positive/negative discrimination on grounds of ethnicity can be tracked over all the immigrant generations.

In most European countries outside the UK, the concept used to define people of ethnic ‘difference’ is the state of being foreign. ‘Foreign’ can be defined in many different ways – being born in a foreign country, possessing citizenship of a foreign country, speaking the language of a foreign country as

mother tongue. The opposite of 'foreign' is being 'native', born in the current country of residence or having the citizenship of that state or speaking the main language of the country fluently. Immigrants who are foreign[21] will give birth to and raise children who are 'native'. Sometimes questions are included in censuses and surveys to identify the country of birth or citizenship of the respondent parents or, more rarely, grandparents as well. The consequence of this view of difference is over time people with immigration origins become 'invisible'. They are assumed to have assimilated into the national population and monitoring disadvantages/discrimination in terms of foreign background is not possible.

In some countries, such as Poland, use of a question of birth to capture foreign status is problematic. Depending on decade of birth in the 20th century, Polish speakers with Polish nationality who may not have moved at all in their lives could have been born in Germany, Austria, Russia or the Soviet Union if the country as constituted at time of birth is used. For example, according to the 2002 census data 70% of all foreign-born arrived from three countries: Belarus, Lithuania and Ukraine, and 82% of those persons were 60 years of age or older. People born in pre-war Polish territory, were displaced after the Second World War to Poland, because that land was annexed to Soviet Union (Okólski 2010). Hence the Polish census statistics we use in the research uses the nationality question rather that the country of birth question.

Which view of European populations that are different is correct – ethnicity or foreign/national status? The answer, of course, is that both concepts reveal important information about the social groups concerned, but both are inadequate. People's lives are influenced by both these factors. In this study we used data on Polish and foreign nationality in our Warsaw analysis, and data on White British and other ethnic groups in the Leeds analysis.

3.2.3. Census data for disability

There are two questions related to health condition and disability in the both cities (see Table 3.6 and Table A.1.4 in Appendix 1). However, in the British census all people were asked to assess their general health over the last year (on a scale bad, fairly good, good) and in the second question to specify whether they have a limiting long term illness, health problems or disability. In the Polish census people were asked about disability in two questions – in one about the legal disability (i.e. state approved conditions giving right to incapacity benefit) and in another one – about biological disability. The second category is the most similar to the British disability question. However, in the dataset we were working on, the disability data were not available for Warsaw.

[21] Note that the popular and demographic definitions of immigrant are different. Under popular definition immigrants are often assumed to be foreign, but under the demographic definition, necessary for balancing national and subnational population accounts, immigrant may be either foreign or native.

Table 3.6. Comparison of disability variables in the British and Polish censuses

Variable	Availability		Variable attributes
	UK	Poland	
General health	✓	x	No question in Polish census
Limiting long-term illness, health problem or disability	?	?	UK 'limiting long-term illness, health problem or disability' is a broader category than Polish 'legally' or 'biologically disabled'
Health and disability by household	✓	x	No variable in Polish census

Notes: ✓ - variable present and comparable, x – variable absent, ? - variable present, but not comparable
UK = United Kingdom, PL = Poland.

3.2.4. Final list of compatible variables

The final list of variables that were available in both the British and Polish censuses is as follows:

Table 3.7. List of the final variables available in the British and Polish censuses

Number	Variable
1	People aged 0-14
2	People aged 15-24
3	People aged 25-44
4	People aged 65 and more
5	Single people
6	Married people
7	Married people with children
8	Married people without children
9	Couples with children
10	Couples without children
11	Lone parents
12	1 person households
13	2 persons households
14	3 persons households
15	4 persons households
16	5 and more persons households
17	People with manual occupations
18	People with no qualifications/education
19	People with 1 level of qualifications/primary education
20	People with 2 level of qualifications/secondary education
21	People with 3 level of qualifications/postsecondary education
22	People with 4-5 level of qualifications/higher education
23	Households with occupancy rating -1 / substandard conditions
24	Household rented/owned by council
25	Non White British / non Polish nationality
26	People with limiting long-term illness (LLTI), health problem or disabled (UK)
27	People with good health (UK)
28	People with bad health (UK)

4. METHODS

In this section of the paper we discuss methods used to determine the nature of diversity in our two cities. They are designed to combine information of all groups to form profiles for zones and to cluster the zones into types on the basis of these profiles. This is a type of classification known as *cluster analysis*, widely used to understand how urban spatial communities are similar or different in the mixes of people and housing. The second set of methods involve measuring the level of diversity at zone or study area (the sum of zones) level, either examining one group by itself or comparing one group to another or summarising the degree of mixing of group within degree of mixing of groups within zones or study area. Details of this second set of methods are given in Appendix A2, where we use a common notation to present the algebraic formulae.

Cluster analysis is the assignment of objects belonging to a given group into a set of subgroups that are called clusters, so that the objects in the clusters are similar to one another in some sense. Clustering is very helpful in analysing big sets of data. The idea of clustering is reduction of a large amount of variation across N observational units to a smaller amount of variation across C clusters or types of area. During clustering two reductions occur. Firstly, V variables are reduced to W indicators through selection and through standardization (sometimes through factor analysis). Secondly, N zones are reduced to T types through one of the methods of cluster analysis. The process of reduction is illustrated in Figure 4.1.

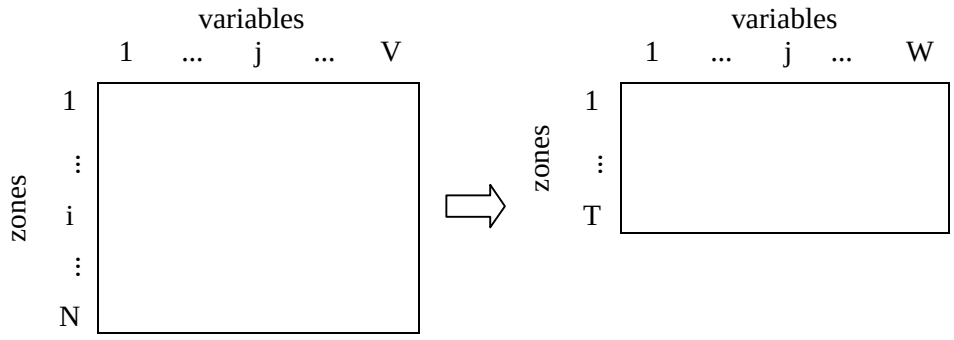


Figure 4.1. Data matrix

A comprehensive introduction to clustering methods is provided by Everitt et al. (2001), while Harris et al. (2005) show how clustering is used to develop geo-demographic types of neighbourhood widely used in marketing. Milligan and Cooper (1987) provide a particularly useful account of the procedures of cluster analysis, ordered step by step. Vickers et al. (2003) develop a classification of UK local authorities which has been widely used as model of cluster development, building on earlier work by Webber and Craig (1978). Vickers et al. (2005) describe the design of the official Output Area Classification (OAC) of 2001 Census Output Areas, while the key features are described in Vickers and Rees (2006, 2007). We did not use the OAC because it was too general for our purposes including

a wide range of census variables beyond those linked to the dimensions of difference which are the focus of this paper.

Cluster analysis applied in a socio-geographic context involves the steps set out in Table 4.1.

Table 4.1. Steps in the implementing a cluster analysis

Step	Description
(1)	The specification of the *aims* of the classification
(2)	The selection of *variables* and *spatial units*
(3)	Decisions about the *scaling or transformation* of variables or their conversion to principal components
(4)	Decision about the range of *numbers of cluster types* in which to search for a solution;
(5)	Choice of a *distance metric* to measure the separation in multi-dimensional social space of populations living in residential areas and a specification for *cluster centres*
(6)	Selection of an *algorithm and associated optimisation criteria* to create cluster types of residential areas
(7)	Choice of *software* for implementing a cluster analysis
(8)	Decisions about the starting assumptions for *cluster centres* in multidimensional social space because it is rarely possible with more than a small number of areas to evaluate all combinations of areas into cluster types
(9)	Decisions about selecting the best solution for *number of cluster types*
(10)	Construction of *cluster type profiles* using the input variables, a *succinct description* of each cluster type inventing a *name* for each type

4.1. Aims of the classification

The aims of the classification are to produce a typology of spatial communities differentiated in terms of the diversity of the people who reside there. The typology will help understand the spatial character of diversity in a city context. The typology will also be used to structure an interview survey of a sample of residents in the two cities. The types serve as strata for the sampling design in the survey, which will ensure equal samples within each type for robust comparison.

4.2. Selection of the variables and the spatial units

The first step in cluster analysis is to select relevant variables and to choose the method of standardization. We experimented with three different approaches to variable selection: (1) a simple analysis using three variables only – each variable representing one social dimension, (2) a complex analysis using a larger set of variables and (3) a diversity index analysis.

4.2.1. Variables selected for simple analysis

In the first phase of analysis, the simple analysis, only four variables have been selected. The idea was to represent each dimension by one variable. The variables were converted into percentage variables and the default z-score standardization of the variables was implemented later on. The variables for each dimension are presented in Table 4.2.

Table 4.2. Variables used in the simple cluster analysis

Diversity Dimension	Indicator used
Demographic diversity	Percentage of people aged 65 and more (v1)
Socioeconomic diversity	Percentage of people with manual occupations (5-9 SOC) (v2)
Ethnic/Affinity diversity	Percentage of non White British ethnicity or of non Polish nationality (v3)
Disability diversity (only for Leeds)	Percentage of population with limiting long-term illness (LLTI), health problem or disabled (v4)

Each percentage in the table refers to the share of the residential area population taken by the social group identified.

There were two ways that the simple analysis could be implemented: a simple clustering of the residential communities of each city separately and a simple clustering of the residential communities of both cities in combination. We wished to know whether the residential community types in the two cities could be similar or different. We report on both analyses.

4.2.2. Variables selected for complex analysis

It can be argued that four variables are too few and arbitrary to capture the variation across residential areas in social diversity. We drew up a long list of variables that were associated with our diversity dimensions and that were similarly measured in the UK and Poland censuses (Appendix A.1). These were then winnowed down to a final set of variables for analysis by excluding variables that did not have much explanatory value or were highly correlated one with another. The results of correlation analysis, multi-collinearity regression and variance inflation were examined to remove redundant

variables. The set of variables was reduced from 28 to 14. The final variables used in the complex analysis are shown in Table 4.3.

Table 4.3. Variables used in the complex cluster analysis

Diversity Dimension	Indicator used
Demographic diversity	Percentage of people aged 15-24 (v1)
	Percentage of people aged 25-44 (v2)
	Percentage of people aged 65 and more (v3)
	Percentage of married people (v4)
	Percentage of couples without children (v5)
	Percentage of lone parents (v6)
	Percentage of 4 and more persons households (v7)
Socioeconomic diversity	Percentage of people with manual occupations (v8)
	Percentage of people with higher education / 4-5 level of qualifications (v9)
	Percentage of household rented/owned by council (v10)
	Percentage of households with standard rate -1 /substandard (v11)
Ethnic diversity	Percentage of non White British / of non Polish nationality (v12)
Disability (only for Leeds)	Percentage of population with limiting long-term illness (LLTI), health problems or disabled (v13)
	Percentage of people with good health (v14)

4.2.3. Diversity analysis

In a diversity analysis we replaced the percentage variables with Simpson's Index of Diversity[22] using three groups for each dimension where possible (e.g. young, middle aged, older). This had value, but was difficult to interpret because diversity was low on the demographic and socio-economic dimension and high (in Leeds) on the ethnicity dimension. So we do not report in detail of the cluster analysis using diversity indexes, but do compute some indexes for clusters that were produced in the complex analysis (with percentages). Diversity indexes were used by in a recent analysis of the changing ethnic population of Leeds, showing that "the majority of Leeds' community areas have a low level of ethnic mixing" (Stillwell, Phillips 2006: 1150). We present the indexes in three and four dimensional bubble graphs. These were Simpson's Index of Diversity, Dissimilarity Index and Exposure Index. For each of social dimensions of diversity the indexes set out in Table 4.4 were computed:

[22] Simpson's Index of Diversity measures the possibility that two randomly selected individuals from a sample/group will belong to the same category. See Appendix A2 for more detailed the definition and the formula.

Table 4.4. Diversity, dissimilarity and exposure indexes used

Diversity Dimension	Index used	Social groups
Demographic diversity	Simpson Index	9 age groups (0-14, 15-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84, 85+)
Socioeconomic diversity	Simpson Index	9 occupation groups (9 SOC)
Ethnic/Affinity diversity	Simpson Index	9 ethnic groups in Leeds (White British, White Irish, White Other, Mixed, Indian, Pakistani, Other Asian, Black, Chinese)
	Simpson Index	Polish residents and foreign immigrants in Warsaw
	Dissimilarity Index	Polish residents and foreign immigrants in Warsaw
	Exposure Index	Polish residents to foreign immigrants in Warsaw
	Exposure Index	Foreign immigrants to Polish residents in Warsaw
Disability	Simpson Index	People with LLTI, bad health and disability and people without LLTI, bad health and disability in Leeds
	Dissimilarity Index	People with LLTI, bad health and disability and people without LLTI, bad health and disability in Leeds
	Exposure Index	People with LLTI, bad health and disability to people without LLTI, bad health and disability in Leeds
	Exposure Index	People without LLTI, bad health and disability to people with LLTI, bad health and disability in Leeds

4.3. The scaling or transformation of variables

We decided against transformation to either the percentage variables or the diversity indexes (e.g. by applying logarithm transforms) in order to maintain transparency in the final results. Here we follow the argument put forward by Senior (2002) in his review of indexes of deprivation.

4.4. The numbers of cluster types

The number of cluster types to be sought depends on the purpose of the classification and whether a hierarchy of types is needed. So, for example, the National Statistics Output Area Classification (Vickers et al. 2005, Vickers and Rees 2007) has three nested tiers: super-groups with 7 types, groups with 21 types and subgroups with 52 types. Other geo-demographic classifications adopt a similar nested hierarchy. For the present purposes we only sought to create one tier of types for descriptive purposes. Geo-demographic classifications employ between 5 and 20 clusters in their upper tier. Above 20 it is difficult to remember and understand a classification.

4.5. A distance metric and cluster centres

Cluster analysis works with matrices of distances between objects and clusters in the *n*-dimensional space of the variables. Our objects are the Community Areas of Leeds and the Urban Regions of Warsaw. Our variable space in the simple clustering has 3 dimensions and in the complex clustering has 14 dimensions. A large number of distance metrics are available and users have experimented extensively with different metrics. Aggarwal et al. (2001) demonstrated that Manhattan distance (the sum of the distance on the opposite and adjacent sides of a right angle triangle) was theoretically superior to Euclidean distance (the hypotenuse of the right angle triangle or the square root of the sum of squared adjacent and opposite sides) or its square. Dennett and Stilwell (2009) showed that Manhattan distance produced better, clearer clusters of local authorities in terms of the migration behaviour of their inhabitants. We also experimented with squared Euclidean distance and Manhattan distance and found the latter gave more coherent and interpretable cluster solutions.

The simple cluster analysis was carried out also using another distance measure – squared Euclidean distance. However, the absolute average difference from mean cluster sizes were higher than when Manhattan distance was used, and a comparison of silhouette plots revealed that the cluster solutions were of worse quality. So only Manhattan distance was used in the combined simple analysis and the complex analysis.

It is necessary in all clustering algorithms to compute a representative set of coordinates for a cluster centre. Two alternatives are available. The first is the mean value of the variable values of objects belonging to the cluster. The second is the median of the variable values of objects in the cluster. We used the latter, which avoids excessive influence of objects distant from the centre of the cluster.

4.6. A clustering algorithm

There are two general methods of clustering: hierarchical and non-hierarchical.

Hierarchical clustering groups objects by creating a cluster tree or dendrogram. The tree is not a single set of clusters, but rather a multi-level hierarchy, where clusters at one level are joined as clusters at the next level. You do not have to decide the final number of clusters, but instead decide on which level of clustering is most appropriate (Beale et al. 2009). Ward's algorithm for hierarchical clustering (Ward 1963) has proved popular but suffers from disadvantages. It is difficult to implement with a large number of objects in many software packages. The clusters it produces tend towards a single large cluster centred on the mean values of the variables with small outlier groups.

Because we aimed to produce only a single tier classification, we adopted *a non-hierarchical method*. This can be used to generate a hierarchy if needed by application to the clusters of each tier successively (e.g. Vickers et al. 2005).

The non-hierarchical method we used is called *K-means*. The K-means method of clustering iteratively swaps objects into and out of clusters searching for the best solution aiming for maximum between cluster variation and minimum within cluster variation. K-means clustering partitions the *n* points in the *n*-by-*p* data matrix *X* into *k* clusters. Each observation treated is like an object having a location in the multi-dimensional space of the chosen input variables. K-means uses an iterative algorithm that minimizes the sum of distances from objects to the cluster centre, over all clusters (Beale et al. 2009). Unlike hierarchical clustering, K-means creates a single level of clusters. K-means can be run using different distance measures such as Euclidean, squared Euclidean, Manhattan, cosine, correlation or Hamming distance (section 4.5).

4.7. Software for implementation

To carry out the clustering we used the software package MATrix LABoratory Statistics Toolbox (MATLAB) (Beale et al. 2009) in preference to the frequently used Statistical Package for the Social Sciences (SPSS). Many cluster analyses in the past (Vickers et al. 2003, 2005) have used the SPSS clustering routine which does not offer automatic replicates. MATLAB was preferable because the software enables the user to request that the clustering be repeated T times with different selections of the starting cluster centres (a random sample of the objects). The algorithm was run for a specific number of times starting each replicate with a randomly selected initial cluster centres. From the resulting set of T cluster analyses, MATLAB provides a suggested optimum typology.

4.8. Replication of initial cluster centres

MATLAB enables the user to run clustering replicates. What these are is alternative starting allocations for the centres of potential clusters. The cluster solutions depend on the starting allocations of objects as centres. If we wish to find the best solution for *c* clusters, then there are *n!(n-c)!* ways starting the search where *n* is the number of objects to be classified (106 Community Areas and 91 urban regions). The term *n!(n-c)!* is always a very large number. What MATLAB does is run a small number of replicates and search for the best solution among the set in terms of the solution that maximizes between cluster variance and minimizes within cluster variance. This does not guarantee that the optimum solution has been found, of course, but experience with the package suggests that running several hundred replicates produces solutions close to the optimum. So in cluster analysis analyses reported here, we used MATLAB to produce 200 replicates from which to produce the best choice with 1000 replicates.

4.9. The best number of cluster types

K-means clustering generates multiple solutions k=2, 3, 4 up to a maximum of n-1 clusters. From the set of different cluster typologies, the user must select the best using one or more of several criteria. Large numbers of clusters provide little gain in generalization, so we focused on examining from 2 up to 16 cluster solutions only. Cluster membership distributions across cluster solutions for 106 Community Areas and 91 Urban Regions generated using Manhattan distance measure and replicated 200 times are presented in Appendix A6. The best cluster solution was sought by applying two types of diagnostics: average absolute intra-cluster distances and silhouette plots.

Absolute average differences from mean cluster sizes for the cluster solutions were compared. It is preferable to have clusters with a similar size in term of number of members, clusters with few members will be ignored and one cluster with a large membership fails to supply information. The best cluster solution will have an average value closer to 0. This test enabled us to select cluster solutions with the most even distribution of cases. Relatively low average differences were obtained for six and higher cluster solutions for Leeds, for four and higher cluster solutions for Warsaw and for five and higher cluster solutions for the combined analysis (see Figures 4.2, 4.3 and 4.4). The lower cluster solutions did not perform well so we excluded solutions with only 2 to 5 clusters.

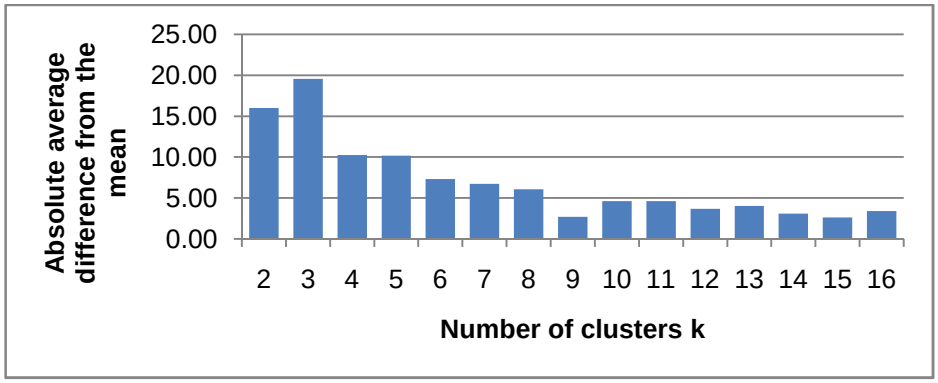


Figure 4.2. Absolute average difference from mean cluster size for 106 Community Areas in Leeds (Manhattan distance, 200 replicates): complex cluster analysis

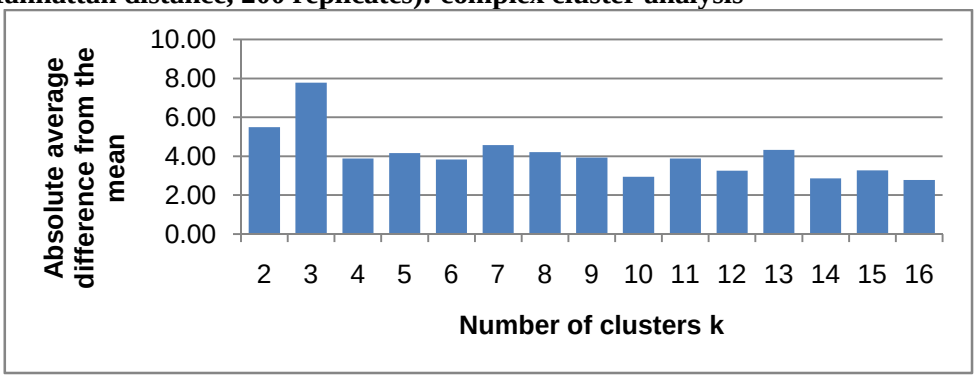


Figure 4.3. Absolute average difference from mean cluster size for 91 Urban Regions in Warsaw (Manhattan distance, 200 replicates): complex cluster analysis

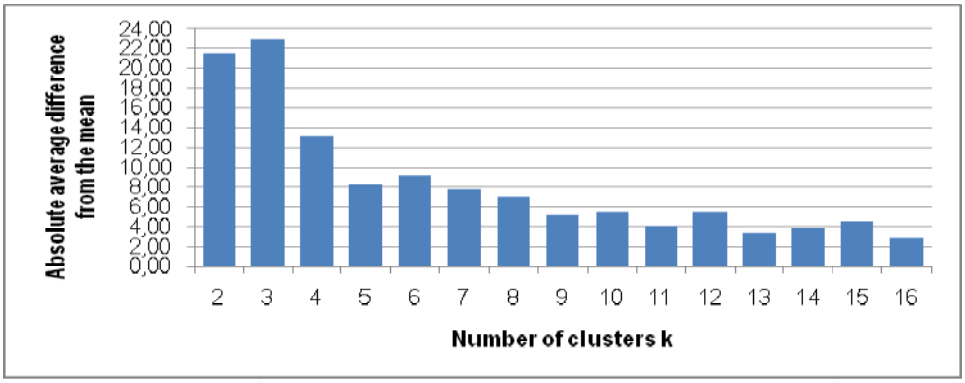


Figure 4.4. Absolute average difference from mean cluster size for 197 Community Areas and Urban Regions in Leeds and Warsaw (Manhattan distance, 200 replicates): simple analysis

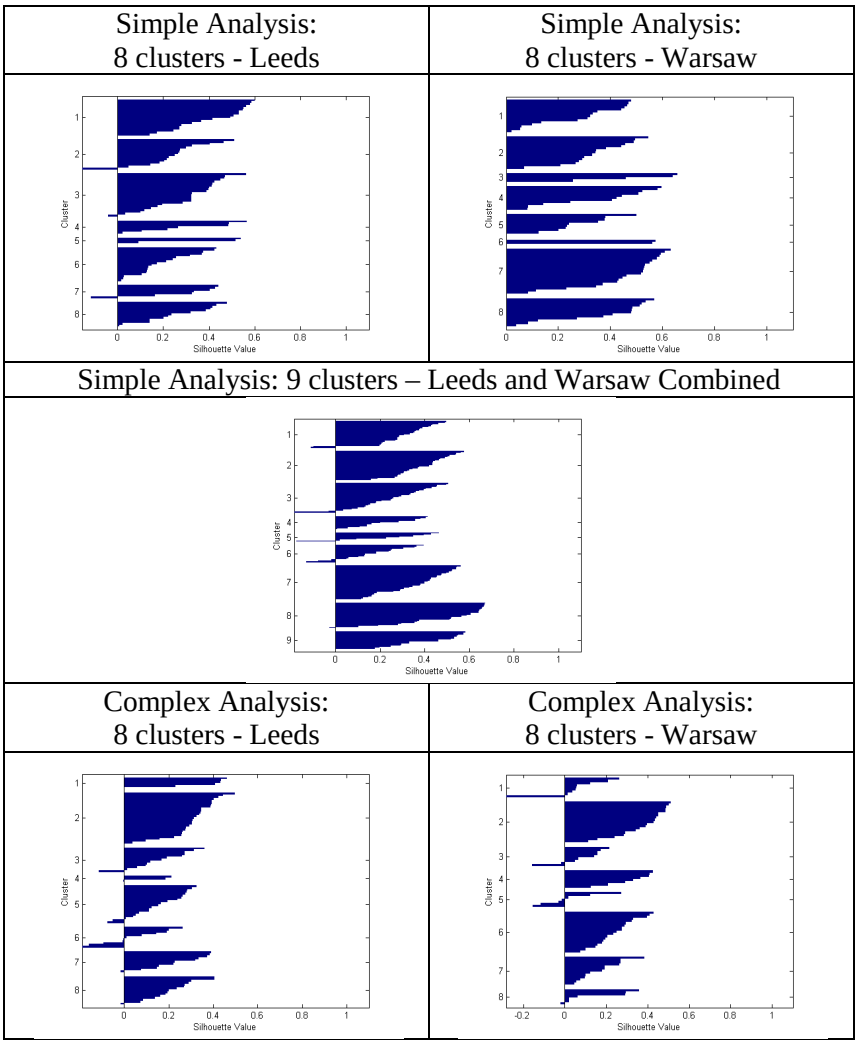


Figure 4.5. Silhouette plots for the final cluster solutions for Leeds and Warsaw

In a second step, the silhouette plots of cluster solutions were compared. A silhouette plot is a way to graphically display a sorted list of similar objects which displays a measure of how close each point in one cluster is to points in the neighbouring clusters (Beale et al. 2009). Not only did this enable us to select the cluster solution with the most uniform cluster size, but also we could exclude poor quality

solutions, i.e. cluster solutions that had the highest number of cases that could be ascribed to another cluster (these are observations with negative values in the plots). All silhouette plots for 2-16 cluster solutions are presented in Appendix A7. Figure 4.5 displays the silhouette plots for the final, selected cluster solutions for Leeds and Warsaw, using the 4 variables (simple clustering) and the 14 variables (complex clustering) and for the combined, simple analysis.

In the separate analyses for both cities, the eight cluster solutions proved to be the most coherent and had equally distributed cases, while for the combined analysis the nine cluster solution was the best. Finally, a 1000 replicate k-means algorithm was run for these cluster solutions to improve the final classifications.

Moreover, the cluster analysis was run using diversity indexes (see Appendix A2 for formulae) rather than percentages for three/four other variables: the Diversity Index for 3 age groups, the Diversity Index for 3 occupational groups, the Diversity Index for 3 ethnic groups and additionally for Leeds the Diversity Index for 3 health groups (with both distance measures). However, the diversity indexes proved to be bad variables for cluster analysis. The silhouette plots were of poor quality, i.e. many cases could be classified into another cluster and consequently the interpretation of achieved cluster solutions was difficult. Diversity maps are provided in the Appendix A8.

Table 4.5 summarizes what was done.

Table 4.5. Cluster analyses undertaken in search of a perfect cluster solution

Cluster label	Number of variables	Variables metric	Distances used	Diagnostics	
				Average intra-cluster distance	Silhouette plots
A	Simple Separate (3)	Percentages	Squared Euclidean	✓	✓
B			Manhattan	✓	✓
C		Diversity Indexes	Squared Euclidean	✓	✓
D			Manhattan	✓	✓
E	Simple Combined (3)	Percentages	Squared Euclidean	x	x
F			Manhattan	✓	✓
G	Complex (12/14)	Percentages	Squared Euclidean	x	x
H			Manhattan	✓	✓

Note: ✓ - diagnostic implemented, x – diagnostic not implemented.

4.10. Description of the clusters

A final step in the analysis is to describe the clusters. Several devices are used to help understand the nature of each diversity cluster: a graphical plot of cluster mean variable values, summary statistics for the cluster, a map showing the location of areas which belong to the cluster type and a list of area names. Several graph types can be used for profiling. Vickers et al. (2005) used radial plots, for example. In this paper we bar charts of the z-score for the variable means of clusters. This information enables the writing of a summary of the characteristics of the cluster type. Adding a naming phrase is the last thing to do, which is fraught with difficulty as it is very difficult to capture the essence of the diversity of an area in a few words. In the Output Area Classification, the then Director of National Statistics vetoed the use of cluster names, wishing to avoid any official stigmatization of Output Areas. Users need to consult the academic papers and web pages to learn about the names invented by the authors (Vickers et al. 2005).

5. DIVERSITY DESCRIBED ONE DIMENSION AT A TIME

The aim of the analysis is to understand the nature of multi-dimensional diversity in Leeds and Warsaw, using three or four selected dimensions of diversity: age, occupation, ethnicity and disability. In this section we present the results employing just one variable to represent each dimension at a time.

Statistical data for the whole population in both cities give a first overview of the difference between the two cities as well as a variation of the selected variables in each city (Table 5.1). In both cities reside a similar percentage of people aged 65 and more – 16%. While in Warsaw we have larger economically active population with non-manual occupations, in Leeds the share of minorities ethnic is noticeably higher. More detailed statistics are presented in a table below and spatial distribution of population with selected characteristics is provided in the next section.

Table 5.1. Statistical profiles for Leeds (2001) and Warsaw (2002) across the dimensions of diversity

Statistic[1]	Percentage of the population						
	Aged 65 and over		Manual occupations		Ethnic minorities or foreign immigrants		Disabled[2]
	Leeds	Warsaw	Leeds	Warsaw	Leeds	Warsaw	Leeds
Mean	15.5	16.5	47.5	32.0	10.8	0.6	18.0
Median	15.8	14.2	48.3	33.3	5.3	0.5	17.9
Std. Dev.	3.6	6.8	13.7	12.4	11.2	0.4	4.0
Minimum	3.9	2.4	21.7	11.0	1.4	0.0	8.9
Maximum	23.1	29.3	72.7	76.7	68.6	2.7	30.4
Percentiles 10	10.8	6.3	29.2	22.2	2.8	0.2	13.2
50	15.8	14.2	48.3	33.3	5.3	0.5	17.9
90	20.3	26.0	66.8	49.2	24.1	1.0	24.1

Notes:
1. We use 106 Community Areas in Leeds and 91 Urban Regions in Warsaw.
2. No Disability variable is available for Warsaw.

5.1. Demographic diversity (age)

The range of age diversity is higher for Warsaw – it is almost 27 percentage points between the minimum – 2.4%, and the maximum – 29.3%. In Leeds the lowest share of people aged 65+ at the level of Community Areas was 4% and the highest was 23%, with an average of 19%. In Leeds the median value for population aged 65+ was slightly higher 16% in comparison to a 15% median value for Warsaw. Figures 5.1 and 5.2 show the distributions of areas in Leeds and Warsaw across the percentage 65+ range.

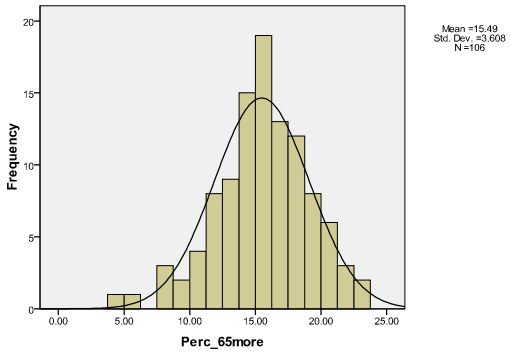


Figure 5.1. Percentage of population aged 65 and more in Community Areas in Leeds, 2001

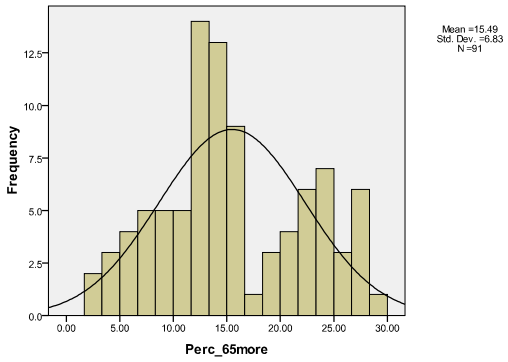


Figure 5.2. Percentage of population aged 65 and more in Urban Regions in Warsaw, 2002

Maps of the age variable (Figures 5.3 and 5.4) reveal substantial differences between the cities. While in Leeds the older population tends to live in the outer areas of the city, in Warsaw they live in the city centre. The 'oldest' Community Areas in Leeds are suburban areas Moor Allerton, Garforth West, Harewood and District, Collingham and Linton and in the city centre New Wortley. In Warsaw the 'oldest' Urban Regions are Stare i Nowe Miasto, Centrum-Północ, Muranów Zachodni, Muranów Wschodni, Mokotów Centrum, Żoliborz Zachodni and Pole Mokotowskie, all located in the centre of the city.

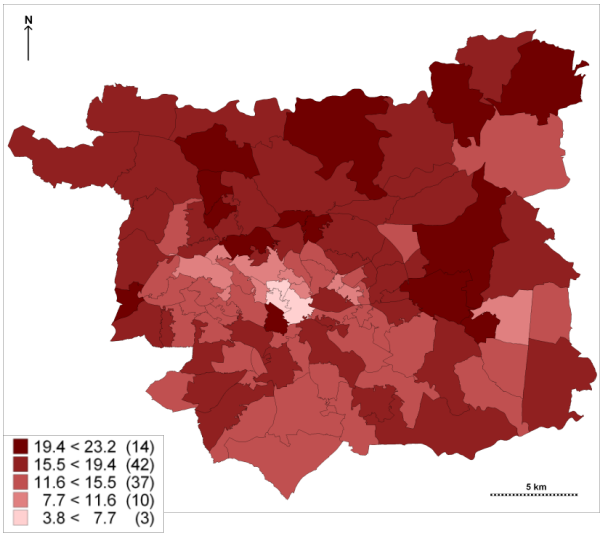


Figure 5.3. Percentage of population aged 65 and over in Community Areas in Leeds, 2001

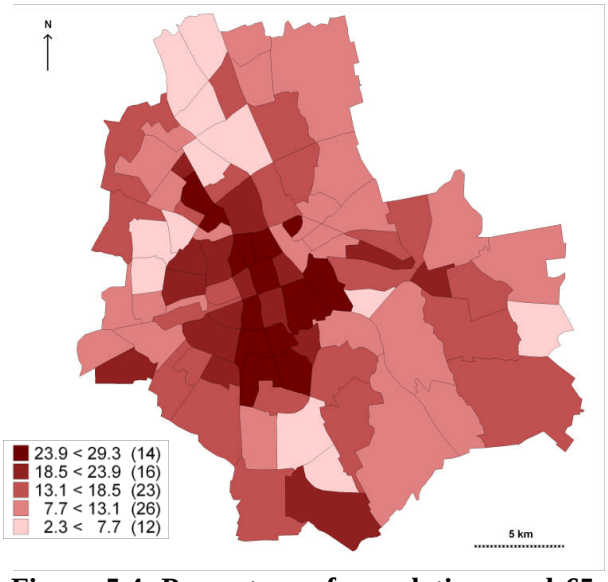


Figure 5.4. Percentage of population aged 65 and over in Urban Regions in Warsaw, 2002

In Leeds the areas with the lowest concentration of older people are found in the residential districts northwest of the city's two centrally located universities (University of Leeds and Leeds Metropolitan University). Student households have replaced households with older residents, and these households have migrated away or when the death of a single old person living alone has created a vacancy. The

ring of suburbs with older populations results from ageing of people who moved into these areas when younger. Once families are settled in the suburbs, they stay a long time and age in place.

Warsaw is a good example of a city with quite distinct mosaic of areas dominated by various age groups. This is a direct consequence of the post-war development of the city, with new, often large, areas of the city being developed over a relatively short period of time and settled with people of similar age. The more general process of increasing spatial segregation of elderly (65+) was noted by Węcławowicz (1988a, 1991). An overview of the distribution of elderly in Polish cities (Węcławowicz 1988b) shows quite persistent pattern of their concentration in city centres and correlation between higher shares of elderly in population and the share of housing constructed prior 1944, generally indicating the forming of areas of old and impoverished population. This process seems to has petrified or even accelerated after the fall of communism (Węcławowicz et al. 2006)

In comparison to Leeds, Warsaw is in a different stage of 'city life cycle' experiencing suburbanization and population de-concentration on an agglomeration scale (Kupiszewski et al. 1998; Śleszyński 2004). The territorial expansion of city followed sectoral pattern of development with large new estates being built in suburban zones: Ursynów in 1970s, Natolin and Kabaty from 1980s till now and Wilanów and Białołęka expanding in the last decade. The areas outside the city centre have not yet aged, but probably we will witness the same gentrification processes and ageing of outer Urban Region in near future. The youngest population lives in newer housing estates in northern, south-western and southern Urban Regions in Warsaw, but also new families are moving outside the city (Potrykowska, Śleszyński 2001).

5.2. Socio-economic diversity (occupation)

The working population of Warsaw was more diversified than that of Leeds in terms of type of occupation. Note that the classification of occupations in the British and Polish censuses is very close as both are aligned to the International Standard Classification for Occupations (1988 version – see Appendix A1, Table A1.2). The difference in manual/non-manual occupational shares results from the different functional roles played by the two cities. Leeds is a regional metropolitan centre, with employment diversified across a range of manufacturing and service industries, but without the national government business headquarters and international agencies that are found in Poland's capital city.

Although the average share of people with manual occupations was lower in Warsaw than in Leeds – i.e. 35% rather than 49%, there is a similar range. There are Urban Regions in the Polish capital city with only 11% of people with manual occupations, but also some Urban Regions with 77%. In Leeds the value of the manual/non-manual variable varies between 21 and 72%. The median value of

residents with manual occupations for the Community Areas of Leeds is 48%, while for Warsaw the median value is much lower at 33%. The distribution of the manual occupation percentages in Leeds is graphed in Figure 5.5 and that for Warsaw in Figure 5.6.

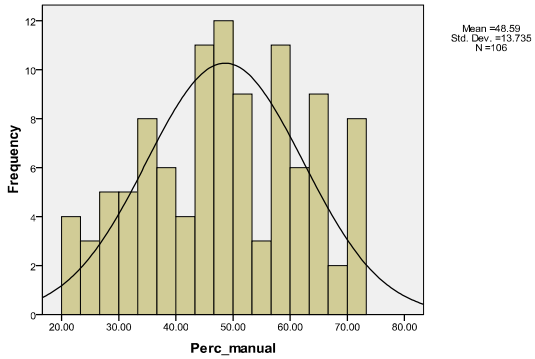


Figure 5.5. Percentage of population with manual occupations in Community Areas in Leeds, 2001

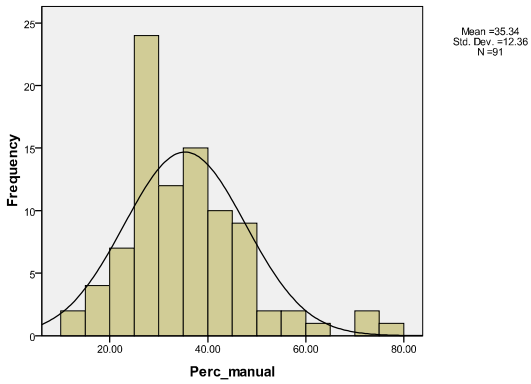


Figure 5.6. Percentage of population with manual occupations in Urban Regions in Warsaw, 2002

There are some visible patterns in the spatial distribution of people with manual occupations in both cities (Figures 5.7 and 5.8). In Leeds the working class population tends to live in the city centre and the southern parts of the city. These are Community Areas such as Osmondthorpe, New Wortley, Seacroft South, Swarcliffe, Seacroft North and Cross Green. The spatial structure is different in Warsaw – the city centre is inhabited by people with non-manual occupations, and working class people are located in some former-manufacturing areas in the northern and peripheral areas, e.g. Huta Warszawa, Targówek Przemysłowy, PKP-Wola, Odolany, Mańki-Brzeziny and Żerań Wschodni.

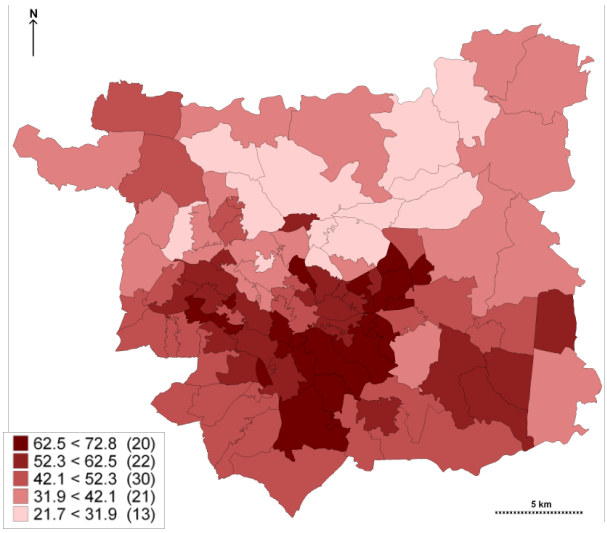


Figure 5.7. Percent of population with manual occupations in the Community Areas of Leeds, 2001

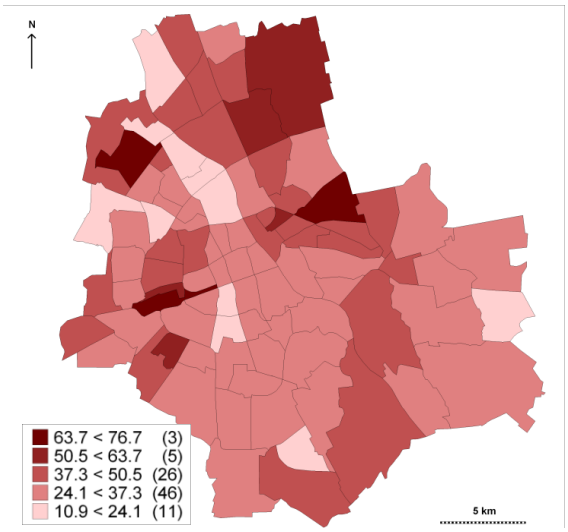


Figure 5.8. Percent of population with manual occupations in the Urban Regions of Warsaw, 2002

In Leeds the geography of social class has its origins in the Industrial Revolution in the 18th century and in the industrialization of the 19th century. The city's iron works and engineering industries were located south of the river Aire which flows through the city from north west to south east. Working class areas of the city such as Hunslet/Stourton, Holbeck and Middleton house the 'shock troops' of the Industrial Revolution. Between Hunslet and Middleton ran the one of first steam locomotive railways and around the city centre were established large clothing factories which employed female sewing operatives. Workers lived in dense streets of terrace and back to back housing, the most of which has been replaced by public or social housing in the 20th century. A detailed account of the social class geography of Leeds is provided in Stillwell and Shepherd (2004).

In Poland studies based on factor analysis method show that the 'socioprofessional position', largely determined by the level of education, dominates the process of spatial segregation, with uneducated people being pushed away from attractive areas of the city, in Warsaw, in particular form the Vistula river escarpment (Węcławowicz 1979; 1988). There are some areas of concentration of people belonging to working class in Warsaw. The socio-economic spatial disparities are a legacy of Warsaw being a 'socialist city' for several decades after Second World War. Communists aimed at reduction of pre-war socio-economic disparities, settling working class representative into the Warsaw most central space. MDM (*Marszałkowska Dzielnica Mieszkaniowa*) was perhaps the best example of such policy. Węcławowicz et al. (2003) noted that the socio-economic differentiation has reduced until 1970s and started to rise since then, as Dangshat and Blasius (1987) have shown.

The another general tendency was to 'plant' factories into urban areas in order to lower the percentage of middle-class people and increase share of working class (proletariat) living in cities (Węcławowicz 2005). Therefore, the 'darker' areas in the Figure 5.8 reflect spatial differentiation created under socialism, when factories were located there such as the steelworks set up in 1952 in northern Warsaw (Huta Warszawa) or the car factory founded in 1951 in the north-eastern Warsaw (*Ż*erań Wschodni). Working class concentration stems from the fact that in these Urban Regions the communist authorities provided housing flats for people working in the factories on previously rural land. However, some studies undertaken under socialist have shown that socio-economic inequalities have risen and Warsaw urban structure was 'mosaic' like, i.e. flats of poor people were neighbouring those of better-off people (Węcławowicz 1979).

Recently Węcławowicz et al. (2006) noted that the wealth connected with socio-economic status is the main driver of spatial segregation, leaving poorer people in the city centre unable to benefit from post-transition social and economic improvements. The transition allowed the expansion of housing for more affluent households on the outskirts of the city. Outside the city centre new gated housing areas were formed, separating those who have from those who do not.

5.3. Ethnic diversity (non White British / foreign immigrants)

Foreigners are more evenly distributed in Warsaw than are ethnic minorities in Leeds. However, the share of residents without Polish citizenship is very low in the Polish capital city – 0.6% –and reached a maximum of 2.7% in one Urban Region. In Leeds the lowest share of ethnic minorities was 1.4%, and at the same time there were some Community Areas where ethnic minorities constituted in fact the majority of the population – up to 70%. Still most of the Community Areas in Leeds have between 1.4% and 15% non White British population. Because of the contrast between an ethnically diverse Leeds and an ethnically homogenous Warsaw, we cannot use the same class intervals on the graphs (Figure 5.9 and 5.10) and the maps (Figure 5.11 and 5.12)

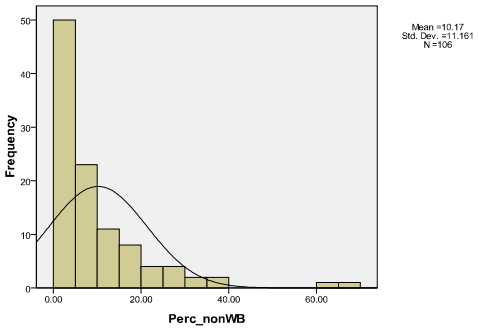


Figure 5.9. Percentage of non White British in the Community Areas of Leeds, 2001

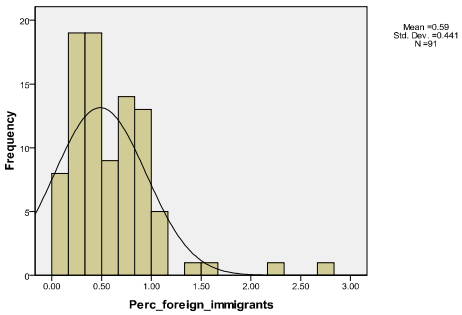


Figure 5.10. Percentage of foreign immigrants in the Urban Regions of Warsaw, 2002

The highest share of non White British people lives in the city centre and central-north areas of Leeds. These are Chapeltown, the Harehills Triangle, Burley Lodge and Little Woodhouse and Little London. Similarly in Warsaw foreigners live in the central part of the city, in regions like in some areas in the south, e.g. Czerniaków Wilanów and Mokotów-Centrum, and in the northern part of the city, e.g. Młociny, Dąbrówka, Żoliborz Przemysłowy, but also in city centre: Centrum-Północ, Mirów, Ochota-Centrum and Szosa Krakowska (Szczęśliwice).

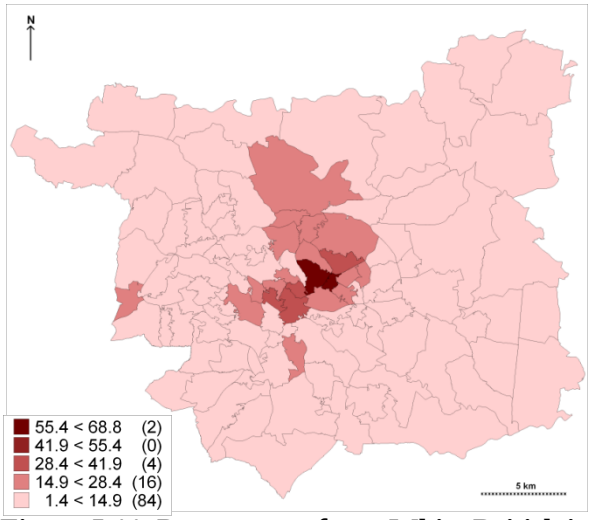


Figure 5.11. Percentage of non White British in Community Areas in Leeds, 2001

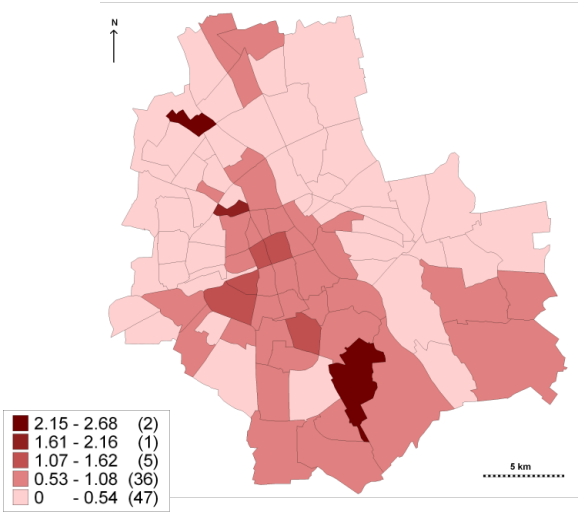


Figure 5.12. Percentage of foreign immigrants in Urban Regions in Warsaw, 2002

The Leeds map (Figure 5.11) embeds the history of immigration and subsequent internal migration of ethnic minorities. Inner north Leeds neighbourhoods served as the 'port of entry' for newcomers. These areas were selected because housing was cheap and former middle-class houses could be converted into multi-occupied dwellings. As the minority ethnic groups became more established and improved its socio-economic positions, minority group families were able to migrate further out from the city centre. The port-of-entry neighbourhoods shifted as slum housing was replaced by business use in the near city centre districts. This sequence of immigration and outward movement characterized the Jewish immigration of the late 19th and early 20th centuries, the West-Indian migration of the 1950s into Chapeltown and later into Chapel Allerton and the Indian migration later into Moortown and Roundhay. Other Asian groups (Pakistanis and Bangladeshis) established new housing enclaves in Burley Lodge and Little Woodhouse and the Harehills Triangle and south of the river Aire in Beeston Hill. The residential neighbourhoods around the two Leeds Universities (Hyde Park, Woodhouse and City Centre) have seen increased numbers of foreign students from a range of diverse backgrounds. The areas surrounding the zone of minority ethnic groups concentration have small numbers who are not White British with one exception, the Priesthorpe area on the western boundary on Leeds. Here the ethnic minority population is Asian (Pakistani and India) who have migrated outwards to southern Leeds from the inner wards of Bradford. More details of the minority ethnic geography of Leeds can be found in Phillips et al. (2004) and Harland and Stillwell (2007) who used 2001-2006 Pupil Level Annual School Census (PLASC) data to prove ethnic minority dispersal in Leeds.

The residential neighbourhoods in the city centre of Warsaw with their concentration of housing flats and high rented housing stock, prove to be the most frequently chosen by newcomers, including foreign immigrants. Vietnamese and other Asians tend to live in clusters close to open markets, for

example in the triangle located south-west from the city centre, close to the open market in Banacha Street (Ochota-Centrum and Szosa Krakowska/Szczęsliwice neighbourhoods). Another centrally location of foreign residence is the 'Iron Gate' estate, huge block of flats with small apartments, ('Osiedle za Żelazną Bramą') that had good communication with the 10th Anniversary Stadium placed at the other bank of the Vistula river, where thousands of petty-traders with diverse ethnic background used to work in 1990s and in the beginning of 2000s. The Stadium served as a social space for many foreign communities, especially the Vietnamese (Grzymała-Kazłowska 2002). However, with preparations for Euro 2012 football cup, the Stadium was closed for traders and in turn it affected the interurban migration of foreign residents. Many of them have moved outwards from the city centre to the south-west areas of Warsaw, near Okęcie airport, because these areas are relatively close located to big wholesale markets in the south-west suburbs outside the city (namely in Wólka Kossowska commune). In contrast to Vietnamese, Ukrainian and other former Soviet Union immigrants are much more dispersed in the city area and also their city concentration is lower, as some of them live outside city boundaries where flats are cheaper (Grzymała-Kazłowska, Piekut 2007). Refugee centres are other foreign immigrant concentration spaces. There used to be three of them in the city and six more in Warsaw suburbs (operating in different periods). Some are already closed, but during 2002 Census one centre in the northern Warsaw was operating, that is why we can see a concentration of foreigners in the Młociny Urban Region (the refugee centre was in Improwizacji Street). Finally, due to the concentration of international companies and foreign representative institutions, there are some 'expatriate communities' in Warsaw. Especially in the southern Warsaw and Wilanów district, with its prestigious housing estates and international schools, there are some Urban Regions where immigrants live, who come from West European and North American countries, and also from other countries but undertaking highly skilled jobs in big companies or being diplomats. Some of the estates are closed communities, e.g. with guards, fences and can be accessed by their residents only, and they could be called 'gated communities'.

5.4. Health diversity (LLTI, disabled)

About 18% of Leeds residents declared in the 2001 census that they have limiting long-term illness (LLTI), health problems or are disabled. In some of the Community Areas LLTI or disability was reported by three of ten residents. The lowest value of this variable was 9%. Figure 5.13 shows the distribution of Community Areas.

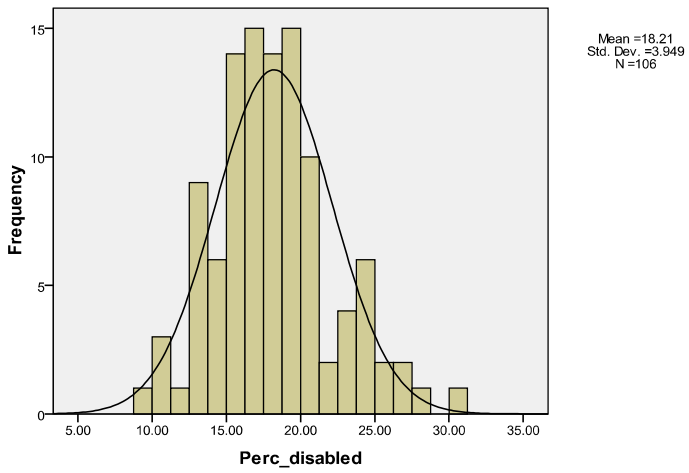


Figure 5.13. Percentage of population with limiting long-term illness, health problem or in the Community Areas of Leeds, 2001

Residents with LLTI and disabled are unevenly distributed in the city. The highest share of this population lives in the following CA: New Wortley, Moor Allerton, Gipton South, Hunslet/Stourton, Burmantofts, Lincoln Green and Ebor Gardens and Osmondthorpe (Figure 5.14). The lowest share of population with LLTI and disabled tend to live in Hyde Park and the City Centre (student residential areas) or in Scarcroft and Garforth East (suburbs with young families).

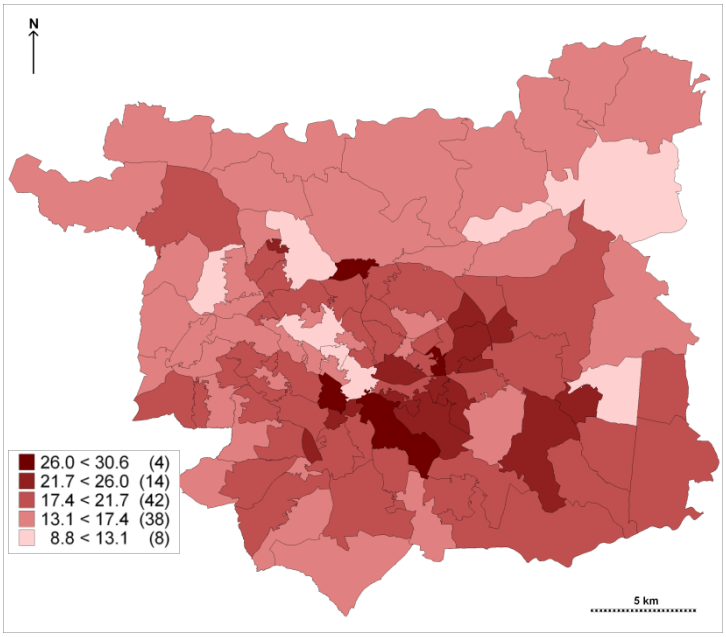


Figure 5.14. Percentage of population with limiting long-term illness or health problem in Community Areas in Leeds, 2002

Limiting long-term illness increases exponentially with increasing age (Rees et al. 2009), so that most studies of illness standardise the variable by age and sex and compute a Standardized Illness Ratio or SIR (e.g. Brown, Rees 2006). Regression models of the determinants of SIR variation across local

authorities show that almost 70% of its variance across areas is associated with material deprivation (Brown, Rees 2006). So the first two of our social dimensions – the demographic and the socio-economic dimensions – can be used to reproduce this ill-health dimension. Studies using micro data, not subject to the potential biases of ecological regression, confirm that as people age the chances of being disabled increase, but that this relationship differs in level between the social classes.

There were two disability questions in the Polish census of 2002 (for details see Appendix A1), but the data were not available, so a comparison of disability level could not be made.

6. MULTI DIMENSIONAL DIVERSITY

In previous sections we considered our four social dimensions, as revealed in Leeds and Warsaw, independently. In the current section we consider them jointly at Community Area and Urban Region level. We aim to discover which types of Community Area or Urban Region are similar in their degrees of diversity across the three or four social dimensions. To do this we use the tools of cluster analysis as discussed in Section 4.

6.1. Cluster profiles in Leeds

In the simple analysis cluster profiles were created for the final eight cluster solution for Leeds. The bar chart in Figure 6.1 shows the distribution of variables by cluster. The values are standardised. The numbers on the Y scale represent the difference from the mean value for the variable: percentage of aged 65 and more, percentage of people with manual occupations, percentage of non-White British and percentage of population with LLTI or disabled. Figure 6.2 maps the Community Areas by cluster number.

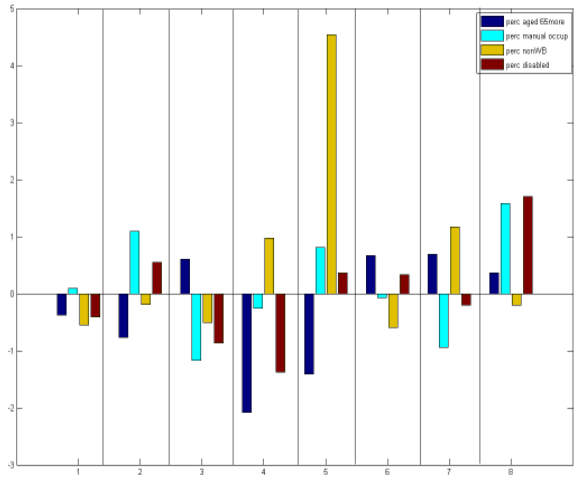


Figure 6.1. Key variables for the 8 cluster solution for Leeds, simple analysis, 2001

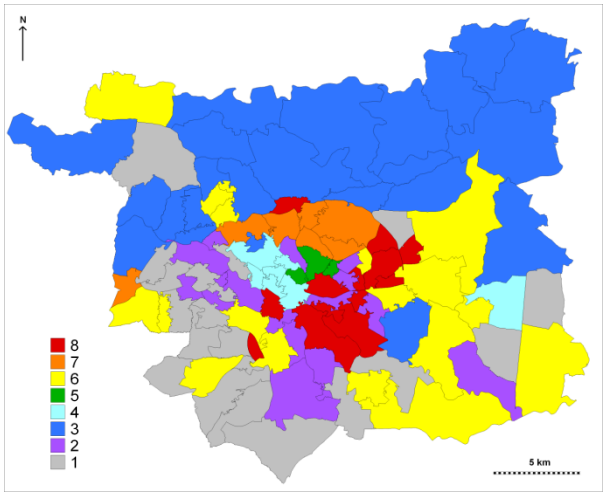
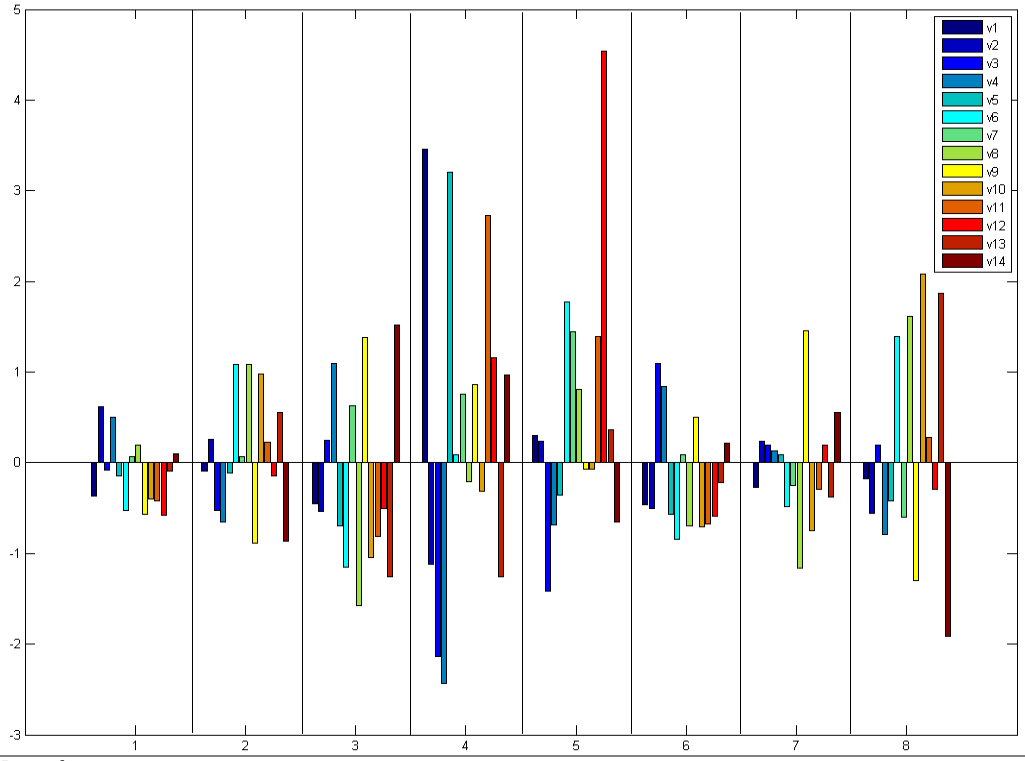


Figure 6.2. A map of the classification of Community Areas in Leeds into 8 clusters, simple analysis, 2001

The simple analysis methodology produced a sensible cluster solution, i.e. the results were easy to interpret and cluster profiles were considerably different one from another. More statistics on the simple analysis and Community Areas classification are available in the Appendix A5. Next we implemented the same clustering methodology including more variables. In the complex analysis the eight cluster solution proved to be most coherent too. The statistics for complex cluster solution are presented in Figure 6.3 and Table 6.2 and the classification is mapped in Figure 6.4.



Legend:

v1 - Percentage of people aged 15-24
v2 - Percentage of people aged 25-44
v3 - Percentage of people aged 65 and more
v4 - Percentage of married people
v5 - Percentage of couples without children
v6 - Percentage of lone parents
v7 - Percentage of 4 and more persons households
v8 - Percentage of people with manual occupations
v9 - Percentage of people with 4-5 level of qualifications
v10 - Percentage of household rented from council
v11 - Percentage of households with standard rate -1
v12 - Percentage of non White British
v13 - Percentage of population with limiting long-term illness (LLTI), health problems or disabled
v14 - Percentage of people with good health

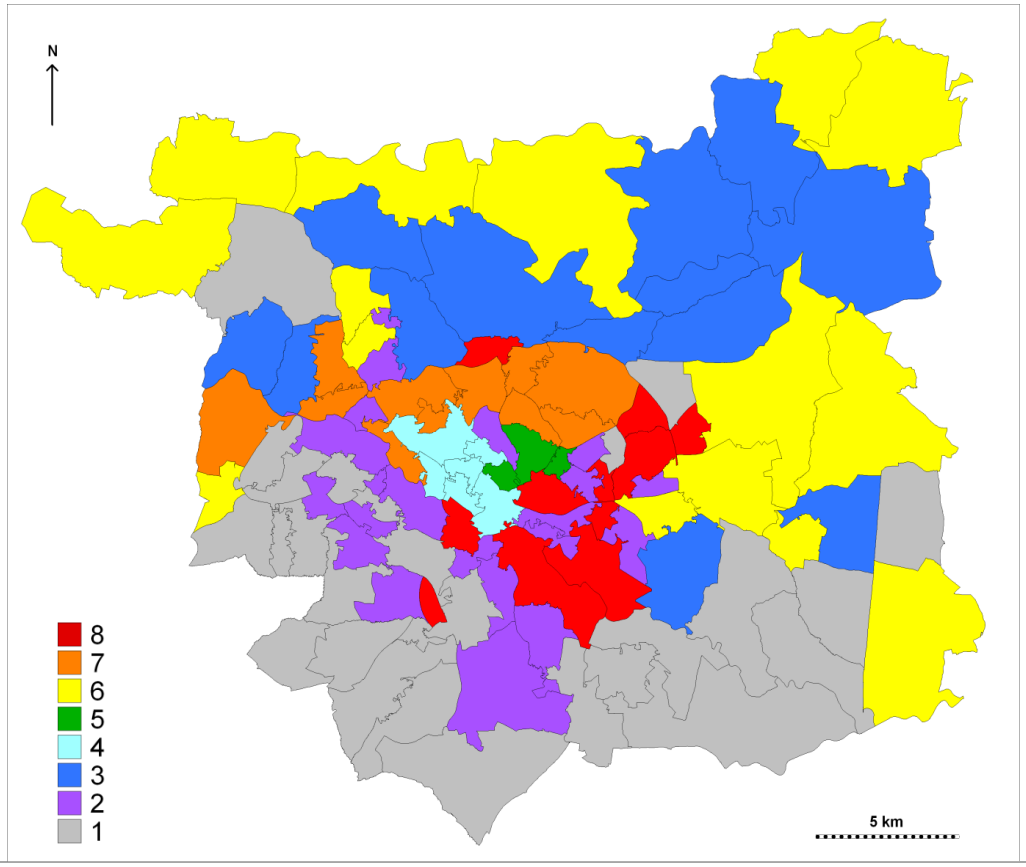
Figure 6.3. Key variables for eight cluster solution in Leeds (complex analysis), 2001

Statistics in Table 6.1 show that the Leeds population was unevenly classified into eight clusters. The cluster size varies from 3% to 30% of Leeds population, and in terms of number of included Community Areas – from 3% to 27%.

Table 6.1. Selected statistics for eight clusters in Leeds (complex analysis), 2001

Cluster solution	No of CAs	Population	Households	% of city population	% aged 65 and more	% with manual occupation	%non White British	% with LLTI and Disabled
1	27	208,622	87,586	29.2	15.3	51.0	4.6	17.8
2	20	126,645	53,867	17.7	13.5	62.3	10.7	20.3
3	13	67,005	26,847	9.4	16.2	29.7	8.4	13.5
4	6	49,373	20,015	6.9	8.3	43.0	21.7	12.5
5	3	23,825	10,147	3.3	10.6	57.8	59.6	18.9
6	15	98,272	40,901	13.7	19.3	42.0	4.8	17.6
7	11	86,678	36,712	12.1	17.1	33.3	16.7	16.9
8	11	54,746	25,526	7.7	17.2	67.3	8.7	25.5
All clusters	106	715,402	301,601	100.0	15.5	47.5	10.8	18.0

Community Areas were mapped by cluster (Figure 6.4). The location of clusters enables us to better interpret the results of cluster analysis and capture the spatial patterning of clusters. Based on the distribution of key variables and a map analysis, the clusters have been named. However, because this was very simple analysis the cluster were given short descriptions only with very general names. The map is followed by cluster characteristics.



#	**Cluster description**
1	Different age groups, Different Social Classes, White British, Average Disability - Diverse Population
2	Working Class, Different age groups, with Lone Parents and Disabled Co-residents with Average share of non White British
3	Different age groups, Middle Class, White British with below Average Disability, Outside City Centre, good quality housing
4	Young Adult, Couples, with Middle Class and Multiethnic Co-Residents City Centre, substandard housing
5	Multi-ethnic, Working Class, Young and Middle aged, with Lone Parents and some Disabled Co-Residents Neighbours, substandard housing
6	Older, White British, Middle Class Outside City Centre, and good quality housing
7	Different age groups, Middle Class with some non White British Co-Residents Average quality of flats, north City Centre
8	Working Class, different age groups, with Lone Parents and Disabled Co-Residents City Centre, council housing

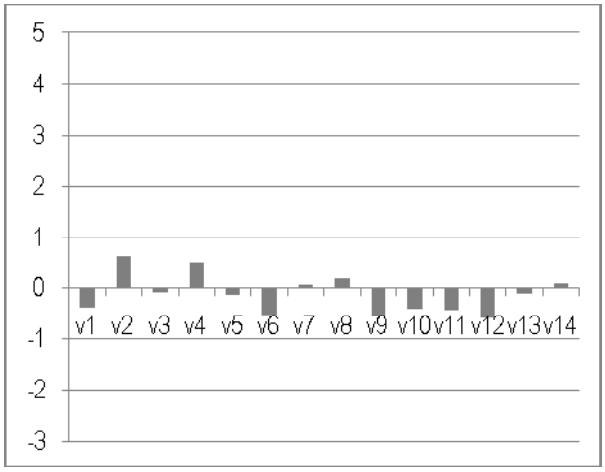
Figure 6.4. Cluster classification of Community Areas in Leeds (complex analysis)

Cluster 1: Different age groups, Different Social Classes, White British, Average Disability - Diverse Population

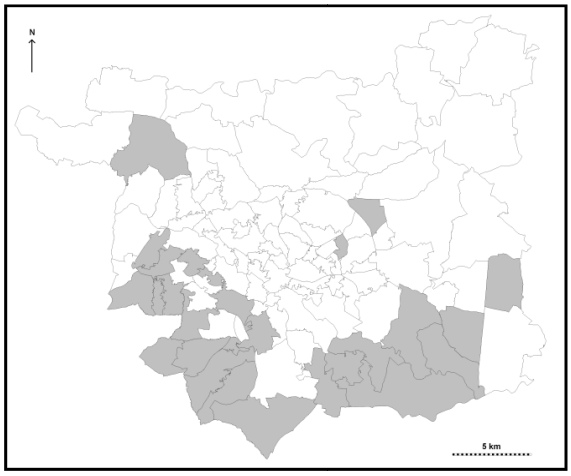
Number of CAs: 27

2001 Census Population: 208,622

Key variables for cluster 1 in Leeds (z-scores)



Map of Community Areas in cluster 1 in Leeds



Selected statistics for cluster 1 in Leeds

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	15.3	12.0	19.2	1.9
B	%manual	51.0	43.2	59.4	5.0
C	%nonWB	4.6	1.4	17.0	3.2
D	%disabled	17.8	15.4	22.4	1.8

Note: SD – social dimension; A – demographic social dimension; B – socio-economic social dimension; C – ethnic social dimension; D – disability social dimension.

List of Community Areas in cluster 1 in Leeds

2 Allerton Bywater &
Great and Little Preston
4 Ardsley East / West
(inc. Tingley)
9 Beeston
15 Bramley
22 Churwell
29 Drighlington
31 Fairbank
34 Farsley & Rodley
35 Fearnville
38 Gildersome
56 Kippax
59 Lofthouse & Robin
Hood
62 Methley
63 Micklefield
67 Morley North
68 Morley South
69 New Farnley
74 Oulton &
Woodlesford
76 Pudsey
77 Pudsey Lowtown
80 Rothwell
88 Stanningley
90 Swillington
93 Upper Armley
97 Whinmoor
98 Wortley
100 Yeadon

Figure 6.5. Profile for cluster 1 in Leeds, 2001

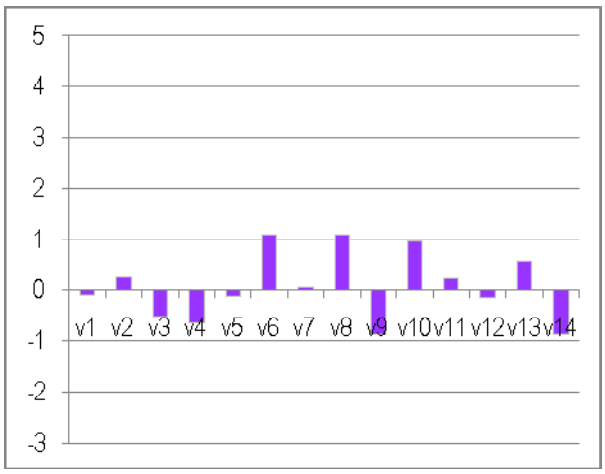
Cluster 1 has the biggest population among clusters: almost 30 per cent of the Leeds population lives in CAs classified as cluster 1. Population living in this cluster are middle aged; families with children are prominent, but people aged 15-24 are underrepresented: and the cluster has an average share of older population. The cluster includes Community Areas with large White British majorities and consequently the location quotient for the non-White British is the lowest among clusters. Regarding socio-economic traits, people with manual occupations are slightly overrepresented and the highly qualified are underrepresented. Areas classified into this cluster have good quality housing and are mainly located in the suburban areas of the south of Leeds.

Cluster 2: Working Class, Different age groups, with Lone Parents and Disabled Co-Residents with Average share of non White British

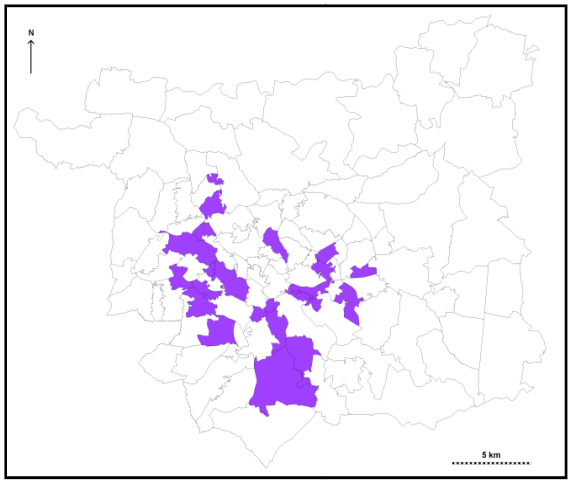
Number of CAs: 20

2001 Census Population: 126,645

Key variables for cluster 2 in Leeds (z-scores)



Map of Community Areas in cluster 2 in Leeds



Selected statistics for cluster 2 in Leeds

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	13.5	9.4	19.2	2.5
B	%manual	62.3	44.5	69.9	5.8
C	%nonWB	10.7	3.6	28.4	6.7
D	%disabled	20.3	16.9	23.9	2.0

Note: SD – social dimension; A – demographic social dimension; B – socio-economic social dimension; C – ethnic social dimension; D – disability social dimension.

Community Areas in cluster 2 in Leeds

5 Armley
10 Beeston Hill
11 Belle Isle
28 Crossgates
30 East Bank
33 Farnley
39 Gipton North
43 Halton Moor
44 Harehills
47 Hawksworth
49 Holbeck
50 Holt Park
64 Middleton
79 Richmond Hill
82 Sandfords, Ganners & Moorside
84 Scott Hall & Miles Hill
91 Swinnow & Fairfields
94 Upper Wortley
99 Wythers
106 Ireland Wood

Figure 6.6. Profile for cluster 2 in Leeds, 2001

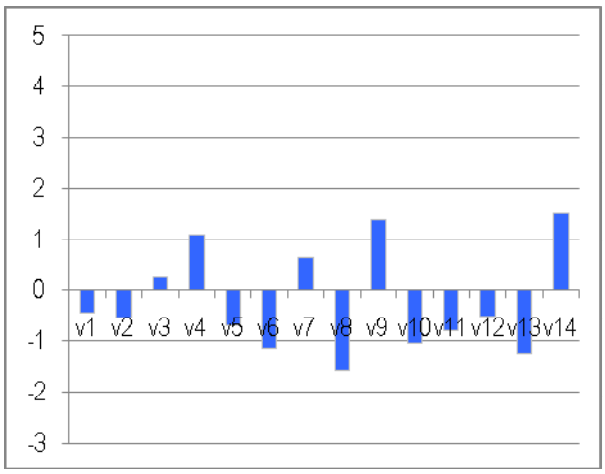
Cluster 2 is inhabited by people in different age groups. A considerably higher share of population working in manual occupations than on average in the city lives in Cluster 2. The location quotient for council housing is the highest of all clusters. With respect to family status, lone parent families with children are overrepresented while married people are underrepresented. Although the percentage of people aged 65 and over is lower than average, the percentage of people with LLTI or disabled is relatively high – over 20 per cent. The share of non White British population is close to the city average.

Cluster 3: Different age groups, Middle Class, White British with below Average Disability, Outside City Centre, good quality housing

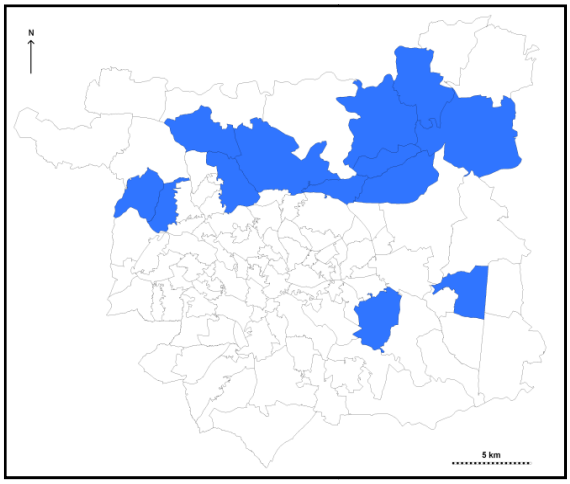
Number of CAs: 13

2001 Census Population: 67,005

Key variables for cluster 3 in Leeds (z-scores)



Map of Community Areas in cluster 3 in Leeds



Selected statistics for cluster 3 in Leeds

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	16.2	8.4	21.4	3.5
B	%manual	29.7	22.3	44.3	6.3
C	%nonWB	8.4	2.4	15.5	4.4
D	%disabled	13.5	10.1	15.4	1.5

Note: SD – social dimension; A – demographic social dimension; B – socio-economic social dimension; C – ethnic social dimension; D – disability social dimension.

Community Areas in cluster 3 in Leeds

1 Adel
3 Alwoodley & Wigton Moor
7 Bardsey & East Keswick
13Bramham
14 Bramhope
24 Collingham & Linton
25 Colton
36 Garforth East
53 Horsforth West End
78 Rawdon
83 Thorner
87 Shadwell
103 Scarcroft

Figure 6.7. Profile for cluster 3 in Leeds, 2001

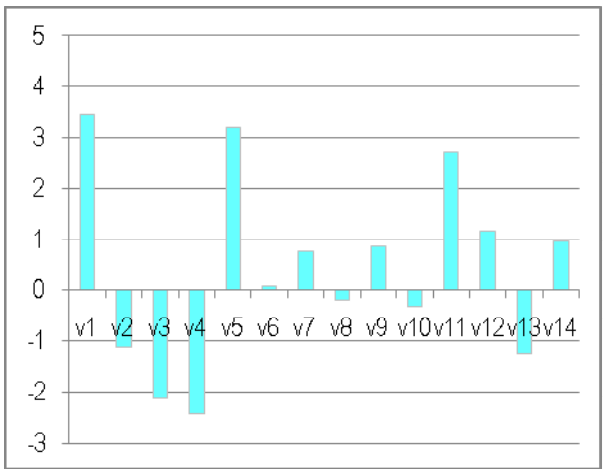
Different age groups live in cluster 3. However, the share of married people is higher, while couples without children and lone parents are underrepresented. Community Areas in cluster 3 are inhabited by White British people. The percentage of population with manual occupations is the lowest among clusters and share of people with highest qualification one of the highest, thus we can say that these are better off, middle class areas. Quality of housing is good and location quotient of council housing is the lowest among clusters (only 0.2).

Cluster 4: Young Adult, Couples, with Middle Class and Multiethnic Co-Residents City Centre, substandard housing

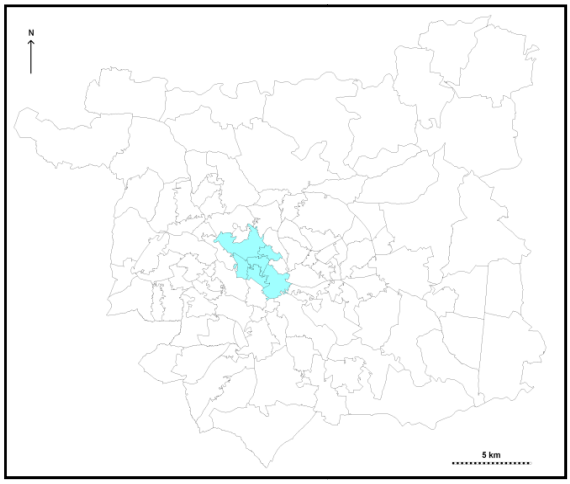
Number of CAs: 6

2001 Census Population: 49,373

Key variables for cluster 4 in Leeds (z-scores)



Map of Community Areas in cluster 4 in Leeds



Selected statistics for cluster 4 in Leeds

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	8.3	3.9	10.6	2.5
B	%manual	43.0	39.2	49.7	4.3
C	%nonWB	21.7	14.7	38.9	9.0
D	%disabled	12.5	8.9	15.6	2.4

Note: SD – social dimension; A – demographic social dimension; B – socio-economic social dimension; C – ethnic social dimension; D – disability social dimension.

Community Areas in cluster 4 in Leeds

16 Burley
17 Burley Lodge & Little Woodhouse
23 City Centre
48 Headingley
55 Hyde Park
58 Woodhouse

Figure 6.8. Profile for cluster 4 in Leeds, 2001

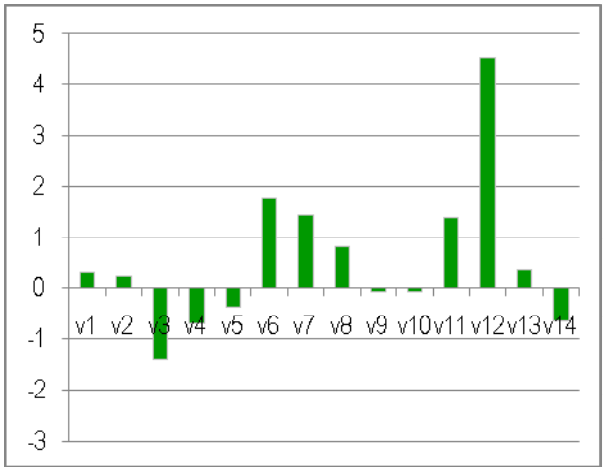
Community Areas ascribed to cluster 4 have the youngest population in Leeds and 50 per cent of residents in the area are aged 15-24. Other age groups, especially people aged 65 and more are underrepresented. Regarding family status, single people and cohabitating couples are in the majority, but the share of married people is very low. Ethnic minorities live there more frequently than in other parts of the city. People with non-manual occupations are over-represented. Although there are not so many council flats, every fourth flat/house has been classified as overcrowded.

Cluster 5: Multi-ethnic, Working Class, Young and Middle aged, with Lone Parents and some Disabled Co-Residents Neighbours, substandard housing

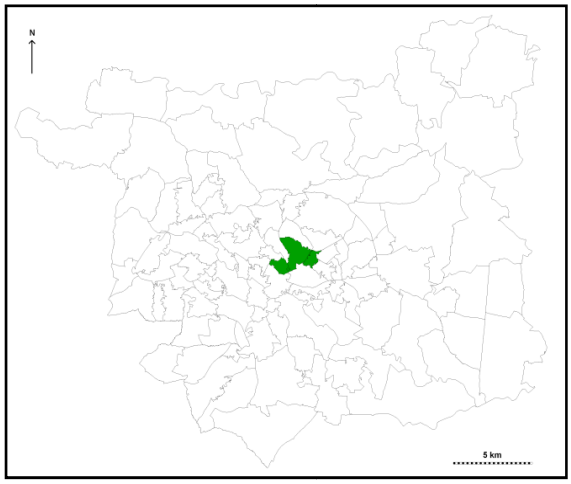
Number of CAs: 3

2001 Census Population: 23,825

Key variables for cluster 5 in Leeds (z-scores)



Map of Community Areas in cluster 5 in Leeds



Selected statistics for cluster 5 in Leeds

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	10.6	8.9	12.0	1.6
B	%manual	57.8	54.5	61.8	3.8
C	%nonWB	59.6	38.8	68.6	15.4
D	%disabled	18.9	17.1	19.9	1.6

Note: SD – social dimension; A – demographic social dimension; B – socio-economic social dimension; C – ethnic social dimension; D – disability social dimension.

List of Community Areas in cluster 5 in Leeds

21 Chapeltown
45 Harehills Triangle
105 Little London

Figure 6.9. Profile for cluster 5 in Leeds, 2001

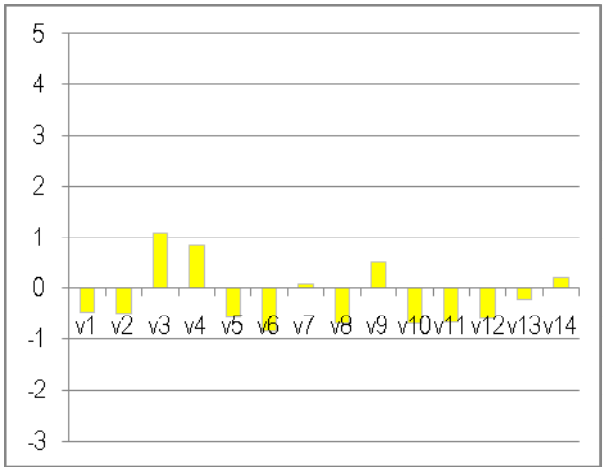
Cluster 5 covers the smallest population of Leeds (3.3 per cent) and only three neighbouring Community Areas, located to the north of the city centre belong to this cluster. It is another young cluster with higher share of people aged15-24, but also people in a middle age are overrepresented. Ethnic diversity is high and percentage of non-White British people is the highest among eight clusters (5.5 location quotient). The share of lone parents is the highest among clusters. The Community Areas in the cluster are more frequently inhabited by working class. In comparison to other clusters the residents of cluster 5 are more often people with LLTI or disabled.

Cluster 6: Older, White British, Middle Class Outside City Centre, and good quality housing

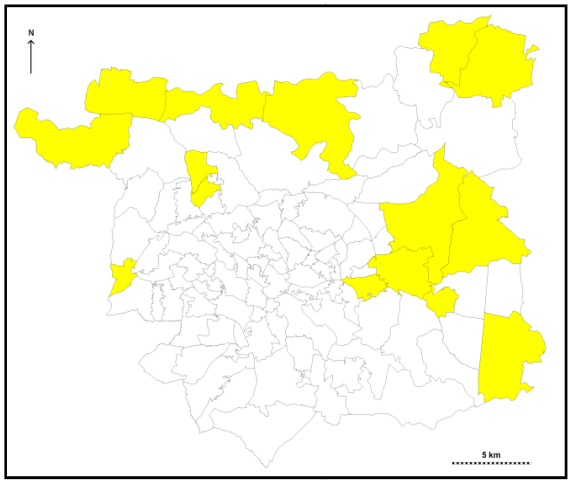
Number of CAs: 15

2001 Census Population: 98,272

Key variables for cluster 6 in Leeds (z-scores)



Map of CAs in cluster 6 in Leeds



Selected statistics for cluster 6 in Leeds

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	19.3	16.1	22.7	1.8
B	%manual	42.0	32.5	47.8	5.1
C	%nonWB	4.8	2.0	26.8	6.1
D	%disabled	17.6	14.8	22.0	2.2

Note: SD – social dimension; A – demographic social dimension; B – socio-economic social dimension; C – ethnic social dimension; D – disability social dimension.

Community Areas in cluster 6 in Leeds

6 Arthington & Pool
8 Barwick & Scholes
12 Boston Spa
26 Cookridge
37 Garforth West
41 Guiseley
42 Halton/Whitkirk
46 Harewood and District
60 Manston,
73 Otley
75 Priesthorpe
92 Tinshill
96 Wetherby
10 Ledston & Ledsham
102 Aberford

Figure 6.10. Profile for cluster 6 in Leeds

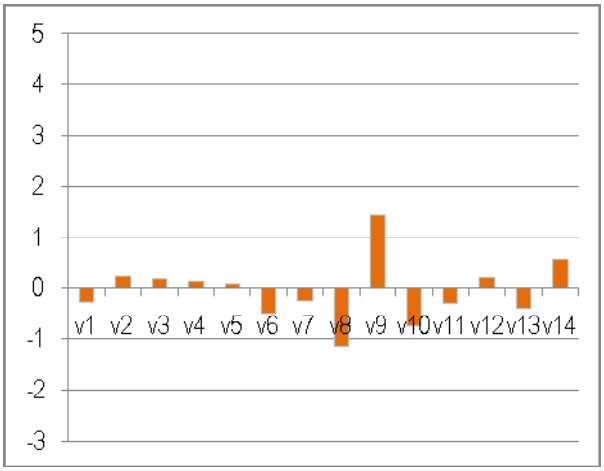
Cluster 6 covers Community Areas with high percentages of older and married population – younger and middle aged residents are underrepresented, as well as other types of families, e.g. couples without children or lone parents. The cluster has low share of ethnic minorities with very low location quotient of non White British population (0.4). Socioeconomic characteristics are on average typical, but there is quite considerable variation in non-manual occupation percentage from 32% to 48%. These are communities with good quality housing located outside city centre.

Cluster 7: Different age groups, Middle Class with some non White British Co-Residents Average quality of flats, north City Centre

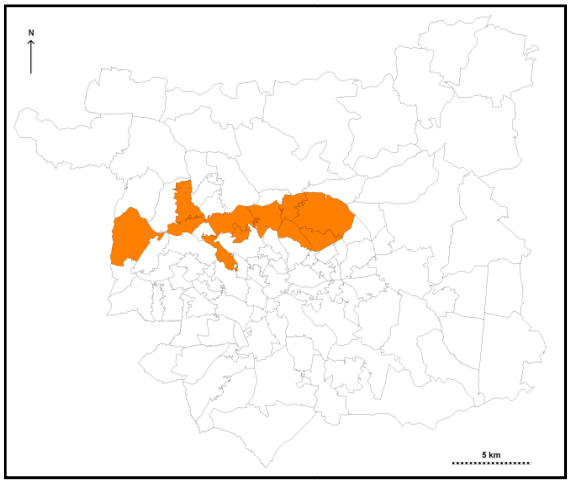
Number of CAs: 11

2001 Census Population: 86,678

Key variables for cluster 7 in Leeds (z-scores)



Map of CAs in cluster 7 in Leeds



Selected statistics for cluster 7 in Leeds

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	17.1	13.6	20.8	2.2
B	%manual	33.3	21.7	41.2	5.8
C	%nonWB	16.7	4.5	31.9	9.2
D	%disabled	16.9	13.2	18.8	1.6

Note: SD – social dimension; A – demographic social dimension; B – socio-economic social dimension; C – ethnic social dimension; D – disability social dimension.

List of Community Areas in cluster 7 in Leeds

19 Calverley
20 Chapel Allerton
32 Far Headingley
51 Horsforth
52 Horsforth Newlaithes & Woodside
57 Kirkstall
61 Meanwood
66 Moortown
71 Oakwood
81 Roundhay
95 West Park

Figure 6.11. Profile for cluster 7 in Leeds, 2001

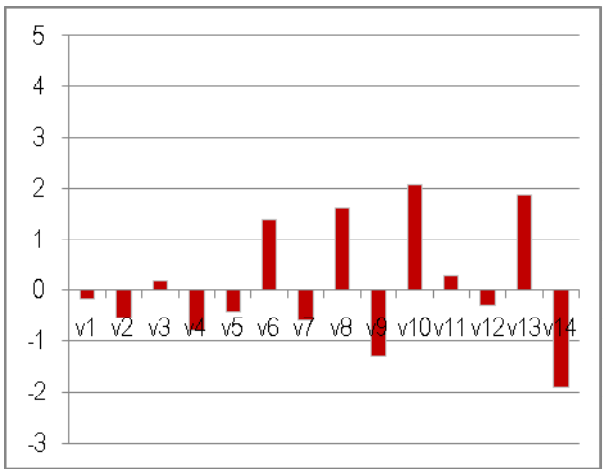
Cluster 7 is inhabited by residents of different age groups and representing different family types. The share of ethnic minorities is high. The areas have one of the lowest shares of manual workers and the highest share of people with the highest qualifications. Percentage of disabled population is slightly below city average. Cluster 7 covers some Community Areas with good quality housing, located in the ring outside city centre and some more suburban areas.

Cluster 8: Working Class, different age groups, with Lone Parents and Disabled Co-Residents City Centre, council housing

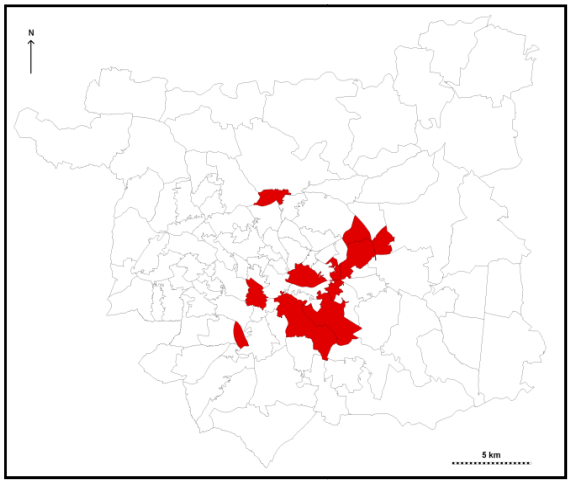
Number of CAs: 11

2001 Census Population: 54,746

Key variables for cluster 8 in Leeds (z-scores)



Map of CAs in cluster 8 in Leeds



Selected statistics for cluster 8 in Leeds

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	17.2	12.6	23.1	3.2
B	%manual	67.3	56.4	72.7	5.6
C	%nonWB	8.7	3.0	19.9	5.1
D	%disabled	25.5	22.6	30.4	2.0

Note: SD – social dimension; A – demographic social dimension; B – socio-economic social dimension; C – ethnic social dimension; D – disability social dimension.

Community Areas in cluster 8 in Leeds

18 Burmantofts, Lincoln Green & Ebor Gardens
27 Cottingley
40 Gipton South
54 Hunslet / Stourton
65 Moor Allerton
70 New Wortley
72 Osmondthorpe
85 Seacroft North
86 Seacroft South
89 Swarcliffe
104 Cross Green

Figure 6.12. Profile for cluster 8 in Leeds, 2001

Cluster 8 contains Community Areas with the highest percentage of population with manual occupations – 67 percent, 20 percent higher than the city average. These Community Areas are inhabited by different age groups with some overrepresentation of older population. The cluster has the higher share of population with LLTI or disabled among eight clusters (every fourth resident). Council flats are considerably overrepresented in the areas comprising 55 percent of flats/houses. Although the CAs are located in the proximity of the city centre, the share of ethnic minorities is lower than in the whole of Leeds.

6.2. Cluster profiles Warsaw

The same clustering procedure was repeated for Warsaw at the level of Urban Regions. The major difference was that the disability variable was not available, so that the simple cluster analysis was based on three variables: percentage of population aged 65 and more, percentage of manual workers and percentage of foreign immigrant population. It should be borne in mind that dataset for Warsaw covered different type of population – not the resident population, but the population de facto living in the city (for definitions see section 3.1.2).

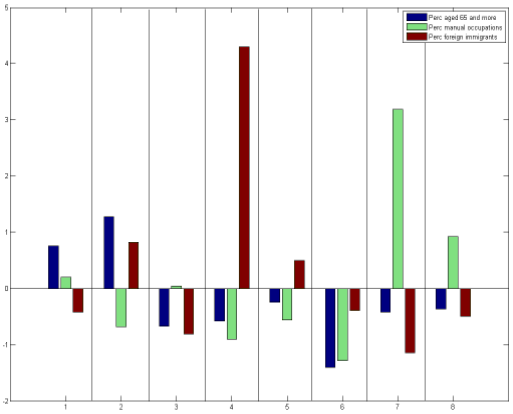


Figure 6.13. Key variables for eight cluster solution and Urban Regions classification for Warsaw (z-scores, simple analysis)

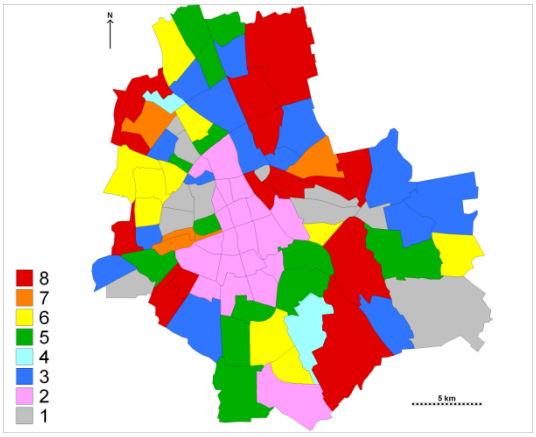
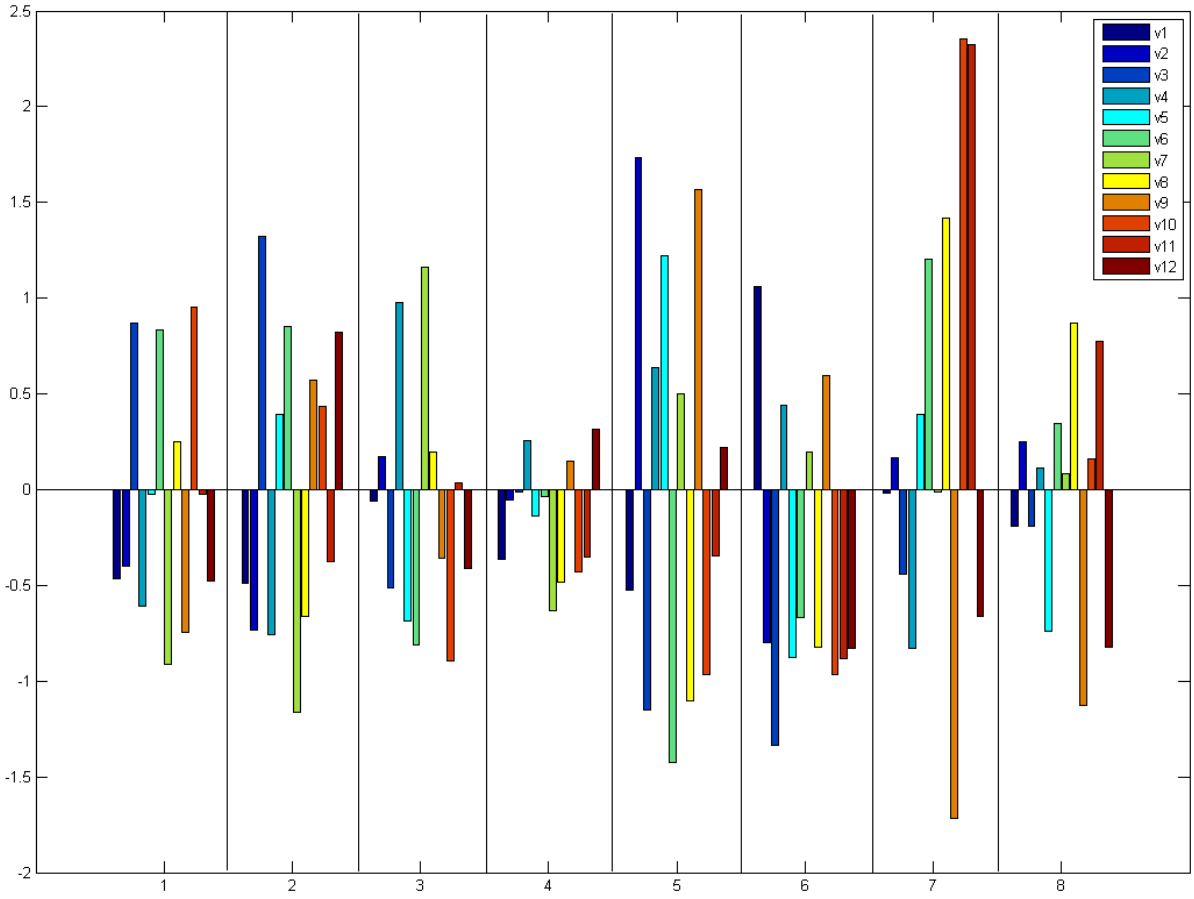


Figure 6.14. A map of the classification of Urban Regions in Leeds into 8 clusters, simple analysis, 2002

Figure 6.13 profiles the clusters in terms of the three variables, while Figure 6.14 maps all the clusters. The simple analysis results were satisfactory, so the same methodology was applied for more complex cluster analysis. In case of Warsaw the complex analysis was based on 12 variables – without the percentages of disabled people and people in good health. Again, the eight cluster solution proved to be the most coherent. Key variables for eight cluster solution are presented in the Figure 6.15 and in Table 6.2.

Like Leeds, the Warsaw population was unevenly classified in terms of population and number of Urban Regions into eight clusters. The cluster size varies from 3 percent to 30 percent of Warsaw population too, but the range of included Urban Regions was not as wide as for Leeds – from 7 to 19. Urban Regions were mapped by cluster. Figure 6.16 presents the 8 cluster solution from the complex analysis. Then follow short descriptions of clusters and interpretation of the results.



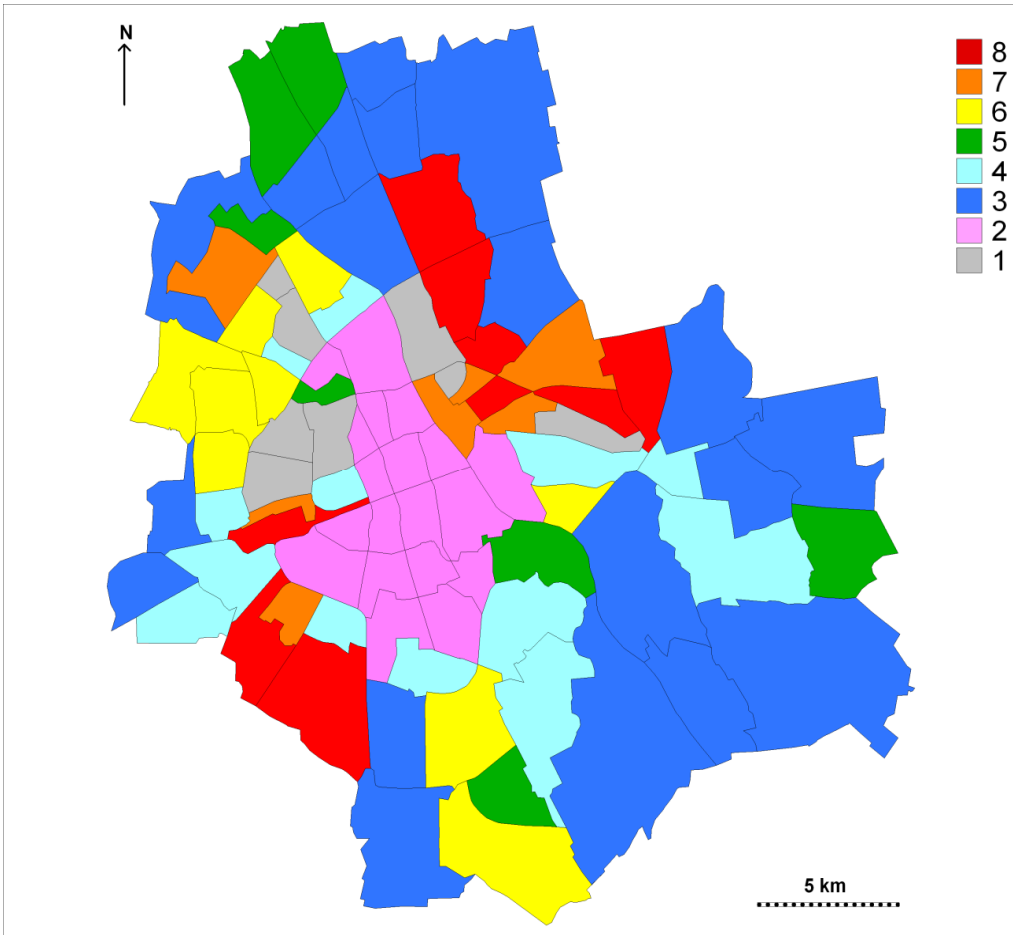
Legend:

v1 - Percentage of people aged 15-24
v2 - Percentage of people aged 25-44
v3 - Percentage of people aged 65 and more
v4 - Percentage of married people
v5 - Percentage of couples without children
v6 - Percentage of lone parents
v7 - Percentage of 4 and more persons households
v8 - Percentage of people with manual occupations
v9 - Percentage of people with higher education
v10 - Percentage of household rented/owned by council
v11 - Percentage of substandard housing
v12 - Percentage of foreign immigrants

Figure 6.15. Key variables for eight cluster solution in Warsaw (complex analysis)

Table 6.2. Selected statistics for eight clusters in Warsaw (complex analysis)

Cluster solution	No of URs	Population	Households	% of city population	% aged 65 and more	% manual occupation*	% foreign immigrants
1	8	202,620	96,056	12.0	22.0	40.2	0.4
2	19	503,204	254,561	29.8	24.5	27.8	0.9
3	19	150,756	54,303	8.9	10.7	38.6	0.4
4	13	321,769	143,578	19.1	15.9	29.9	0.7
5	7	77,979	30,600	4.6	4.2	18.5	0.7
6	9	239,759	93,296	14.2	6.9	27.5	0.4
7	7	47,928	20,854	2.8	12.5	50.5	0.4
8	9	144,287	61,802	8.5	12.5	45.8	0.2
All clusters	**91**	**1,688,302**	**755,050**	**100.0**	**16.5**	**32.0**	**0.6**



	Legend
1	Working class, Polish, Older population and Lone parents, Inferior areas, Substandard housing
2	Older, Middle Class, Couples and Lone Parents, with Foreign Immigrant population, City Centre Regions, with council housing
3	Big Families and Different age groups, with prevalence of Working Class, Polish, outside City Centre, average standard (big UR)
4	Middle Class, Different age groups with Foreign Immigrant population, scattered outside City Centre, good quality housing
5	Middle aged, higher educated, couples and married/families, with Some Foreigners, outside City Centre, good quality housing
6	Young, Middle Class, Married/Families, Polish, new estates outside City Centre, good quality housing
7	Hyper-Working Class, Polish, Middle aged, with Lone Parents, disadvantaged flats (high share of council and substandard housing)
8	Polish, Working Class, different age groups, and with above share of lone parents, Outside City Centre, substandard housing

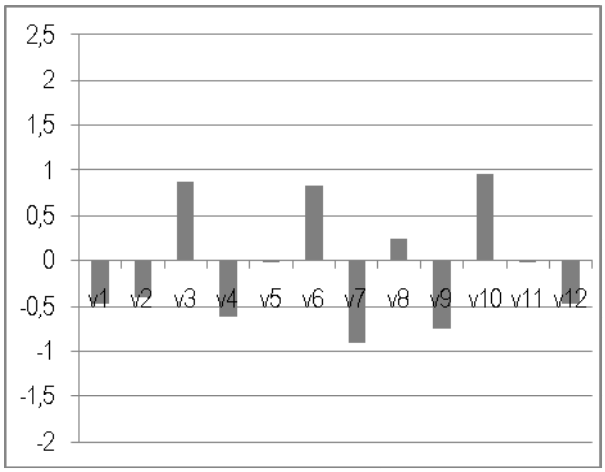
Figure 6.16. Cluster classification of Community Areas in Warsaw (complex analysis)

Cluster 1: Working class, Polish, Older population and Lone parents, Inferior areas, Substandard housing

Number of URs: 8

2002 Census Population: 202,620

Key variables for cluster 1 in Warsaw (z-scores)



Map of URs in cluster 1 in Warsaw

Selected statistics for cluster 1 in Warsaw

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	22.0	16.5	26.2	2.9
B	%manual	40.2	35.1	49.0	4.6
C	%foreign	0.4	0.1	0.6	0.1

Note: SD – social dimension; A – demographic social dimension; B – socio-economic social dimension; C – ethnic social dimension.

Urban Regions in cluster 1 in Warsaw

5 Bielany
7 Brzeziny
17 FSO
20 Grochów-Północ
28 Koło-Północ
29 Koło-Poludnie
38 Młynów
61 Praga II i III

Figure 6.17. Profile for cluster 1 in Warsaw, 2002

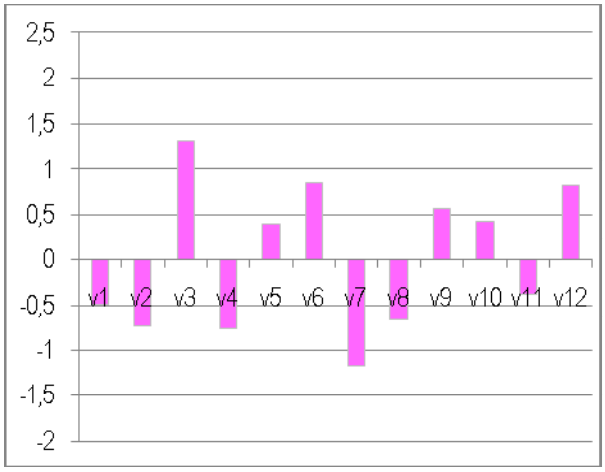
Older people are overrepresented in the cluster 1, while younger and middle aged residents constitute similar share to the city mean. Regarding family status lone parents with children live there more frequently than general in the city. The share of foreigners is lower than the city average. The areas are more often inhabited by working class and also are perceived by Warsaw residents as inferior, comprising high share of council flats.

Cluster 2: Older, Middle Class, Couples and Lone Parents, with Foreign Immigrant population, City Centre Regions, with council housing

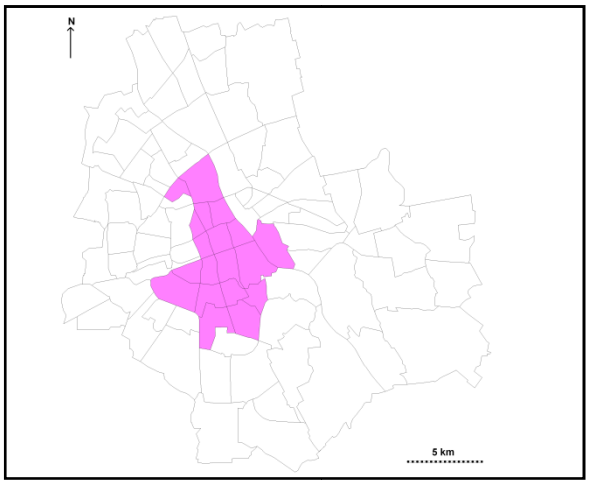
Number of URs: 19

2002 Census Population: 503,204

Key variables for cluster 2 in Warsaw (z-scores)



Map of URs in cluster 2 in Warsaw



Selected statistics for cluster 2 in Warsaw

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	24.5	22.2	29.3	2.1
B	%manual	27.8	22.5	33.3	2.9
C	%foreign	0.9	0.3	1.6	0.3

Note: SD – social dimension; A – demographic social dimension; B – socioeconomic social dimension; C – ethnic social dimension.

Urban Regions in cluster 2 in Warsaw

8 Centrum Północ
9 Centrum Południe
35 Mirów (Śródm. Zach.)
39 Mokotów Centrum
40 Mokotów Stary
41 Mokotów Wschód
42 Muranów Wschodni
43 Muranów Zachodni
47 Ochota Centrum
56 Pole Mokotowskie
58 Powiśle Północ
59 Powiśle Południe
65 Saska Kępa
67 Sielce
70 Stare i Nowe Miasto
72 Szosa Krakowska
83 Wierzbno
89 Żoliborz Centralny
91 Żoliborz Zachodni

Figure 6.18. Profile for cluster 2 in Warsaw, 2002

Urban Regions in Cluster 2 are located in the city centre and cover the biggest population, almost 30 percent of Warsaw residents. They are inhabited by older population – every fourth resident is at least 65 years old. However, married people are underrepresented in contrast to couples and lone parents. The middle class population lives in this cluster, as the share of people with manual occupations is relatively low. The city centre Urban Regions have the highest share of foreign immigrants (location quotient 1.5).

Cluster 3: Big Families and Different age groups, with prevalence of Working Class, Polish, outside City Centre, average standard (big UR)

Number of URs: 19

2002 Census Population: 150,756

Key variables for cluster 3 in Warsaw (zscores)

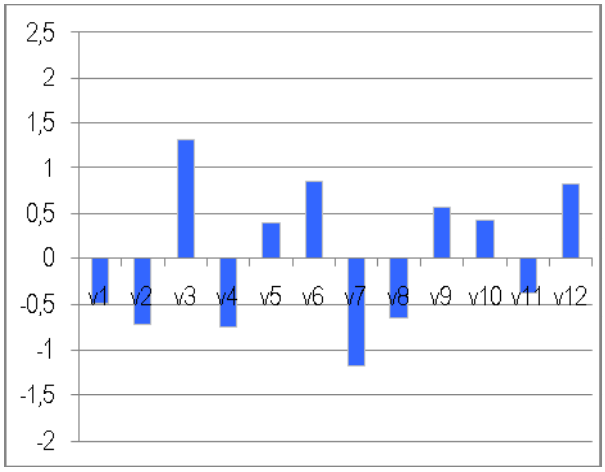
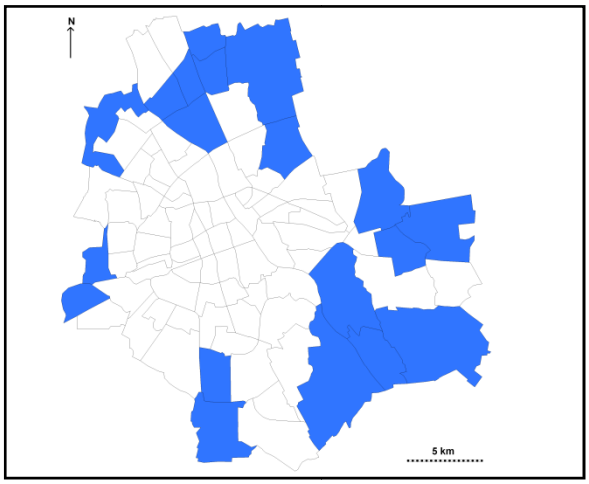


Figure 6.28. URs in cluster 3 in Warsaw



Selected statistics for cluster 3 in Warsaw

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	10.7	4.6	15.6	3.1
B	%manual	38.6	28.6	59.2	7.3
C	%foreign	0.4	0.2	1.1	0.3

Note: SD – social dimension; A – demographic social dimension; B – socioeconomic social dimension; C – ethnic social dimension.

Urban Regions in cluster 3 in Warsaw

4 Białołęka Dworska,
10 Choszczówka,
11 Chrzanów,
16 Falenica,
18 Gocław,
22 Henryków,
25 Julianów,
33 MańkiBrzeziny,
37 Młociny Las,
53 Park Sobieskiego,
60 Powsin,
63 Pyry,
64 Rembertów,
69 Służewiec Południowy,
73 Tarchomin,
76 Ursus Gołąbki,
81 Wesoła Północ,
86 Zacisze,
88 Żerań Zachodni

Figure 6.19. Profile for cluster 3 in Warsaw, 2002

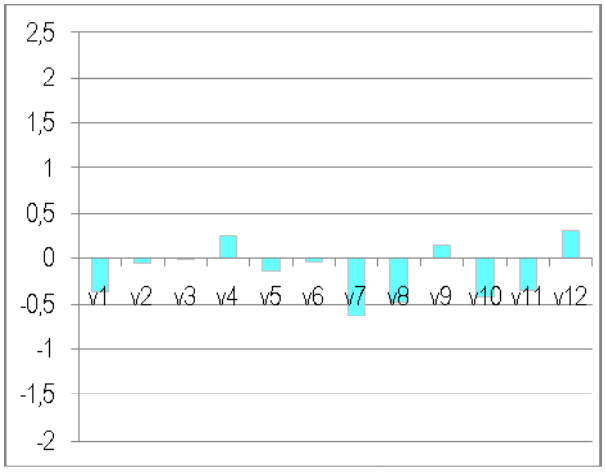
Cluster 3 includes Urban Regions with different age groups, but much lower share of older population. Married people with children dominate among family types. Working class people prevail over middle class people. Although share of council housing is low, almost every tenth flat/house was classified as substandard. These are big Urban Regions located outside city centre and spread in all directions.

Cluster 4: Middle Class, Different age groups with Foreign Immigrant population, scattered outside City Centre, good quality housing

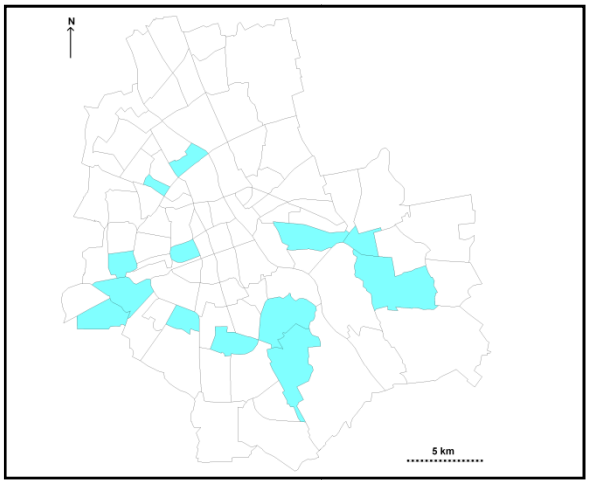
Number of URs: 13

2002 Census Population: 321,769

Key variables for cluster 4 in Warsaw (zscores)



Map of URs in cluster 4 in Warsaw



Selected statistics for cluster 4 in Warsaw

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	15.9	11.8	21.3	2.7
B	%manual	29.9	21.6	37.7	4.5
C	%foreign	0.7	0.2	2.7	0.6

Note: SD – social dimension; A – demographic social dimension; B – socioeconomic social dimension; C – ethnic social dimension.

Urban Regions in cluster 4 in Warsaw

12 Czerniaków
13 Czerniaków Wilanów
14 Czyste
21 Grochów Południe
24 Jelonki
34 Marymont
49 Okęcie Nowe
54 Piaski
68 Służew
77 Ursus Skorosze
79 Wawer
84 Włochy
85 Wygoda

Figure 6.20. Profile of cluster 4 in Warsaw, 2002

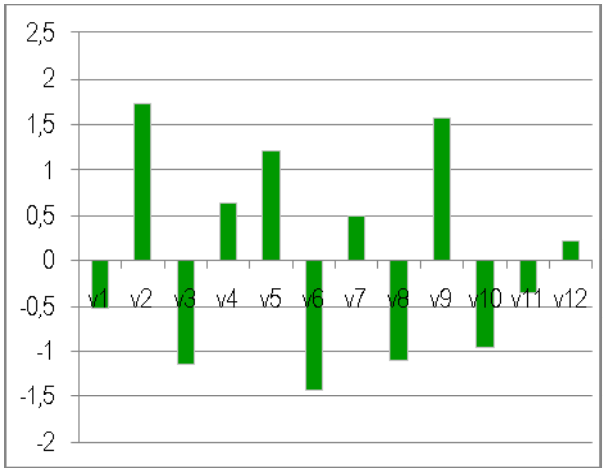
Cluster 4 is inhabited by almost 20 percent of Warsaw residents. People of different age groups live in these Urban Regions and also family statuses are equally represented. The share of foreigners is relatively high. The areas are inhabited by lower percentage of population with manual occupations than the city average, but the share of highly educated people is similar to the city mean. The quality of housing is good.

Cluster 5: Middle aged, higher educated, couples and married/families, with some foreigners, outside City Centre, good quality housing

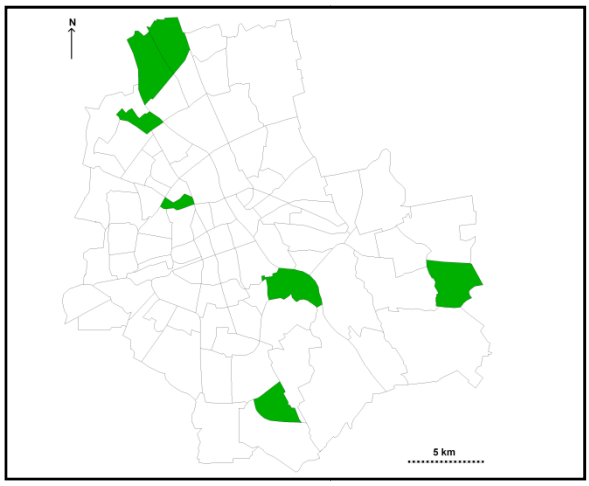
Number of URs: **7**

2002 Census Population: **77,979**

Key variables for cluster 5 in Warsaw (zscores)



Map of URs in cluster 5 in Warsaw



Selected statistics for cluster 5 in Warsaw

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	4.2	2.9	8.9	2.4
B	%manual	18.5	16.1	38.5	8.5
C	%foreign	0.7	0.5	2.3	0.8

Note: SD – social dimension; A – demographic social dimension; B – socioeconomic social dimension; C – ethnic social dimension.

Urban Regions in cluster 5 in Warsaw

15 Dąbrówka,
36 Młociny,
44 Natolin,
46 Nowodwory,
66 Siekierki,
82 Wesoła Południe,
90 Żoliborz Przemysłowy

Figure 6.21. Profile for cluster 5 in Warsaw, 2002

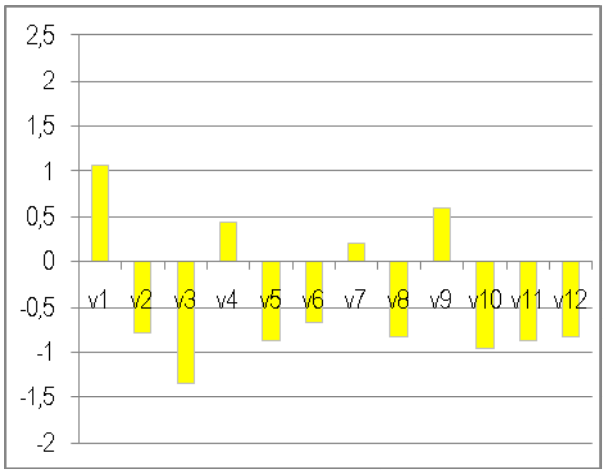
Urban Regions assigned to Cluster 5 are inhabited by a middle aged population and married people with children dominate. However, couples without children are overrepresented and the location quotient is the highest among clusters (1.8). People aged 65 and more are underrepresented. People with manual occupations are in the minority and the regions could be described as middle class areas, as the highest share of highly educated people lives there among eight clusters. These are Urban Regions with relatively new housing estates, and consequently the best housing standards among clusters with the lowest share of council flats/houses.

Cluster 6: Young, Middle Class, Married/Families, Polish, new estates outside City Centre, good quality housing

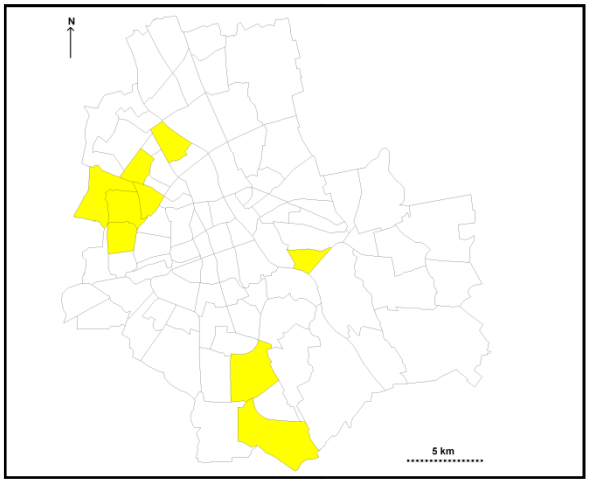
Number of URs: 9

2002 Census Population: 239,759

Key variables for cluster 6 in Warsaw (zscores)



URs in cluster 6 in Warsaw



Selected statistics for cluster 6 in Warsaw

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	6.9	2.4	15.2	3.5
B	%manual	27.5	14.9	31.8	5.7
C	%foreign	0.4	0.0	1.0	0.3

Note: SD – social dimension; A – demographic social dimension; B – socioeconomic social dimension; C – ethnic social dimension.

Urban Regions in cluster 6 in Warsaw

1 Babice
2 Bemowo Wschód
3 Bemowo Zachód
19 Górce
30 Las Bielański
31 Las Kabacki
32 Lotnisko
78 Ursynów
80 Wawrzyszew

Figure 6.22. Profile for cluster 6 in Warsaw, 2002

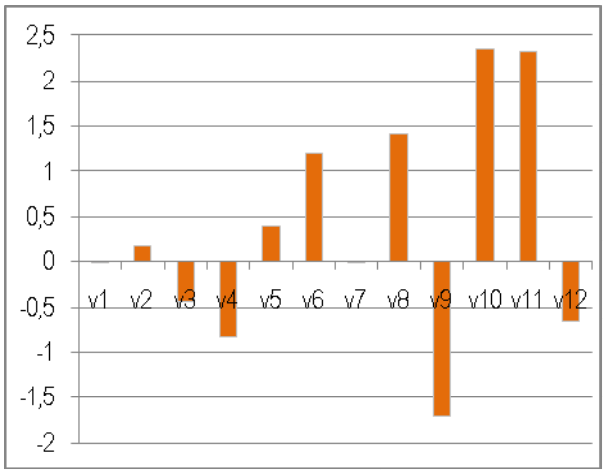
Cluster 6 is a young people cluster, the location quotient of people aged 15-24 is the highest among clusters (1.3). These Urban Regions are located in the suburban areas with relatively new estates, which explains why the share of older people is so low (only 7 percent, city average 16 percent). Cluster 6 population is better educated and more often possesses non-manual occupations. Foreign inhabitants do not live in these Urban Regions frequently. Housing conditions are good in the cluster.

Cluster 7: Hyper-Working Class, Polish, Middle aged, with Lone Parents, disadvantaged flats (high share of council and substandard flats, Post-manufacturing regions outside City Centre

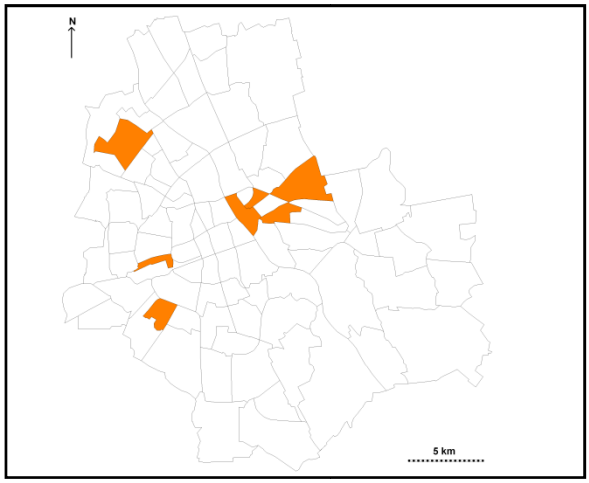
Number of URs: 7

2002 Census Population: 47,929

Key variables for cluster 7 in Warsaw (zscores)



Maps of URs in cluster 7 in Warsaw



Selected statistics for cluster 7 in Warsaw

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	12.5	11.9	14.8	1.0
B	%manual	50.5	42.3	76.7	13.2
C	%foreign	0.4	0.0	0.7	0.2

Note: SD – social dimension; A – demographic social dimension; B – socioeconomic social dimension; C – ethnic social dimension.

Urban Regions in cluster 7 in Warsaw

23 Huta Warszawa
26 Kamionek
45 Nowa Praga
48 Odolany
50 Okęcie Przemysłowe
62 Praga Centrum
75 Targówek Przemysłowy

Figure 6.23. Profile of cluster 7 in Warsaw, 2002

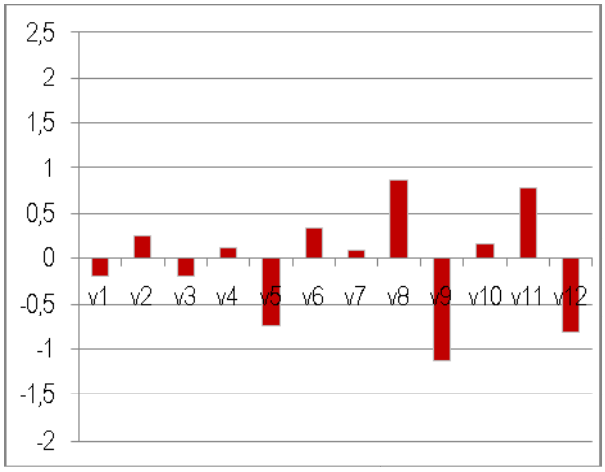
Cluster 7 covers the smallest population with only seven Urban Regions. Middle aged people are overrepresented. Regarding family status the share of married people is lower than city average, but share of lone parents is higher. The percentage of manual occupations reaches 51 percent and working class share is the highest among eight clusters. At the same time the cluster has one of the lowest numbers of foreign inhabitants. Over 55 percent of housing is owned by city council and 22 percent has substandard living conditions (location quotients 3.3).

Cluster 8: Polish, Working Class, different age groups, and with above share of lone parents, Outside City Centre, substandard housing

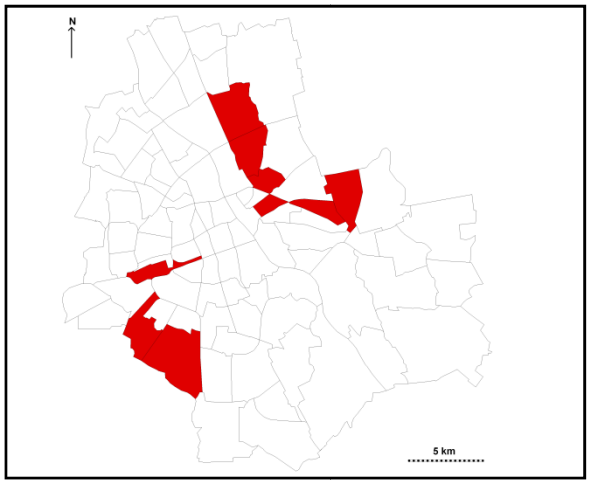
Number of URs: 9

2002 Census Population: 144,287

Key variables for cluster 8 in Warsaw (zscores)



Map of URs in cluster 8 in Warsaw



Selected statistics for cluster 8 in Warsaw

SD	Variable	Average	Min	Max	St.Dev.
A	%65+	12.5	9.5	15.6	1.7
B	%manual	45.8	38.6	74.5	10.5
C	%foreign	0.2	0.0	0.6	0.2

Note: SD – social dimension; A – demographic social dimension; B – socioeconomic social dimension; C – ethnic social dimension.

Urban Regions in cluster 8 in Warsaw

6 Bródno
27 Kawęczyn
51 Okęcie Lotnisko
52 Okęcie Opacz
55 PKP Wola
57 Postojowa
71 Szmulowizna
74 Targówek Mieszkaniowy
87 Żerań Wschodni

Figure 6.24. Profile of cluster 8 in Warsaw, 2002

A relatively high percentage of working population lives also in cluster 8, e.g. people with manual occupations (46 percent, city average 32 percent). Inhabitants comprise of people in different age groups, but middle aged are over- and older people under-represented. The Urban Regions do not have many foreign inhabitants. These Urban Regions are located outside the city centre, but in comparison to cluster 7 living conditions are much better and the share of council flats is close to city mean.

6.3. Synthesis

In order to present four dimensions of diversity in a more comprehensive way, four dimensional bubble graphs were prepared. The age dimension is represented on axis Y, the occupation dimension is represented on the X axis, bubble colour represents percentage of people with LTTI and disabled and bubble size represents percentage of non White British residents. The applet is one of Google's tool, but it was originally created by Swedish statistician and physician Hans Rosling[23]. The graphs are called Motion Charts, because you can apply data for different time series and then play the chart and watch the bubble changing location on X and Y axes, and changing colour and size[24].

In Figure 6.25 the bubble colour stands for cluster number instead of percentage of people with LLTI and disabled. In this graph it is visible how Community Areas that belong to the same cluster do cluster together, which means that they have similar age, occupation and ethnic characteristics. The classification of similar cases in the respect of the four selected variables was, of course, the aim of the cluster analysis.

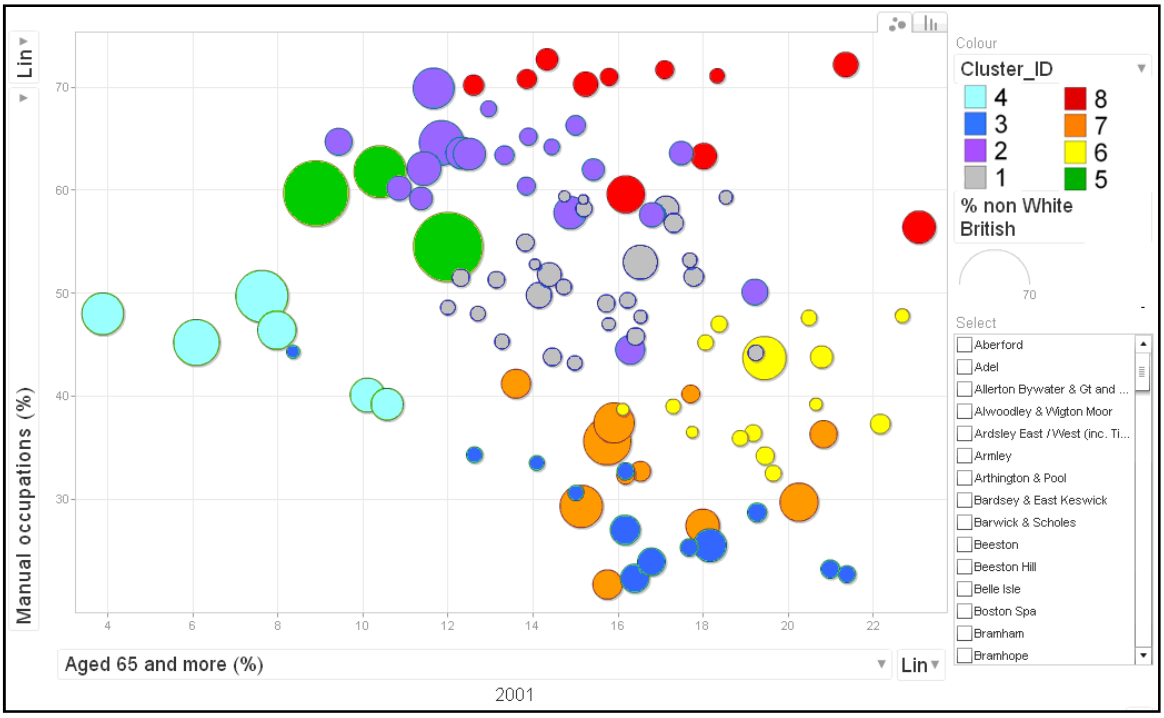


Figure 6.25. Community Areas in Leeds plotted on a graph showing the four social dimensions

The graph can be accessed online at https://spreadsheets0.google.com/ccc?pli=1&hl=en&key=tQ_SfAJ2dotTqJ2xBVO1l3Q&authkey=CJ-aq88N&hl=en#gid=5

[23] See: http://www.gapminder.org/.
[24] See: http://www.google.com/ig/directory?url=www.google.com/ig/modules/motionchart.xml.

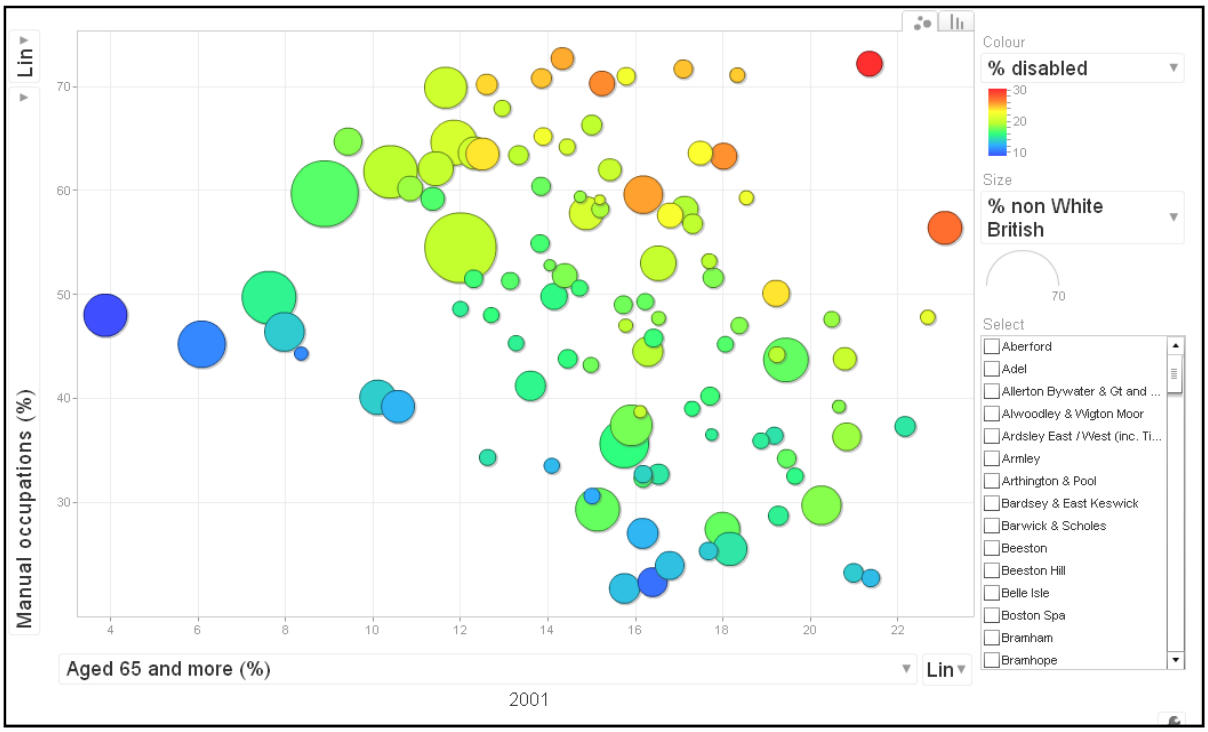


Figure 6.26. Community Areas in Leeds plotted on a graph showing the three social dimensions

The graph can be accessed online at
https://spreadsheets.google.com/ccc?key=0AqFpSb0sQUl4dFFfU2ZBSjJkb3RUcUoyeEJWTzFsM1E&hl=en#gid=3

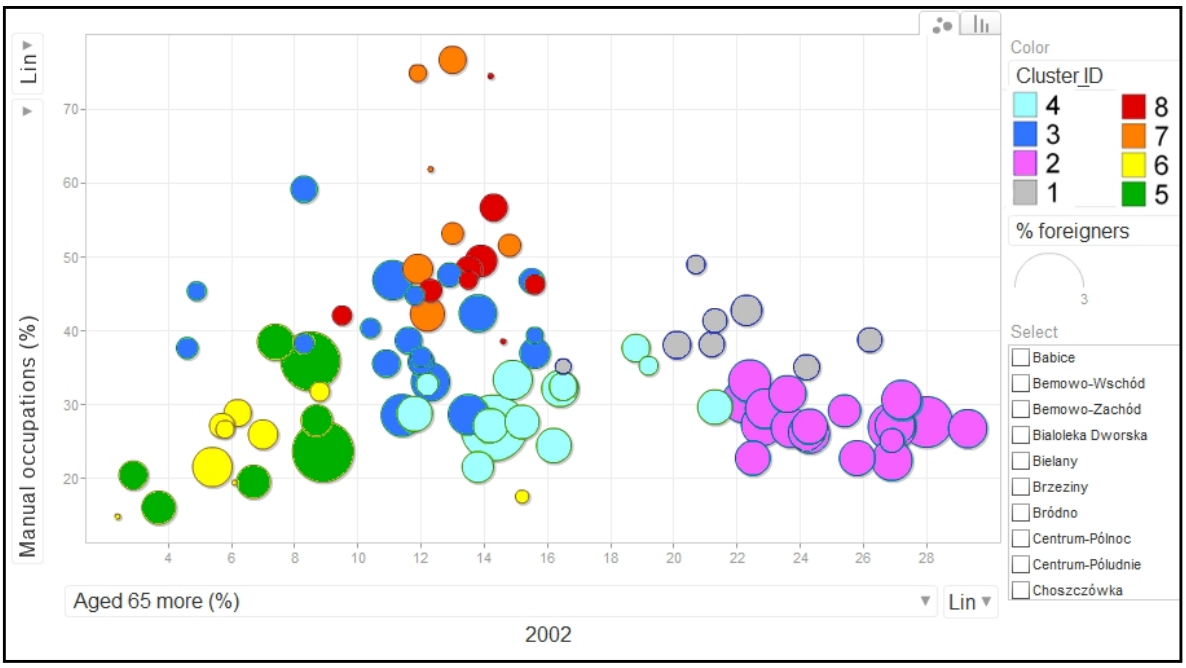


Figure 6.27. Urban Regions in Warsaw plotted on a graph showing the three social dimensions

The graph can be accessed online at
https://spreadsheets0.google.com/ccc?pli=1&hl=en&key=tQ_SfAJ2dotTqJ2xBVO1l3Q&authkey=CJ-aq88N&hl=en#gid=16

In Figure 6.26 of Community Areas in Leeds we can deduce that the LTTI and disability variable is a function of age and socioeconomic status. The higher share of people aged 65 and more and the higher percentage of people with manual occupation, the higher is percentage of population with LLTI or disabled in a Community Area.

Because the disability variable was not used for cluster analysis in Warsaw, three dimensional graphs were prepared for the Polish capital city. The age dimension is represented on axis Y, the occupation dimension is represented on the X axis and bubble size represents percentage of non Polish population. Bubble colour stands for cluster number.

The three dimensional bubble graph for Warsaw also confirms that the cluster analysis did minimise intra-cluster variability. Urban Regions with similar percentage of older population, people with manual occupations and share of foreign residents were ascribed to the same cluster.

The comparison of bubble graphs for Leeds and Warsaw reveals some striking differences among two cities. Firstly, Warsaw Urban Regions have broader range of values of age and occupation variables. Secondly, the Community Areas in the British city are more evenly distributed across variables, while in case of Polish city, the Urban Regions more often have values close to the mean of the variables. What this means is that some clusters in the Polish city are more similar to one another. We conclude that diversity in the two cities is of a very different kind.

6.4. Diversity between cities – combined analysis for the two cities

Presenting separate cluster analysis for Leeds and Warsaw allows us to show diversity within each urban space. But the differences between the two cities are not visible. To bring the intercity diversity into light, a common cluster analysis for Leeds and Warsaw combined was implemented, using only three social dimensions as measures of disability were not available for Warsaw.

However, there are also drawbacks of running common cluster analysis for both cities. The final picture is more flat, specific characteristics for cities are lost and there may be clusters present only in one city. Thus, the common analysis is not a good solution for structuring the survey sample but a common analysis does identify the key differences in history and urban development.

The procedure was one again the same – k-means cluster analysis, Manhattan distance, replicated 200 times and the final cluster solution replicated 1000 times to improve the classification (for details see section 4). We standardized variables together for Leeds and Warsaw, using the standard z score formula. Figure 6.28 profiles the clusters, Figures 6.29 and 6.30 show the spatial distributions of the

clusters in Leeds and Warsaw respectively and Table 6.3 describes the clusters and identifies in which city they are concentrated.

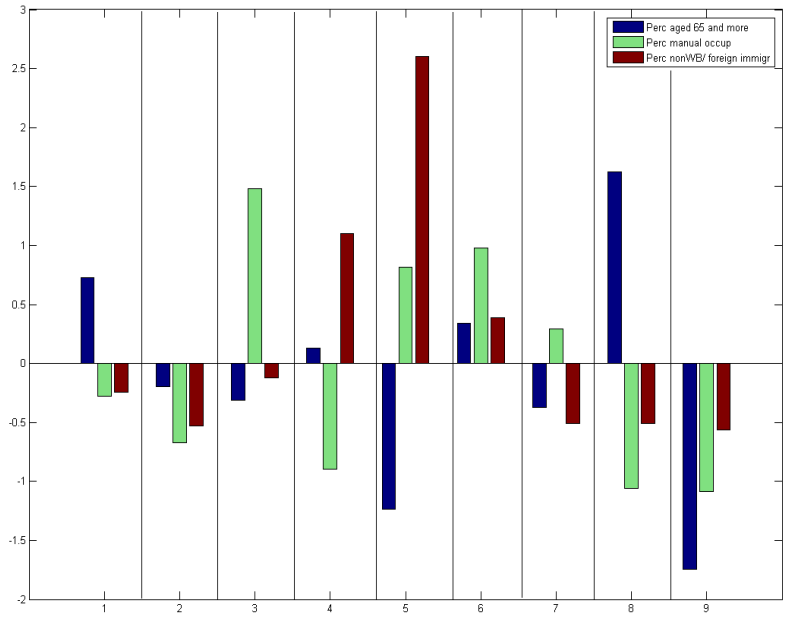


Figure 6.28. Key variables for the 9 cluster solution for both cities (%, k-means, 1000 replicates, Manhattan distance)

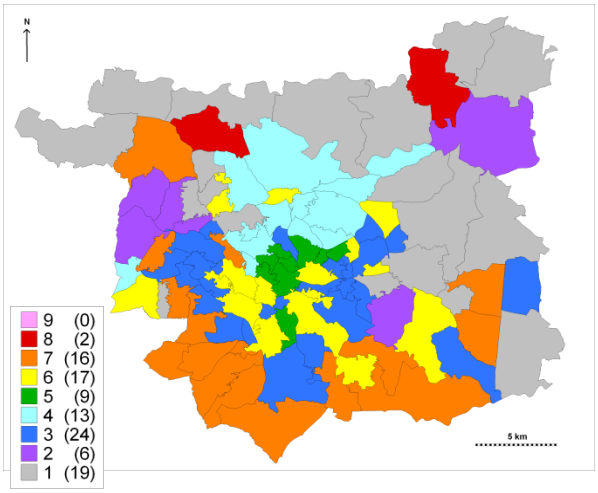


Figure 6.29. CAs classification in Leeds – 9 cluster solution (%, k-means, 1000 replicates, Manhattan distance)

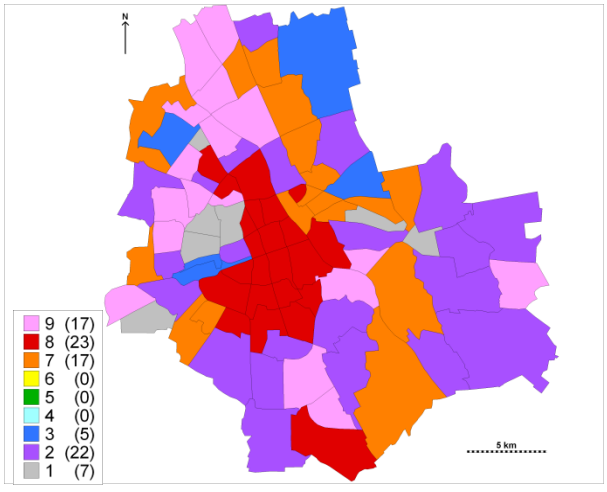


Figure 6.30. URs classification in Warsaw – 9 cluster solution (%, k-means, 1000 replicates, Manhattan distance)

One look on the maps makes it clear that we find two totally different configurations of diversity.

Table 6.3. Clusters produced in the combined analysis for both cities

No	CAs	URs	Location	Descriptions
1	19	7	Mainly Leeds	Older population, different classes, average share of non WB/foreign co-residents
2	6	22	Mainly Warsaw	Different age groups, middle class, White British/Polish residents
3	24	5	Mainly Leeds	Different age groups, working class, White British/Polish residents
4	13	0	Only Leeds	Different age groups, middle class, with non WB/foreign co-residents
5	9	0	Only Leeds	Young population, working class, with the highest share of non WB/foreign co-residents
6	17	0	Only Leeds	Older population, working class, with some non WB/foreign co-residents
7	16	17	Shared	Diverse population, typical traits on average, White British/Polish residents
8	2	23	Mainly Warsaw	The oldest population, middle class, White British/Polish residents
9	0	17	Only Warsaw	The youngest population, middle class, White British /Polish residents

The social geography in two cities is very different. There is only one cluster, cluster 7, that is equally shared by the both cities. Four clusters belong to one city only, namely clusters 4, 5 and 6 to Leeds and cluster 9 to Warsaw. In clusters 1, 3 more than three quarters of member zones in found in Leeds only, while in clusters 2 and 6 most of the member zones is located in Warsaw only.

Clusters that are located only or mainly in Leeds have some non White British co-residents. Due to a much lower share of foreign immigrants residents, the more ethnically diverse clusters are missing in Warsaw (clusters 1, 4, 5 and 6). Meanwhile clusters with higher share of people with non-manual occupation were more specific for the Polish city (clusters 1, 8 and 9) and there was only one middle class cluster in Leeds (cluster 4). The only cluster, cluster 7, that was evenly shared by both cities covers populations with typical age and class traits on average.

6.5. Diversity within cities

The final step in the spatial analysis was to calculate diversity indexes for three or four dimensions. We calculated Simpson's Diversity Indexes for 9 age groups and 9 occupation groups for the both cities and for 9 ethnic groups for Leeds. For disabled population in Leeds and foreign immigrants in Warsaw, beside Simpson's Diversity Indexes we have computed Dissimilarity Indexes, Exposure Indexes and percentages, which better reflect diversity in this dimensions, because we had here dichotomous data, i.e. disabled/nondisabled and foreign/native.

Values of indexes for clusters in Leeds and Warsaw are presented in Tables 6.4 and 6.5. We can compare these with index values for the whole cities (but remember this is a different spatial scale) and to the median for each clusters. We decided to use median instead of mean value, because the first is better for comparisons within non-symmetrical distributions. When determining the median value for each cluster we sorted them from the one with the lowest value of a given index to the one with the highest value, and then they were weighted by the number of CAs or URs in a cluster. We consider the index to be 'high' (red) if it is higher from the median value, 'low' (blue) when lower the median value, and 'medium'(green) when equals the median.

Table 6.4. Diversity indexes for clusters in Leeds

Cluster ID	SI age	SI occupation	SI ethnic	SI disabled/not disabled	DI disabled /notdisabled	EI disabled /notdisabled	EI notdisabled /disabled	% disabled
1	0.865	0.883	0.089	0.293	0.003	0.238	0.052	17.8%
2	0.856	0.877	0.200	0.323	0.028	0.160	0.035	20.3%
3	0.868	0.856	0.161	0.234	0.028	0.061	0.013	13.5%
4	0.695	0.864	0.380	0.219	0.026	0.042	0.009	12.5%
5	0.839	0.879	0.767	0.306	0.002	0.028	0.006	18.9%
6	0.873	0.878	0.093	0.290	0.003	0.111	0.024	17.6%
7	0.870	0.862	0.301	0.281	0.009	0.095	0.021	16.9%
8	0.865	0.869	0.165	0.380	0.039	0.081	0.018	25.5%
Leeds	**0.866**	**0.884**	**0.203**	**0.295**	**0.069**	**81.497**	**17.868**	**18.0%**
Median	**0.865**	**0.877**	**0.161**	**0.293**	**0.009**	**0.111**	**0.024**	**17.8%**

Notes:
Indexes: SI – Simpson's Index of Diversity; DI – Dissimilarity Index; EI – Exposure Index; % – percentage.
Populations: age – 9 age groups; occupation – 9 occupation groups; ethnic – 9 ethnic groups, disabled – people with LLTI, bad health and disability.

Both cities have relatively high age and occupational diversity but the latter is lower in case of Polish city. Ethnic diversity is very low in Warsaw in comparison to Leeds. In the fourth disability dimension, diversity remains relatively low if we compare it with the age and the occupational diversities, but it is higher than ethnic diversity in Leeds. Generally, in Leeds the diversity index was the highest for socioeconomic dimension, in Warsaw for the demographic dimension, and in both cities the lowest diversity indexes were for the ethnic dimension of diversity.

Table 6.5. Diversity indexes for clusters in Warsaw

Cluster ID	SI age	SI occupation	SI foreign/PL	DI foreign/PL	EI foreign/PL	EI PL/foreign	% foreign
1	0.873	0.830	0.008	0.040	0.080	0.001	0.42%
2	0.876	0.789	0.019	0.155	0.448	0.003	0.94%
3	0.857	0.841	0.008	0.029	0.061	0	0.42%
4	0.868	0.812	0.014	0.021	0.210	0.001	0.69%
5	0.830	0.760	0.014	0.006	0.052	0	0.71%
6	0.836	0.807	0.008	0.054	0.088	0.001	0.38%
7	0.860	0.831	0.009	0.008	0.020	0	0.44%
8	0.862	0.841	0.005	0.052	0.034	0	0.24%
Warsaw	**0.868**	**0.815**	**0.012**	**0.012**	**0.183**	**99.279**	**0.62%**
Median	**0.862**	**0.812**	**0.009**	**0.029**	**0.080**	**0.001**	**0.44%**

Note:
Indexes: SI – Simpson's Index of Diversity; DI – Dissimilarity Index; EI – Exposure Index; % – percentage.
Populations: age – 9 age groups; occupation – 9 occupation groups; foreign – foreign immigrants, PL – Polish residents.

Comparisons of diversity indexes among clusters enable us to make further interesting observations. SI for disabled/not disabled in Leeds was classified in the same way as the percentages for disabled groups. Similarly, in case of Warsaw SI for foreign/native and percentage of foreign immigrants classified the same clusters as low, medium or high. In next section we will analyse SI and percentage values only.

Bubble graphs will help the reader to follow the diversity indexes' values. We present in Figures 6.31 and 6.32 diversity for Simpson's Diversity Indexes in two or four dimensions and percentage values for disabled population in Leeds and foreign immigrants in Warsaw. More graphs are available in Appendix A7. Diversity indexes in 4 out of 8 clusters for Leeds have mixed diversity level, namely in clusters 1, 2, 3 and 8 we can find low, medium and high diversity level in the fourth social dimensions. In other clusters we found different configurations of diversity levels. Cluster 4 has three low indexes and one classified as high and it was the ethnic dimension. On contrary, cluster 5 has three high indexes and one classified as high – the demographic dimension. Clusters 6 and 7 have two high diversity indexes and two – low. In case of cluster 6 high indexes values were obtained for the ethnicity and the disability dimensions, and for cluster 7 – the occupational and the disability dimensions. The summary of the diversity indexes and their values for Leeds are provided in the Table 6.6 and for Warsaw in Table 6.7.

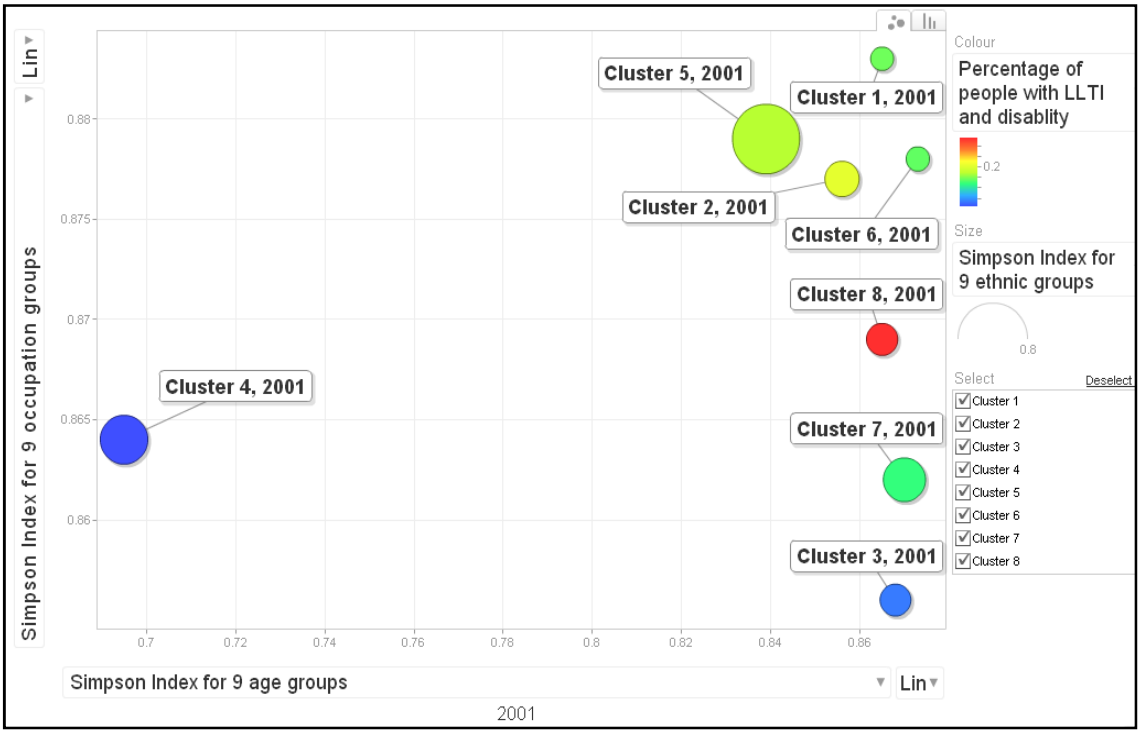


Figure 6.31. Simpson indexes for 3 dimensions and percentages for 4 dimensions

Table 6.6. Summary of the diversity indexes values for Leeds

Cluster ID	SI age	SI occupation	SI ethnic	SI/% disabled/notdisabled	SUMMARY
5	low	High	high	high	3H1L
8	medium	low	high	high	2H1M1L
2	low	medium	high	high	2H1M1L
6	high	high	low	low	2H2L
7	high	low	high	low	2H2L
1	medium	high	low	medium	1H2M1L
3	high	low	medium	low	2L1M1H
4	low	low	high	low	3L1H

For Warsaw the diversity indexes are presented in the bubble graph (Figure 6.32).

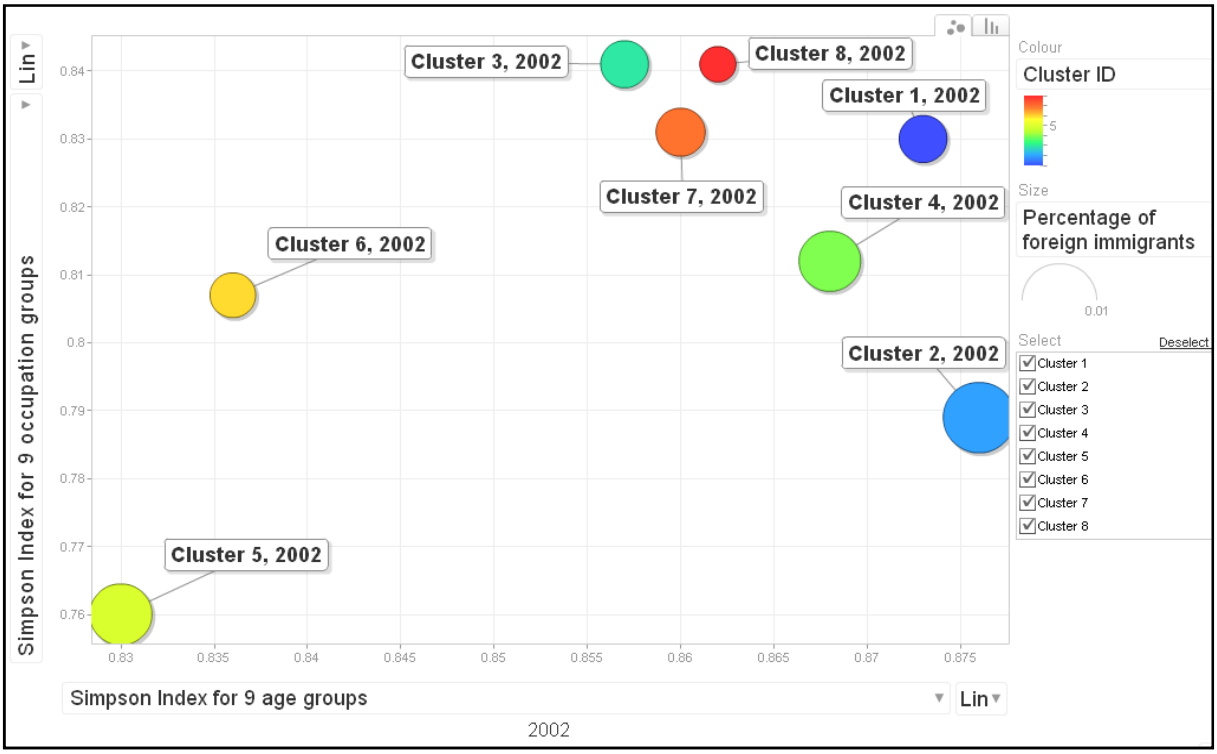


Figure 6.32. Simpson indexes for 2 dimensions and percentages for 1 dimension

Most of the clusters have mixed diversity level measured with SI – either two high or low indexes and one low or high, or each index belonging to different diversity level (i.e. low, medium and high). The highest diversity is in cluster 4 which has two indexes high and one medium. It is followed by clusters 1 and 2 which have two high and two low values of diversity indexes. Next two clusters, 7 and 8, have low, medium and high values of diversity. Clusters 3 and 5 have two low values and 1 high. Finally, there is one cluster in Warsaw that has all three Simpson's Diversity Indexes at a low level – cluster 6. The summary of the diversity indexes and their values for Warsaw are provided in the Table 6.7.

Table 6.7. Summary of the diversity indexes values for Warsaw

Cluster ID	SI age	SI occupation	SI/% foreign/PL	SUMMARY
4	high	medium	high	2H1M
1	high	high	low	2H1L
2	high	low	high	2H1L
7	low	high	medium	1H1M1L
8	medium	high	low	1H1M1L
3	low	high	low	2L1H
5	low	low	high	2L1H
6	low	low	low	3L

It would be possible to prepare one overall diversity index for each cluster, e.g. by adding the dimensional indexes and weighting them. However, we argue that preparing one overall diversity

index for each cluster is not reasonable. Diversity is a complex phenomenon. As we have shown throughout the paper, it cannot be reduced to one dimension. Operating with three or four different dimensions that refer to different and important life courses of people, namely family/demographic, socioeconomic, ethnicity/affinity and disability, allows to keep the diversity picture as comprehensive and as detailed as possible. Consequently, we prefer to keep the dimensional indexes of diversity for further comparisons that will be undertaken when the survey data will be analysed.

7. RECOMMENDATIONS FOR THE SURVEY AND QUALITATIVE STUDIES

Through our cluster analysis we have identified in Leeds and Warsaw a set of community types, each of which has a distinctive profile in terms of social dimensions. It is vital therefore to place the planned survey of attitudes towards people 'of difference' and of experiences in encounters with others in the context of our typology. We will, in due course, wish to test whether community types have an independent effect on people's attitudes and encounters.

It is important in the planned survey to find respondents who live in the different community types in Leeds and Warsaw, and, at the same time, to ensure the survey sample is representative. These twin goals can be achieved by using the Leeds and Warsaw community types as sampling strata in the survey. We have recommended the following strategy:

- We wish to procure a representative sample of the population of Leeds and Warsaw through two interview surveys.
- To ensure representativeness of the sample across the dimensions of social diversity, we have classified the neighbourhood communities of Leeds into 8 clusters (CL) and the neighbourhood communities of Warsaw into 8 clusters.
- For the Leeds clusters we will supply lists of the unit postcode areas which constitute the clusters of diversity. We can sort the postcodes by CL, CA and OA and output a SPSS and an EXCEL file containing the sorted list of unit postcodes. That means that we have lists of postcodes associated with our Diversity Clusters and we can define sample structure for every cluster. For Warsaw we will work with the survey company to access address information, as postcodes are too large in Warsaw to be useful.
- The total sample size for the survey will be 1500 interviews in each city (3000 in total). We wish these interviews to be assigned equally across the 8 clusters as sampling strata in each city. By having equal sample sizes the cluster results will have the same confidence intervals and we can measure confidently the differences between cluster 'mean' responses. The cluster interviews will be reweighted after delivery of the interview results to obtain a representative sample for the two cities as a whole.
- We will work with from the survey company to construct a representative sample of the population of each city. By representative sample we mean one in which every member of the target population has a known probability of selection.
- Finally, we intend to quality assure the representativeness of the sample against a simulated Leeds 2011 population and against 2010 register counts in Warsaw and later against the results of the 2011 UK and Poland Censuses.

8. EXECUTIVE SUMMARY

The aim of the paper was to understand social diversity in two European cities: Leeds and Warsaw. The analyses are part of the European Research Council (ERC) project called *Living with Difference* and one of its sub-projects *Mapping Social Diversity: Opportunities for Encounters*. In order to find out what opportunities are available for encountering difference we have analysed statistical data for residential communities in the two cities. The social survey, which will follow the secondary data analysis, will explore what types of encounter are sought or avoided and by whom, and what are people's experiences of discrimination or exclusion. We argue that the results will differ for communities depending on their population composition across many dimensions of diversity. This proposition will be tested when the survey is analysed.

The key findings of the research are as follows.

Understanding of diversity

(1) We reviewed quantitative and qualitative studies on spatial diversity and residential patterns in cities. While in previous studies on diversity the emphasis was put on the dimensions of social class and ethnic background, in the qualitative studies diversity was captured more broadly including other characteristics, age and generation, gender, disability, sexual orientation or religion. **We aimed to incorporate the broader thinking about diversity into the mapping social diversity**.

(2) We reviewed studies of social diversity to find out different connections between such characteristics as, **age and generation, gender, disability, sexual orientation, ethnicity, religion and socio-economic status**. We concluded that many different factors determine how diverse residential communities are, where different social groups concentrate or where they are underrepresented. However, the diversity of residential patterns cannot be explained by one set of factors only, e.g. age, gender or ethnicity. All dimensions need to be considered simultaneously.

(3) Using concepts of social group, social category and life course we have divided diversity into **four dimensions: family demographic status** (depending on gender, sexual orientation and family type), **socio-economic status** (depending on education, occupation and income), **ethnicity/affinity** (depending on nationality, ethnicity and religion) **and disability** (which is a function of on health, social definition and individual perspective). Throughout the paper we

argue that it is a set of characteristics, organized by us into diversity dimensions that will determine the nature of social diversity in the two studied cities.

Comparability of census data

(4) The diversity analyses were based on census data. We **compared the nature of 2001 British census and 2002 Polish census data** – census organisation, census populations, census geographies and availability of census data.

(5) We decided to use **Community Areas (CAs) for Leeds and Urban Regions (URs) for Warsaw** as our spatially comparable study zones. These units were the most appropriate study zones in relation to the aim of the study: to understand how people encounter diversity in their residential neighbourhoods. CAs and URs have similar sizes and their names are meaningful to people. Census data for smaller statistical units (Output Areas in Leeds and Statistical Regions in Warsaw) were aggregated to form the CAs and URs.

(6) Two census dataset were compared variable by variable, definition by definition. We decided which variables were defined and measured in a similar way to conduct unbiased analysis. We obtained a list of 28 comparable variables that was reduced by eliminating highly correlated variables to the **final list of 14 variables** (12 in case of Warsaw, because we did not have health and disability data in the Polish dataset).

Measurement of social diversity

(7) We used cluster analysis to determine the nature of diversity in Leeds and Warsaw residential communities. The clustering method was tested by running a simple analysis with four variables only (one per each diversity dimension). We then ran **k-means clustering, using Manhattan distance and 200 replicates for 2-16 solutions**. The best cluster solutions, those with the most even distribution of cases and the lower number of cases that could be ascribed to other clusters, were selected on the basis of absolute average difference from the mean cluster size and through use of silhouette plots.

(8) In order to measure diversity in the four dimensions at the residential community level we used **Simpson's Diversity Index, the Dissimilarity Index and the Exposure Index**. But the indexes proved to be poor variables for the cluster analysis and we used percentage variables in the cluster analysis only.

Diversity in Leeds and Warsaw

(9) Diversity between the two cities was compared by mapping the variables one at a time and running a cluster analysis which combines the varaibles. **The two cities differ considerably**. While in Leeds the older population lives in the suburban areas, in Warsaw they tend to concentrate in the city centre. On average the share of residents with manual occupations is much lower in Warsaw than in Leeds, so there are only a few residential areas, post-industrial regions, where people belong to the working class are visibly overrepresented. Meanwhile in Leeds 4 out of 10 communities had more residents with manual occupations than with non-manual. Regarding the ethnic composition of the two cities, Warsaw is a homogenous area, with a share of foreign residents lower than 1%, while in Leeds the share was almost 11%.

(10) Consequently, the combined cluster produced a typology with only one cluster shared by both cities; the rest belonged mainly to one city only. **The two cities have different social geographies** – this is a result of different political histories, different roles played by the cities within their respective regions and countries, and their unique urban development.

(11) Diversity within the two cities was measured with several diversity indexes. The differences between the diversity levels for the four dimensions have shown that the **diversity cannot be reduced to one dimension and measured by one indicator**. Operating with three or four different dimensions that refer to different and important life courses of people, namely family/demographic, socioeconomic, ethnicity/affinity and disability, allows us to keep a more comprehensive and detailed picture of diversity.

Typology of spatial communities

(12) We produced 8-cluster solutions for both cities. **Multidimensional diversity was presented using graphs of key variables, selected key statistics, lists of CAs and URs and maps for all clusters together and each cluster separately**. In both cities population was unevenly distributed across clusters and the cluster sizes varied from 3% to 30% of resident population.

(13) **Each cluster had different characteristics in the four dimensions of diversity**, meaning that combinations of diversity varied, in their stage in the demographic or socio-economic cycle, in their shares of ethnic minorities and in the health conditions of the resident

population. All clusters were provided with short descriptions and interpretations for the four dimensions of diversity.

(14) The clusters were presented in four or three dimensional bubble graphs (the graphs can be browsed online). The multidimensional graphs confirmed that the two cities differ in their diversity make-up. Community Areas in Leeds are more dispersed across variables, while in case of the Polish city the Urban Regions more often have values close to the mean of the variables. What this means is that **some clusters in the Polish city are more similar to one another** than those in Leeds.

Social survey

(15) **Clusters will be used as sampling strata in social surveys in the two cities**. This will enable us to test the hypothesis that community types, with different diversity mixtures, have an independent effect on people's attitudes and encounters.

9. DISCUSSION AND CONCLUSIONS

The two cities have different social geographies resulting from different political histories, different roles played by the cities within their respective regions and countries, and their unique urban development. Consequently, there exist different opportunities to have contacts with people who are different in terms of age, ethnicity, religion/belief, disability and socio-economic status. The diversity maps and graphs presented in this paper show that the two cities differ considerably in terms of diversity. Employing a broader understanding of diversity enabled us to present more comprehensive profiles of communities in both cities.

Societies have always been diverse, but different groups are not evenly distributed across cities reflecting residential preferences and other processes (like housing policy) that are responsible for diversity patterns in the urban context. While ethnic diversity can be positive for cities based on educationally intensive industries, because it supports the information sharing and creativity (Lee 2010), the effects for social relations at the neighbourhood level may be different (Talen 2009). We have shown that the opportunities to encounter diversity vary within both Leeds and Warsaw. The survey analysis, in-depth interviews and focus groups will examine whether proximity generates meaningful contact among diverse social groups and respect and understanding of difference between majority and minority populations (Valentine 2008)[25].

The simple analysis cluster solution proved to be a good 'guess' for the diversity cluster in the 14/12 variable analysis across two cities, which means that variables used in the simple analysis, namely age, occupation, ethnicity/nationality and disability, have descriptive value. However, adding more variables enabled us to refine the cities' cluster profiles and the cluster solutions were improved. Interestingly, diversity cluster solutions from the simple and complex analyses are almost identical for the central residential areas of Leeds and Warsaw. In both cases the cluster solutions persisted unchanged for the central areas of both cities.

Diversity analysis has shown that measurement of diversity cannot be reduced to one indicator, because diversity is too diverse to be captured in one index. Diversity cluster maps and three/four dimensional graphs of diversity indexes clearly exposed the variations. A low/medium/high scale of diversity for clusters should be determined separately for three or four diversity dimensions, because applying 'overall diversity index' would flatten the picture and consequently valuable information would be lost.

[25] That will be done in the Projects B and C of the "Living with Difference" project.

We have analysed social diversity at the community level. However, we do not assume that people living in the communities are the same. People with different characteristics from the average profile of the cluster population may live there. For example, the socio-spatial structure of Warsaw is believed to have a 'mosaic' pattern (Węcławowicz 2005). So within middle class regions there exist estates or streets with people belonging to the working class. Moreover, the analysis was based on 2001 and 2002 population censuses and we are aware that the cluster profiles could have changed in last years. The inter-community differences and changes in the characteristics of communities will be the subject of in-depth qualitative studies undertaken in the "Living with Difference" project.

The composition of diversity may affect attitudes towards 'other' people, prejudice and discriminatory behaviours. While in Leeds the clusters with younger and both middle and working class populations are the most ethnically diverse, in Warsaw minority ethnic groups more often share urban space with older population and residents with non-manual jobs. Age and economic status constitute two of the explanatory factors in shaping people's attitudes (see the 'in-group/out-group' concept of Tajfel 1982). People in a different phase of a life course and those recruiting from different social classes have dissimilar needs and life styles that could be contradictory. Another factor is the potential to encounter social diversity (see the 'social learning' concept of Bandura 1977). The number of non-native residents is considerably lower in Warsaw, so the contact with other ethnic groups is limited and less frequent than in Leeds. Still, even people who live in proximity can self-segregate within particular micro spaces (Valentine 2008), so the greater residential mixing does not necessarily lead to social integration (Phillips 2006).

The opportunities to encounter diversity will vary from community to community, but they may be accompanied by various social processes at the community level. Depending on the attitudes of local population the different groups living in the communities may share values and identities, live separately, or even in conflict. There is different space and consent for diversity in communities that want to assimilate different groups, integrate them, become more socially cohesive or multicultural. The social diversity that we have presented for the two cities in the paper may bring different results. If social diversity is accompanied by social distance, the people that live together as neighbours will be 'strangers'/'others' in the sociological meaning (Simmel 1908). Co-existing in the same urban space is not necessarily enough for living harmoniously with difference. Attitudes and beliefs also matter.

As we declared in the theoretical introductory parts of the paper, we hope that the mapping exercise is a good starting point for a further discussion on understanding diversity in the past and present. Obviously, an 'ethnicity limited' perspective is not sufficient. As our diversity cluster solutions have

shown, communities with similar level of ethnic diversity happen to have different levels of diversity in other dimensions – some of them are inhabited more often by older populations, others by young people, in some more frequently live people that belong to the middle class and in other areas to the working class. In every cluster in Leeds and Warsaw there exist different opportunities to encounter diversity.

Diversity is usually equated with 'ethnic diversity'. The results show here that it is worthwhile to think about diversity beyond the ethnicity dimension. The super-diversity concept recognises social diversity among minority ethnic groups regarding their migration status, religious faith, age, gender or working status (Vertovec 2007). The argument is that ethnic origins very often simplify and homogenise our research perspective on migration. However, focusing on ethnic, religious or language minorities simplifies research perspective in another way. People with different ethnic backgrounds not only represent heterogeneous groups, but also live among diverse populations regarding age, gender, religious faith, (dis)ability conditions or working status. Thus, social processes, such as integration, are related to the composition of diversity at the local community level.

Moreover, shifting the discussion from the ethnicity oriented perspective to social diversity allows us to eliminate minority ethnic stigmatisation. We should avoid putting too much emphasis on the cultural agenda of different ethnic groups and rather focus on the shared characteristics and common interests of the whole society, e.g. stemming from being co-workers or neighbours (Vertovec, Wessendorf 2004). Taking into account the fact that the concept of multiculturalism has recently been disparaged by a number of politicians in Europe, the discussions on diversity need to be reshaped. Thinking broadly about social and cultural diversity will help to 'de-racialise' the political debates in the UK, and help to start the political debate on diversity in Poland with a broader perspective.

REFERENCES

Abramowicz M. (ed.) (2007). *Situation of bisexual and homosexual persons in Poland 2005 and 2006 report*. Warsaw: Campaign Against Homophobia. Accessed 31 January 2011, online at: http://world.kph.org.pl/pdf/report_homophobia_poland_2007_en.pdf.

Abramsson M., Borgegard L.E., Fransson U. (2002). Housing Careers: Immigrants in Local Swedish Housing. Markets. *Housing Studies* 17(3): 445-464.

Aggarwal C.C., Hinnenburg, A. Keim, D.A. (2001). On the surprising behaviour of distance metrics in high dimensional space. In: Van den Bussche J., Vianu V. (eds.). *International Conference on Database Theory 2001, Lecture Notes in Computer Science*, vol. 1973, pp. 420-434.

Ali A.H. (2006). *The Caged Virgin: An Emancipation Proclamation for Women and Islam*. New York: Free Press.

Alvanides S., Openshaw S., Rees P. (2002). Designing your own geographies. In: Rees P., Martin D., Williamson P. (eds.). *The Census Data System*. Chichester: John Wiley and Sons, pp. 47-66.

Bandura A. (1977). *Social Learning Theory*. New York: General Learning Press.

Beale M.H., Hagan M.T., Howard B.D. (2009). *MATLAB Statistics Toolbox 7 User's Guide*. The MathWorks Inc., Natick, MA. Accessed 10 January 2011, online at**:** http://www.mathworks.com/help/pdf_doc/nnet/nnet.pdf**.**

Becker G. (2003). Meanings of place and displacement in three groups of older immigrants. *Journal of Aging Studies* 17(2): 129-149.

Bijak J., Kicinger A., Kupiszewski M., in cooperation with P. Śleszyński (2007). *Studium metodologiczne oszacowania rzeczywistej liczby ludności Warszawy [Feasibility study for the estimation of the actual population size of the city of Warsaw]*. CEFMR Working Paper 2/2007. Warsaw: CEFMR.

Blom S. (1999). Residential Concentration Among Immigrants in Oslo. *International Migration* 37(3): 617-642.

Bolt G., Van Kempen R. (2002). Moving Up or Moving Down? Housing Careers of. Turks and Moroccans in Utrecht, the Netherlands. *Housing Studies* 17(3): 401-422.

Bondi L. (1992). Gender symbols and urban landscapes. *Progress in Human Geography* 16(2):157-170.

Booth C. (1889). *Life and Labour of the People. First Series (i) East, Central and South London*. London: Macmillan, republished 1969.

Bowes A., Dar N., Sim D. (2002). Differentiation in housing careers: the case of. Pakistanis in the UK. *Housing* Studies 17(3): 381-400.

Brown D., Rees P. (2006). Trends in Local and Small Area Mortality and Morbidity in Yorkshire and the Humber: Monitoring Health Inequalities. *Regional Studies* 40: 437-458.

Buck N., Gordon I., Hall P., Harloe M., Kleinman M. (2002). *Working Capital: Life and Labour in Contemporary London*. London: Routledge.

Burgess E.W. (1925). The growth of the city: an introduction to a research project. In: Park R.E., Burgess E.W., McKenzie R.D. *The City*. Chicago: University of Chicago Press, pp.47-62.

Burgess S., Wilson D., Lupton R. (2005). Parallel lives? Ethnic segregation on schools and neighbourhoods. *Urban Studies* 42(7): 1027-1056.

Butler R., Bowlby S. (1997). Bodies and spaces: an exploration of disabled people's experiences of public space. *Environment and Planning D: Society and Space* 15(4): 411-433.

Castles S., Miller M.J. (1998). *The Age of Migration: International Population Movements in the Modern World*. London: Macmillan.

CBOS 2010). *Postawy wobec gejów i lesbijek [Attitudes towards gays and lesbians]*. Komunikat z Badan. BS/95/2010. Accessed 31 January 2011, online at: www.cbos.pl/SPISKOM.POL/2010/K_095_10.PDF.

Cesari J. (2005). Mosques in French Cities: Towards the End of a Conflict? *Journal of Ethnic and Migration Studies* 31(6): 1025-1043.

Chauncey G. (1995). Gay New York: The Making of the Gay Male World, 1890–1940. London: Flamingo.

Chouinard V. (1997). Making space for disabling differences: challenging ableist geographies. *Environment and Planning D: Society and Space* 15(4): 379-387.

Cowan K. (2007). *Living Together: British Attitudes to Lesbian and Gay People*. Stonewall. Accessed 31 January 2011, online at: www.stonewall.org.uk/documents/living_together_final_web.pdf.

Daley P.A. (1998). Black Africans in Great Britain: spatial concentration and segregation. *Urban Studies* 35(10): 1703-1724.

Dangshat J., Blasius J. (1987) Social and spatial disparities in Warsaw in 1978: An application to correspondence analysis to a 'Socialist City'. *Urban Studies* , 24, 173-191.

Darroch G., Marston W.G. (1971). The Social Class Basis of Ethnic Residential. Segregation: The Canadian Case. *The American Journal of Sociology* 77(3): 491-510.

Dear M., Gaber S. L., Takahashi L., Wilton R. (1997). Seeing people differently: The socio-spatial construction of disability. *Environment and Planning D: Society and Space* 15(4): 455-480.
Dennett A., Stillwell J.C.H. (2008). Internal migration in Great Britain – a district level analysis using 2001 Census data. Working Paper 08/01. School of Geography, University of Leeds. Accessed 28 December 2010, online at: http://www.geog.leeds.ac.uk/fileadmin/downloads/school/research/wpapers/internal.pdf.
Deurloo R., Musterd S. (2001). Residential Profiles of Surinamese and Moroccans in Amsterdam. *Urban Studies* 38(3): 467-485.
Diamond I., Cruddas M., Woolford J. (2002). A one-number census. In: Rees P., Martin D., Williamson P. (eds.). *The Census Data System*. Chichester: John Wiley and Sons, pp. 295-304.
Dorling D., Rees P. (2003). A nation still dividing: the British census and social polarisation 1971-2001. *Environment and Planning A* 35(7): 1287-1313.
Duncan N. (2009). Social Class. In: Kitchen R., Thrift N. (eds.) *The International Encyclopedia of Human Geography*. Oxford: Elsevier, pp. 179-184.
Duncan O.D., Duncan B. (1955). A methodological analysis if segregation indices. *American Sociological Review* 20: 210-217.
Duncan O.D., Duncan, B. (1957). *The Negro Population of Chicago: a Study in Residential Succession*. Chicago: University of Chicago Press.
Durmaz B. (2009). *Döner, Machos und Migranten: Mein zartbitteres Lehrerleben*. Freiburg: Herder.
Engels F. (1845). *Die Lage der Arbeitenden Klasse in England*. Translated by F.K. Wischnewetzky (1887) as *The Condition of the Working Class in England in 1844*. London: George Allen and Unwin. Accessed 28 December 2010, online at: http://www.marxists.org/archive/marx/works/1845/conditionworkingclass/index.htm.
Everitt B.S., Landau S., Leese M. (2001). *Cluster Analysis 4th Ed*. London: Arnold.
Fadda G., Cortés A., Olivi A., Tovar M. (2010). The perception of the values of urban space by senior citizens of Valparaiso. *Journal of Aging Studies* 24(4): 344-357.
Flint J. (2009). Neighbourhood and Community. In: Kitchen R., Thrift N. (eds.) *The International Encyclopedia of Human Geography*. Oxford: Elsevier, pp. 354-259.
Flint J. (2010). Faith and Housing in England: Promoting Community Cohesion or Contributing to Urban Segregation? *Journal of Ethnic and Migration Studies* 36(2): 257-274.
Freeman L. (2009). Neighbourhood Diversity, Metropolitan Segregation and Gentrification: What Are the Links to the US? *Urban Studies* 46(10): 2079-2101.
Friedrichs J. (1998). Ethnic Segregation in Cologne, Germany, 1984-94. *Urban Studies* 35(10): 1745-1764.
Giffinger R. (1998). Segregation in Vienna: Impacts of market barriers and rent regulation. *Urban Studies* 35(10): 1791-1812.
Gini C. (1912). *Variabilità e mutabilità*. Bologna: Tipografia di Paolo Cuppini. Reprinted in E. Pizetti and T. Salvemini (eds.). *Memorie di metodologica statistica*. Rome: Libreria Eredi Virgilio Veschi (1955).
Gini C. (1921). Measurement of Inequality of Incomes. *The Economic Journal* (Blackwell Publishing) 31(121): 124-126.
Główny Urząd Statystyczny [Central Statistical Office in Poland] (2003a). *Historia spisów [History of censuses]*. Warsaw: Główny Urząd Statystyczny. Accessed 28 December 2010, online at: http://www.stat.gov.pl/gus/7853_PLK_HTML.htm.
Goodman L.A., Kruskal W.H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association* 49: 732-764.
Goodman L.A., Kruskal W.H. (1959). Measures of association for cross classifications. II: further discussions and references. *Journal of the American Statistical Association* 54: 124-163.
Grzymała-Kazłowska A. (2002), The Formation of Ethnic Representations: The Vietnamese in Poland. Sussex Migration Working Papers 8. Accessed 28 December 2010, online at: http://www.sussex.ac.uk/migration/documents/mwp8.pdf.
Grzymała-Kazłowska A., Piekut A. (2007). 'Mała Ukraina' czy Polskie *Viettown*? Społeczno-Przestrzenne Wzory Zamieszkiwania Imigrantów w Metropolii Warszawskiej ['Little Ukraine' or Polish Viettown? Social and Spatial Patterns of Immigrant Settlement in Warsaw Agglomeration]. *Studia Regionalne i Lokalne* 4(30): 77-99.
GUS (2003b). *Ludność. Stan i struktura demograficzno-społeczna. Raport z Narodowego Spisu Powszechnego 2002 [Population. State and demographic-social structure. 2002 Census Report]*. Warsaw: Główny Urząd Statystyczny [Central Statistical Office in Poland]. Accessed 28 December 2010, online at: http://www.stat.gov.pl/gus/8195_PLK_HTML.htm.
GUS (2003c). *Gospodarstwa domowe i rodziny. Raport z Narodowego Spisu Powszechnego 2002 [Households and families. 2002 Census Report]*. Warsaw: Główny Urząd Statystyczny [Central Statistical Office in Poland]. Accessed 28 December 2010, online at: http://www.stat.gov.pl/gus/8197_PLK_HTML.htm.

Hall S. (2001). The multicultural question. *Pavis Papers in Social and Cultural Research* no. 4. Milton Keynes: Open University. Accessed 23 February 2011, online at: http://www.open.ac.uk/socialsciences/includes/__cms/download.php?file=a5rnojn2m6esrh0a3i.pdf&name=pavis_paper_4.pdf.
Harland K., Stillwell J.C.H. (2007). Evidence of ethnic minority dispersal in Leeds. *The Yorkshire and Humber Regional Review* 17(2): 21-24.
Harris R., Sleight P., Webber R. (2005). *Geodemographics, GIS and Neighbourhood Targeting*. London: Wiley.
Helly D. (2003). Social cohesion and cultural plurality. *Canadian Journal of Sociology* 28(1): 19-42.
Hoyt H. (1939). *The Structure and Growth of Residential Neighborhoods in American Cities*. Washington DC: Federal Housing Administration.
Jazeel T. (2009). Difference/Politics of Difference. In: Kitchen R., Thrift N. (eds.) *The International Encyclopedia of Human Geography*. Oxford: Elsevier, pp. 164-169.
Johnston R., Pattie C., Dorling D., MacAllister I., Tunstall H., Rossiter, D. (2000). The neighbourhood effect and voting in England and Wales: Real or imagined? *Journal of Elections, Public Opinion & Parties* 10(1): 47-63.
King M.L. Jr. (1963). *I Have a Dream***.** Speech delivered 28 August 1963, at the Lincoln Memorial, Washington D.C. Accessed 28 December 2010, online at: http://www.americanrhetoric.com/speeches/mlkihaveadream.htm.
Koser K., Lutz H. (1998). The New Migration in Europe: Contexts, Constructions and Realities. In: Koser K., Lutz H. (eds.) *The new Migration in Europe. Social Constructions and Social Realities*. London: MacMillan, pp.1-17.
Kupiszewski M., Durham H., Rees, P.H. (1998). Internal migration and urban change in Poland. *European Journal of Population* 14(3): 265-290.
Lee N. (2010). Ethnic Diversity and Employment Growth in English Cities. *Urban Studies* 48(2): 407-425.
Levine M.P. (1979). Gay Ghetto. *Journal of Homosexuality* 4(4): 363-377. Reprinted in expanded form as 'YMCA': The Social Organization of Gay Male Life. In: *Gay Macho: The Life and Death of the Homosexual Clone*. New York: New York University Press, pp. 30-54.
Li W., Skop E. (2009). Ethnicity. In: Kitchen R., Thrift N. (eds.) *The International Encyclopedia of Human Geography*. Oxford: Elsevier, pp. 615-619
Lieberson S. (1958). Ethnic Groups and the Practice of Medicine. *American Sociological Review* 23(5): 542-549.
Lorenz M.O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Association* 9(70): 209-219.
Magnusson L., Özüekren S. (2002). The housing careers of Turkish households in middle-sized Swedish municipalities. *Housing Studies* 17(3): 465-486.
Martin D. (2002). Output areas in 2001. In: Rees P., Martin D., Williamson P. (eds.). *The Census Data System*. Chichester: John Wiley and Sons, pp. 37-46.
Massey D.S., Denton N.A. (1988). The dimensions of residential segregation. *Social Forces* 67(1): 281-315.
Matthews M. H., Vujakovic P. (1995). Private worlds and public places: mapping the environmental values of wheelchair users. *Environment and Planning A* 27(7): 1069-1083.
Maxey L. (2009). Ageism and Age. In: Kitchen R., Thrift N. (eds.) The International Encyclopedia of Human Geography. Oxford: Elsevier, pp. 42-47.
McCullough C. (1991) *The First Man in Rome*. London: Arrow Books.
McDowell L. (1983). Towards an understanding of the gender division of urban space. *Environment and Planning D: Society and Space* 1(1): 59-72.
McDowell L. (1993). Space, place and gender relations: Part I. feminist empiricism and the geography of social relations. *Progress in Human Geography* 17(2): 157-179.
McDowell L. (1999). *Gender, Identity and Place: Understanding Feminist Geographies*. Oxford: Blackwell.
McDowell L., Perrons D., Fagan C., Ray K., Ward K. (2005). The contradictions and intersections of class and gender in a global city: placing working women's lives on the research agenda. *Environment and Planning A* 37(3): 441-461.
Mez J., Bühler E. (1998). Functional and spatial segregation in the Swiss financial sector: pink-collar ghetto and male bastion. *Environment and Planning A* 30(9): 1643-1660.
Miller V. (2009). Gay Geographies. In: Kitchen R., Thrift N. (eds.) The International Encyclopedia of Human Geography. Oxford: Elsevier, pp. 302-308.
Milligan G.W., Cooper M.C. (1987). Methodological review: clustering methods. *Applied Psychological Measurement* 11: 329-354.
Miranne K.B., Young A.H. (2000) (eds.). *Gendering the city: Women, boundaries and visions of urban life*. Lanham, MD: Rowman & Littlefield.

Munoz S.A. (2010). Geographies of Faith: The Differing Residential Patterns of the Indian-Hindu, Indian-Sikh and Indian-Muslim Populations of Dundee and Glasgow. *Population, Space and Place* 16: 269-285.
Murdie R.A. (2002). The Housing Careers of Polish and Somali Newcomers in Toronto's Rental Market. *Housing Studies* 17(3): 423-443.
Murray S.O. (1979). The Institutional Elaboration of a Quasi-Ethnic Community. *International Review of Modern Sociology* 9(2): 165-177.
Musterd S., Deurloo M.C. (2002). Instable immigrant concentrations in Amsterdam; Spatial Segregation and Integration of Newcomers. *Housing Studies* 17(3): 487-503.
Office for Foreigners (2004). *Rejestr Urzędu do Spraw Cudzoziemców obejmujący cudzoziemców posiadających ważne karty pobytu w dniu 1 wrześNia 2004 [Office for Foreigners' register with foreigners possesing valid stay cards In Poland, as of 1 September 2004]*. Warsaw: OfF.
Office for Foreigners (2010). *Dane liczbowe dotyczące postępowań prowadzonych wobec cudzoziemców za lata 2007-2009. Tabela 70. Liczba osób, które posiadają ważne karty pobytu (stan na 31.12.2009) [Quantitative data of procedures concerning foreigners in the years 2007-2009. Table 70. Number of persons with valid stay cards in Poland (as of 31 December 2009)]*. Accessed 28 December 2010, online at: http://www.udsc.gov.pl/files/statystyki/biuletyn/biul_2007_2009_pol.xls.
Office for National Statistics (2010a). *Population by Country of Birth & Nationality, Apr 2009 to Mar 2010*. Office for National Statistics, United Kingdom. Accessed 28 December 2010, online at: http://www.statistics.gov.uk/downloads/theme_population/populationbybirthcountrynationalityapr09mar10.zip.
Okólski M. (2010). General Introduction. In: A. Górny, I. Grabowska-Lusińska, M. Lesińska, M. Okólski (eds.). *Immigration to Poland: Policy Labour Market Integration*. Warszawa: Scholar, pp. 17-53.
ONS (2004). *Census 2001 Definitions*. Office for National Statistics, General Register Office for Scotland, Northern Ireland Statistics and Research Agency, United Kingdom. Accessed 28 December 2010, online at: http://www.statistics.gov.uk/downloads/census2001/definitions_chapters_1_5.pdf.
ONS (2009). *The 2011 Census: Final questionnaire content for England and Wales*. Accessed 28 December 2010, online at: http://www.ons.gov.uk/census/2011census/2011censusquestionnairecontent/questionandcontentrecommendationsfor2011/index.html.
ONS (2010b). *Measuring Sexual Identity: An Evaluation Report*. September 2010. Authors: T. Joloza, J. Evans, R. O'Brien, Office for National Statistics. Accessed 28 December 2010, online at: http://www.statistics.gov.uk/articles/nojournal/measuringsexualidentityreport.pdf.
Orford S., Dorling D., Mitchell R., Shaw M., Davey-Smith G. (2002). Life and Death of the People of London: A Historical GIS of Charles Booth's Inquiry. *Health and Place* 8(1): 25-35.
Pain R., Mowl G., Talbot C. (2000). Difference and the negotiation of 'old age'. *Environment and Planning D: Society and Space* 18(3): 377-394.
Parekh, A., MacInnes, T., Kenway, P. (2010) *Monitoring poverty and social exclusion 2010*. Accessed 3 January 2011, online at: http://www.jrf.org.uk/sites/files/jrf/povertysocialexclusion2010full.pdf.
Park R., Burgess E.W. (1921). *Introduction to the Science of Sociology*. Chicago: The University of Chicago Press.
Park R., Burgess E.W., McKenzie R.D. (1925). *The City*. Chicago: University of Chicago Press.
Peach C. (1996a). Does Britain Have Ghettos? *Transactions of the Institute of British Geographers, New Series* 21(1): 216-235.
Peach C. (1996b). The meaning of segregation. *Planning Practice and Research* 11(2): 137-150.
Peach C. (2009). Slippery Segregation: Discovering or Manufacturing Ghettos? *Journal of Ethnic and Migration Studies* 35(9): 1381-1395.
Peake L. (1993). 'Race' and sexuality: challenging the patriarchal structuring of urban social space. *Environment and Planning D: Society and Space* 11(4): 415-432.
Phillips D. (1998). Black minority Ethnic Concentration, Segregation and Dispersal in Britain. *Urban Studies* 35(10): 1681-1702.
Phillips D. (2006). Parallel lives? Challenging discourses of British Muslim self-segregation. *Environment and Planning D: Society and Space* 24(1): 25-40.
Phillips D., Stillwell J. and Burrage A. (2004). Multicultural Leeds: geographies of ethnic minorities and religious groups. In: Unsworth R., Stillwell J. (eds.). *Twenty-First Century Leeds: Geographies of a Regional City*. Leeds: Leeds University Press, pp.49-74.
Piekut A. (2010). Preference for cities: spatial distribution of settled migrants. In: A. Górny, I. Grabowska-Lusińska, M. Lesińska, M. Okólski (eds.). *Immigration to Poland: Policy Labour Market Integration*. Warszawa: Scholar, pp. 156-160.
Potrykowska A., Śleszyński P. (2001). Mobilność przestrzenna ludności w aglomeracji warszawskiej (1990-98) [Spatial mobility of population in Warsaw Agglomeration]. In: I. Jażdżewska (ed.). *Miasto*

postsocjalistyczne – organizacja przestrzeni miejskiej i jej przemiany (część II) [Postsocialist city – urban space organization and its changes (part II)]. XIV Konwersatorium Wiedzy o Mieście, Katedra Geografii Miast i Turyzmu UŁ, Komisja Geografii Osadnictwa i LudnoŚci PTG, Łódź, pp. 63-76.
Poulsen M.F. (2009). Segregation. In: Kitchen R., Thrift N. (eds.) The International Encyclopedia of Human Geography. Oxford: Elsevier, pp. 63-69.
Pratt G., Hanson S. (1988). Gender, class, and space. *Environment and Planning D: Society and Space* 6(1): 15-35.
Putnam R. (2007). E Pluribus Unum: Diversity and Community in the Twenty-first Century. The 2006 Johan Skytte Prize Lecture. *Scandinavian Political Studies* 30(2): 137-174.
Reardon S.F., Firebaugh G. (2002). Measures of Multigroup Segregation. *Sociological Methodology* 32: 33-67.
Rees P. (1970). The urban envelope: patterns and dynamics of population density. In: B. Berry, F. Horton (eds.). *Geographic perspectives on urban systems*. Prentice Hall, Englewood Cliffs, New Jersey, pp. 276-305.
Rees P. (1979). *Residential Patterns in American Chicago: 1960*. Chicago: The University of Chicago, Department of Geography, Research Paper No. 189.
Rees P., Martin D., Williamson P. (eds.). (2002). *The Census Data System*. Chichester: John Wiley and Sons.
Rees P., Stillwell, J.C.H., Tyler-Jones A. (2004). The city is the people: demographic structure and dynamics. In: R. Unsworth, J. Stillwell (eds.). *Twenty-First Century Leeds: Geographies of a Regional City*. Leeds: Leeds University Press, pp.26-48.
Rees P., Wohland P., Norman P. (2009). The estimation of mortality for ethnic groups at local scale within the United Kingdom. *Social Science and Medicine* 69: 1592-1607.
Robson B.T. (1969). *Urban Analysis*. Cambridge: Cambridge University Press.
Rossi P.H. (1955). *Why Families Move: A Study in the Social Psychology of Urban Residential Mobility*. Glencoe, Ill.: The Free Press.
Rushbrook D. (2002). Cities, queer space, and the cosmopolitan tourist. *GLQ: A Journal of Lesbian and Gay Studies* 8(12): 183-206.
Sansom C.J. (2003). *Dark Fire*. London: Pan Books.
SASI (2010). *International Immigrants. Map No 15*. Worldmapper project. Social and Spatial Inequalities Research Group, Department of Geography, University of Sheffield. Accessed 28 December 2010, online at**:** http://www.worldmapper.org/display.php?selected=15**.**
Scharf T., Phillipson C., Smith A.E. (2003). Older people's perceptions of the neighbourhood: Evidence from socially deprived urban areas. *Sociological Research Online* 8(4). Accessed 28 December 2010, online at: http://www.socresonline.org.uk/8/4/scharf.html.
Seebohm Rowntree J. (1901). *Poverty: A Study of Town Life*. London: Macmillan.
Senior M. (2002) Deprivation indicators. In: Rees P., Martin D., Williamson P. (eds.). *The Census Data System*. Chichester: John Wiley and Sons, pp. 127-137.
Shannon C.E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal* 27(3-4): 379-423.
Shevky E., Bell W. (1955). *Social Area Analysis: Theory, Illustrative Application and Computational Procedures*. Stanford: Stanford University Press.
Shevky E., Lewin M. (1949). *Your Neighborhood*. Los Angeles: University of California Press.
Shevky E., Williams M. (1949). *The Social Areas of Los Angeles: Analysis and Typology*. Los Angeles: University of California Press.
Simmel G. (1908). The Stranger. In: Kurt Wolff (Trans.), (1950). *The Sociology of Georg Simmel*. New York: Free Press, pp. 402 - 408.
Simpson E.H. (1949). Measurement of diversity. *Nature* 163: 688.
Śleszyński P. (2004). Demograficzne przesłanki rozwoju rynku mieszkaniowego w aglomeracji warszawskiej [Demographic premises of housing market development in Warsaw agglomeration]. *Przegląd Geograficzny* 76(4): 493-514.
Southern Standard (2010). Chinese shellfish pickers are at risk. 5 August 2010. Accessed 28 December 2010, online at: http://www.southendstandard.co.uk/news/8311655.Chinese_shellfish_pickers_are_at_risk/.
Stillwell J., Phillips D. (2006) Diversity and change: understanding the ethnic geographies of Leeds. *Journal of Ethnic and Migration Studies* 32(7): 1131-1152.
Stillwell J.C.H., Shepherd P. (2004). The 'haves' and 'have-nots': contrasting social geographies. In: R. Unsworth, J. Stillwell (eds.). *Twenty-First Century Leeds: Geographies of a Regional City*. Leeds: Leeds University Press, pp. 127-146.
Taeuber K.E., Taeuber A.F. (1964). The negro as an immigrant group. *American Journal of Sociology* 69(4): 374-382.
Taeuber K.E., Taeuber A.F. (1965). *Negroes in Cities: Residential Segregation and Neighbourhood Change*. Chicago: Aldine.
Tajfel H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology* 33: 1-39.

Takahashi L.M. (1997). Representation, attitudes, and behavior: analyzing the spatial dimensions of community response to mental disability. *Environment and Planning A* 29(3): 501-524.
Talen E. (2009). The Context of Diversity: A Study of Six Chicago Neighbourhoods. *Urban Studies* 47(3): 486-513.
Timms D. (1970). *The Urban Mosaic*. Cambridge: Cambridge University Press.
Triandafyllidou A., Gropas R. (2009). Constructing Difference: The Mosque Debates in Greece. *Journal of Ethnic and Migration Studies* 3(6): 957-975.
UN (1998). *Recommendation on Statistics of International Migration.* New York.
UN (2009). *United Nations' Trends in Total Migrant Stock: The 2008 Revision*. United Nations Population Division. Accessed 28 December 2010, online at: http://esa.un.org/migration.
Unsworth R., Stillwell J. (2004). *Twenty-first century Leeds: geographies of a regional city*. Leeds: Leeds University Press.
Valentine G. (1993). (Hetero)sexing space: lesbian perceptions and experiences of everyday spaces. *Environment and Planning D: Society and Space* 11(4): 395-413.
Valentine G. (2008). Living with difference: reflections on geographies of encounter. *Progress in Human Geography* 32: 321-335.
Valentine G., Skelton T. (2003). Finding oneself, losing oneself: the lesbian and gay 'scene' as a paradoxical space. *International Journal of Urban and Regional Research* 27(4): 849-866.
Valins O. (2003). Stubborn identities and the construction of socio-spatial boundaries: ultraorthodox Jews living in contemporary Britain. *Transactions of the Institute of British Geographers* 28(2): 158-175.
Van Kempen R., Van Weesep J. (1998). Ethnic Residential Patterns in Dutch Cities: Backgrounds, Shifts and Consequences. *Urban Studies* 35(10): 1813-1833.
Vanderbeck R.M. (2007). Intergenerational Geographies: Age Relations, Segregation and Reengagements. *Geography Compass* 1(2): 200-221.
Varady D. (2008). Muslim Residential Clustering and Political Radicalism. *Housing Studies* 23 (1): 45-66.
Vertovec S. (2006). The Emergence of Super-Diversity in Britain. Centre on Migration, Policy and Society, Working Paper No. 25, University of Oxford. Accessed 28 December 2010, online at: http://www.compas.ox.ac.uk/fileadmin/files/pdfs/Steven%20Vertovec%20WP0625.pdf.
Vertovec S. (2007). Super-diversity and its implications. *Ethnic and Racial Studies* 29(6): 1024-54.
Vertovec S. (2009). Conceiving and Researching Diversity. Working Papers 09-01. Göttingen: Max Planck Institute for the Study of Religious and Ethnic Diversity. . Accessed 28 December 2010, online at: http://www.mmg.mpg.de/fileadmin/user_upload/documents/wp/WP_09-01_Vertovec_Diversity.pdf.
Vertovec S. (2010). Towards post-multiculturalism? Changing communities, conditions and contexts of diversity. *International Social Science Journal* 199: 83-95.
Vertovec S., Wessendorf S. (2004). Migration and Cultural, Religious and Linguistic Diversity in Europe: An overview of issues and trends. Working Paper 1. Centre on Migration, Policy and Society [COMPAS] University of Oxford. Accessed 28 December 2010, online at: http://dare.uva.nl/document/39845.
Vickers D., Rees P. (2006). Introducing the National Classification of Census Output Areas. *Population Trends* 125: 15-29.
Vickers D., Rees P. (2007). Creating the UK National Statistics 2001 output area classification. *Journal of the Royal Statistical Society, Series A* 170(2): 379-403.
Vickers D., Rees P., Birkin M. (2003). New classification of UK Local Authorities using 2001 Census Key Statistics. Working Paper 03/3. School of Geography, University of Leeds. Accessed 28 December 2010, online at: http://www.geog.leeds.ac.uk/wpapers/033.pdf.
Vickers D., Rees P., Birkin M. (2005). Creating the national classification of census output areas: data, methods and results. Working Paper 05/2. School of Geography, University of Leeds. Accessed 28 December 2010, online at: http://www.geog.leeds.ac.uk/wpapers/052.pdf.
Ward J.H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58: 236-244.
Webber R., Craig J. (1978). *Socioeconomic Classifications of Local Authority Areas.* Studies in Medical and Population Subjects 35. London: Office of Population, Censuses and Surveys.
Węcławowicz G. (1988a). *Struktury społeczno-przestrzenne w miastach Polski (Socio-spatial structures in Polish cities)*. Wrocław: Ossolineum.
Węcławowicz G. (1988b). Elderly people in the socio-spatial structure of some Polish towns. *Geographia Polonica* 54:39-149.
Węcławowicz G. (1979). The structure of socio-economic space in Warsaw in 1931 and 1970: a study in factorial ecology. In: French R.A., Hamilton F.E.I. (eds.). *The Socialist City, Spatial Structure and Urban Policy*. Chichester: John Wiley & Sons, pp. 387-424.
Węcławowicz G. (1991). Zróżnicowania społeczno-przestrzenne w aglomeracji warszawskiej (1978 i 1988) [Socio-spatial Diversity in Warsaw Agglomeration (in 1978 and 1988)]. *Zeszyty IGiPZ PAN* 2.

Węcławowicz G. (2005). The Warsaw Metropolitan Area on the eve of Poland's integration into the European Union. In: F.E.I. Hamilton, K. Dimitrovska Andrews, N.Pichler-Milanoviè (eds.). *Transformation of Cities in Central and Eastern Europe – Towards Globalization*. United Nations University Press, Tokyo, New York, Paris.

Węcławowicz G., Bański J. Degórski M., Komornicki T., Korcelli P., Śleszyński P. (2006). Spatial organization of Poland at the beginning of the 21st century. Monographies, 6. Warsaw: Institute of Geography and Spatial Organization PAS.

Węcławowicz G., Kozłowski S., Baje R. (2003). *Large Housing Estates in Poland. Overview of development and problems in Warsaw*. RESTATE report (Restructuring Large – scale Housing Estates in European Cities: Good Practices and New Visions for Sustainable Neighbourhoods and Cities). Faculty of Geosciences, Utrecht University.

White M.J. (1983). The measurement of spatial segregation. *American Journal of Sociology* 88(5): 1008-1018.

White M.J. (1986). Segregation and diversity measure in population distribution. *Population Index* 52(2): 198-221.

Wilton R. D. (2004). From flexibility to accommodation? Disabled people and the reinvention of paid work. *Transactions of the Institute of British Geographers* 29(4): 420-432.

Wirth L. (1938). Urbanism as a way of life: the city and contemporary civilization. *American Journal of Sociology* 44(1):1-24.

Wohland P., Rees P., Norman P., Boden P., Jasińska M. (2010). Ethnic Population Projections for the UK and Local Areas, 20012051. School of Geography University of Leeds. Working Paper 10/2. Accessed 28 December 2010, online at: http://www.geog.leeds.ac.uk/fileadmin/downloads/school/research/wpapers/10_1.pdf.

Żarnowska A. (1985). *Robotnicy Warszawy na przelomie XiX i XX wieku [Workers in Warsaw at the turn of 19th and 20th century]*. Warsaw: PIW.

APPENDIX A1. COMPARISON OF VARIABLES FROM THE BRITISH AND POLISH CENSUSES

Table A1.1. Variables in the family/demographic dimension of diversity

Characteristics	UK Census – definitions and variables	Polish Census – definitions and variables	Comments
Base	Usual resident population	Population de facto living in Poland Resident population Temporal residents	Different
Age	Age is derived from the date of birth question and is the age at a person's last birthday.	Age was calculated by comparison of full date of birth with Census Day 20th May 2002	=Compatible
Age groups / Generations	**Age groups**: 0-4, 5-7, 8-9, 10-14, 15, 16-17, 18-19, 20-24, 25-29, 30-44, 45-59, 60-64, 65-74, 75-84, 85-89, 90 and over	**Age groups:** 0-2, 3-6, 7-12, 13-15, 16-18, 19-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85-89, 90-99, 100 and more **Economic age groups** Pre-productive (0-17) Productive Mobile (18-44) Not mobile (45-64 men, 45-59 women) Post-productive (men over 65, women over 60) Not stated **Biological age groups** 0-14 15-64 65 and more Not stated **Educational age groups**: 0–6, 7–12, 13–15, 16–18, 19–24 Data were divided into **three generations**: Young people – to 29 years old Middle-aged people – 30-59 years Older people – 60 and more years	The age groups that overlap are: 0-15, 16-24, 25-29, 30-44, 45-59, 60-64, 65-74, 75-84, 85 and over Population can be also divided into three generations: Young people – to 29 years old Middle-aged people – 30-59 years Older people – 60 and more years
Household	A household comprises one person living alone, or a group of people (not necessarily related) living at the same address with common housekeeping that is, sharing either a living room or sitting room or at least one meal a day.	A household comprises a group of people, related or not related, who live together and share expenses. If one of the people living together has his/hers own expenses he/she comprises separate household.	Slightly different definitions. While in the UK sharing a common living space is stressed, in Poland – sharing the expenses.
Household composition	One person 01 Pensionable age 02 Other One family and no others 03 All people of pensionable age Married couple family 04 no children 05 with one dependent child 06 with two or more dependent children 07 all children nondependent Cohabiting couple family 08 no children 09 with one dependent child 10 with two or more dependent children 11 all children nondependent Lone parent family with male head 12 with one dependent child 13 with two or more dependent children 14 all children nondependent with female head 15 with one dependent child 16 with two or more dependent children 17 all children nondependent Other household types 18 with one dependent child 19 with two or more dependent children 20 all in fulltime education 21 all of pensionable age	Household composition was fixed/settled on a base of relation of each person to a head of the household and some additional information (date of marriage, information on parenthood, partnership) – number and types of families. On a base of relation of a head of household there were distinguished following types of households: Family households (one-family, two-family, three or more family) One family without others marriages couples mothers with children fathers with children with relatives in a direct line from older generation marriages couples mothers with children fathers with children with other people marriages couples mothers with children	Different. But could be divided into: One person household Family household Not-family household Probably we could select lone parent families.

	22 other X not applicable	fathers with children Two family Three family and more Nonfamily households: One person Multi-person	
Household composition by generation	No variable	Household composition by generation: only young people only middle-aged people only older people young and middle-aged people young people and older people middle-aged people and older people young and middle-aged people and older people Not stated	Only in Polish census
Other household variables	Household size Alternative household composition Alternative household type Household Reference Person country of birth Household Reference Person ethnic group Household Reference Person NSSeC Household Reference Person social grade	Households size / by number of people (1, 2, 3, 4, 5 ,6, 7 and more)	Households sizes can be compared
Marital status	Legal marital status as at Census day 1 Single (never married) 2 Married (first marriage) 3 Remarried 4 Separated (but still legally married) 5 Divorced 6 Widowed	Legal marital status Marital status according to national law, for all people aged 15 and more. Measured for the first time. Categories: single married windowed divorced separated (legal separation was introduced in Poland in 1999).	Slightly different. Married and remarried category in the UK would have to be merged.
Family definition	A family comprises a group of people consisting of a married or cohabiting couple with or without child(ren), or a lone parent with child(ren). It also includes a married or cohabiting couple with their grandchild(ren) or a lone grandparent with his or her grandchild(ren) where there are no children in the intervening generation in the household. Cohabiting couples include same sex couples. Children in couple families need not belong to both members of the couple. (In 2001, the definition of a family has been extended to include a cohabiting couple of the same sex with or without child(ren)).	Biological family – two or more people who are married or live together as a couple (cohabitate) – that are opposite sex couples or live together as a parent and a child. The family includes couple without children or a couple with one or more children, as well as lone parent with one or more children. (In 2002 for the first time cohabitating couples were distinguished as a separate families. Previously they were included in marriages)	The basic difference is that same sex couples were not counted as couples. Samesex couples are excluded from the census in Poland.
Family status	1 Not in a family – of pensionable age 2 Not in a family – other 3 In a couple family – member of couple 4 In a couple family – dependent child of one or both members of the couple 5 In a couple family –non dependent child of one or both members of the couple 6 In lone parent family – parent 7 In lone parent family – dependent child of parent 8 In lone parent family – non dependent child of parent 9 Not in a household	01 head of a household 02 husband – wife 03 partner (cohabitant) 04 son – daughter 05 mother – father 06 mother-in-law – father-in-law 07 son-in-law – daughter-in-law 08 brother – sister 09 grandparent, grand-grandparent 10 grandchild 11 other relative 12 unrelated person	Different classifications
Family composition	This variable is derived from the algorithm that allocates people to families. It classifies families according to the characteristics of the people in generation one. Cohabiting couple families include both opposite sex and same sex couples Variables:	Families without children Marriages Couples Families with dependent children under 24 (by number of children) Variables: Marriages with child/children	Similar categories.

	0 Ungrouped individual 1 Married couple family 2,3,4 Cohabiting couple family 5 A lone parent family (male) 6 A lone parent family (female) (Two people are described as cohabiting if they are living together as a couple but are not married to each other. This includes people living with a partner of the same sex. A cohabiting person might be married (to someone not resident in the household) but will not be shown as married or separated in the living arrangements tables.)	Couples with child/children Mothers with child/children Fathers with child/children	
Generation in family	0 A person not in a family (an 'ungrouped individual') 1 Generation 1: a member of a couple or a lone parent 2 Generation 2: a child (any age) in a family X Not applicable	Against order.	Data against order.
Living arrangements	This derived variable is created by combining the responses to the question on legal *Marital status* MSTPUK, and the responses to the relationship question. This is an ONS harmonised survey classification. Living in a couple 1 Married (including separated) 2 Remarried 3 Cohabiting Not living in a couple 4 Single 5 Married or remarried 6 Separated 7 Divorced 8 Widowed X Not applicable (without people in communal establishments)	Actual/De facto marital status It was defined on the basis of character of relationship a given person e.g. relation to a head of a household, legal marital status, and mutual relations among recorded people. Categories: single married (according to law and being *de facto* in marriage) cohabiting (regardless of legal marital status) windowed divorced separated (legal separation and not being in partnership and people legally married, but not being *de facto* in marriage).	Different

Table A1.2. Variables in the socioeconomic dimension of diversity

Characteristics	UK Census – definitions and variables	Polish Census – definitions and variables	Comments
Base	Usual resident population	Population de facto living in Poland Resident population Temporal residents	Different
Economically active definition	All people who were working in the week before the Census are described as economically active. In addition, the category includes people who were not working but were looking for work and were available to start work within 2 weeks. Full time students who are economically active are included but are identified separately in the classification. The economic activity questions are only asked of people aged 16 – 74.	As economic active population (e.g. labour force) includes all people recognised as employed or unemployed. The economic activity questions are only asked of people over 15 years old.	=Compatible
Employed definition	Any person who did paid work in the week before the Census, whether as an employee or self-employed, is described as employed or in employment. 'Paid work' includes casual or temporary work, even if only for one hour; being on a government-sponsored training scheme; being away from a job/business ill, on maternity leave, on holiday or temporarily laid off; or doing paid or unpaid work for their own or family business.	Persons aged 15 and more, who during the census week: performed for at least one hour any work generating pay or income, i.e. were employed as paid employees, worked on their own (or leased) agricultural farm, or conducted their own economic activity outside agriculture, assisted (without pay) in work on family agricultural farm or in conducting family economic activity outside agriculture, had work but did not perform it: a) due to sickness, maternity leave or vacation, b) due to other reasons, but the break in employment: did not exceed 3 months, exceeded 3 months, but these persons were paid employees and during that period received at least 50% of the hitherto remuneration. Employed persons in line with international standards also include students, with whom employers or physical persons concluded a contract for apprenticeship or teaching a particular job, if they receive wages or salary. The term "any work" shall be understood as work performed for at least 1 hour.	=Compatible In UK – only 'paid work' and unpaid work in case of own or family business, week before census In Poland – work that generates income, even unpaid income, census week
Unemployed definition	A person is defined as unemployed if he or she is not in employment, is available to start work in the next 2 weeks and has either looked for work in the last 4 weeks or is waiting to start a new job. This is consistent with the International Labour Office (ILO) standard classification.	Persons aged 15-74 who simultaneously fulfil three conditions: during the reference week were not employed, were actively seeking work, i.e. they undertook specific actions during 4 weeks (including the reference week as the last one) to find work; were ready (capable) to take up employment in the reference week or the following week. Persons who weren't seeking work, because they had an arranged employment and waited for its beginning for time not longer than 3 months were also counted as unemployed.	=Compatible
Main job definition	The main job is the job in which a person usually works the most hours. Questions on employment relate to each person's main job.	In case of persons carrying out more than one job main employment refers to that job which normally takes most time. If jobs take the same amount of time, the main job is the one that job that generates higher income.	=Compatible

Occupation classification	A person's occupation is coded from the response to the question asking for the full title of the *Main job* and the description of what is done in that job. It is coded to the 2000 edition of the Standard Occupational Classification (SOC). (SOC is aligned as far as possible with the international classification ISCO 88). 1 Managers and senior officials 2 Professional occupations 3 Associate professional and technical occupations 4 Administrative and secretarial occupations 5 Skilled trades occupations 6 Personal service occupations 7 Sales and customer service occupations 8 Process, plant and machine operatives 9 Elementary occupations	A person's occupation is coded to the International Standard Classification of Occupations ISCO1988 (ISCO1988) Group 1 Legislators, senior officials and managers Group 2 Professionals Group 3 Technicians and associate professionals Group 4 Clerks Group 5 Service workers and shop and market sales workers Group 6 Skilled agricultural and fishery workers Group 7 Craft and related trade workers Group 8 Plant and machine operators and assemblers Group 9 Elementary occupations Group 0 Armed forces	=Compatible Different classifications, but compatible categories
Employment type	1 Employee 2 Selfemployed with employees 3 Selfemployed without employees X Not applicable	International Classification of Status in Employment (ICSE) distinguishes the following categories of working people: **Paid employment jobs** – jobs where the employees hold employment contracts at public or provide entrepreneurship, includes civic contracts as well as **employers –** are those workers who, working on their own account (or with one or a few partners), hold the type of job defined as a "self-employment job" and, in this capacity, on a continuous basis (including the reference period) have engaged one or more persons to work for them in their business as "employee(s)" (Fulltime and part-time) **own-account workers** – are those workers who, working on their own account or with one or more partners, hold the type of job defined as a "self-employment job" (cf. paragraph 7), and have not engaged on a continuous basis any "employees" (cf. paragraph 8) to work for them during the reference period. The category includes agents (in all types of agencies) as well as members of agricultural cooperative societies **contributing family workers** – person that helps (without payment) in a market-oriented establishment operated by a related person living in the same household	=Compatible Employee Employers / Self-employed with workers Self-employed without workers Others
Source of livelihood	No variable.	**Source of livelihood** (sourced for consumption basic needs for the last 12 month before the census). Variables: Income (earning) from employment Unpaid income Other incomes (from properties, rents) Being dependent on other people that have income	Different. No data on source of livelihood in the UK census
Socioeconomic position	The National Statistics Socioeconomic Classification (NSSeC) provides an indication of socioeconomic position, based on occupation. It is an Office for National Statistics standard classification. Individual level	Information on (main) livelihood sources were used to classify households into socioeconomic groups: 1. households of workers 2. households of workers possessing a farm 3. households of farmers	Different categories and different level of data aggregation. UK – individual level / based on occupation Poland – household level / based on livelihood source

	Class1 (Higher managerial and professional occupations) Class 2 (Lower managerial and professional occupations) Class 3 (Intermediate occupations) Class 4 (Small employers and own account workers) Class 5 (Lower supervisory and technical occupations) Class 6 (Semi-routine occupations) Class 7 (Routine occupations) Class 8 (Never worked and long-term unemployed) Not classified Occupation not coded 39 XX Not applicable	4. households of own-account workers 5. households of old age pensioners 6. households of pensioners 7. households providing by nonpaid sources of livelihood 8. households providing by property incomes 9. other households	(Farming and non-farming household – not appropriate in case of Warsaw)
Other variables related to employment	Workers in generation one of family Adults in employment, number in Household Employment selected characteristics, household indicator (Not deprived / Deprived in employment dimension) Education selected characteristics, household indicator (Not deprived / Deprived in education dimension)	No more variables.	No data in the Polish census
Industry of employment	A Agriculture, hunting and forestry B Fishing C Mining and quarrying D Manufacturing E Electricity, gas and water supply F Construction G Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods H Hotels and restaurants I Transport, storage and communication J Financial intermediation K Real estate, renting and business activities L Public administration and defence; compulsory social security M Education N Health and social work O Other 'Other' includes other community, social and personal service activities; private households with employed persons; and extraterritorial organisations and bodies.	Data on industry of employment were coded according to Statistical Classification of Economic Activities in the European Community, Rev. 1.1 (2002) by EUROSTAT A Agriculture, hunting and forestry B Fishing C Mining and quarrying D Manufacturing E Electricity, gas and water supply F Construction G Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods H Hotels and restaurants I Transport, storage and communication J Financial intermediation K Real estate, renting and business activities L Public administration and defence; compulsory social security M Education N Health and social work O Other community, social and personal service activities P Activities of households Q Extraterritorial organizations and bodies	=Compatible
Level of qualification / education	Level of qualifications This derived variable provides information on the highest level of qualification. In England and Wales it is based on responses to both the qualifications question and the professional qualification question QU(x)PEW and PQU(x)PEW. Levels: No academic or professional Qualifications Level 1: CSEs (grades 25), GCSEs (grades DG), 14 CSEs (grade 1), 14 GCSEs (grades AC), 14 O levels, NVQ level 1, Foundation GNVQ	Level of education Higher with PhD degree at least with masters degree, doctor in medicine or similar with engineer degree, bachelors Postsecondary high school exams, college of further education without high school exams Secondary Vocational with high school exam Vocational without high school exam Comprehensive with high school exam Comprehensive without high school exam Basic Vocational	Different. UK –qualifications PL – education only

	Level 2: 5+ O levels, 5+ CSEs (grade1), 5+GCSEs (grades AC) etc, 1 'A' level, 13 AS levels, NVQ level 2, Intermediate GNVQ Level 3: 2+ A levels, 4+ AS levels, Higher School Certificate, NVQ level 3, Advanced GNVQ Level 4/5: First degree, Higher degree, NVQ levels 45, HNC, HND. Qualified Teacher status, Qualified Medical Doctor, Qualified Dentist, Qualified Nurse, Midwife, Health Visitor Other qualifications/ level unknown: Other qualifications (eg City and Guilds etc), Other Professional qualifications XX Not applicable	Elementary Elementary unfinished or no education Not stated	
Household space	All household spaces With residents With no residents Vacant Second residence/holiday accommodation	All households With residents Type of a building: Residential/housing Residential stocktaking Communal accommodation Non-residential Holiday accommodation	=Compatible Households with residents Households with no residents
Households variables	Average household size Average number of rooms per household	Average number of persons per apartment Average number of households living in one apartment	Different
Accommodation type	All household spaces which are of accommodation type Whole house or bungalow Detached Semidetached Terraced (including endterrace) Flat, maisonette or apartment Purpose built block of flats or tenement Part of a converted or shared house (including bedsits) In a commercial building Caravan or other mobile or temporary structure	No variable.	No data in the Polish census.
Building age	No variable.	Before 1918 r. 1918-1944 1945-1970 1971-1978 1979-1988 1989-2002 In construction	No data in the UK census
Tenure	Owner occupied Owns outright Owns with a mortgage or loan Shared ownership Rented from Council (local authority) Housing Association/Registered Social Landlord Private landlord or letting agency Other	Privately owned building flat Privately owned (within a housing cooperative) Rented from Landlord housing association Community (local authority) Treasury Working place Building Social Society / Cooperative flats Other Sub-rented Relationship Other Non stated	=Compatible UK –census ask "Does your household own or rent the accommodation?" PL – census ask „On what legal grounds you live in the apartment?" Privately owned Rented + subrented Other

Table A1.3. Variables in the ethnic dimension of diversity

Characteristics	UK Census – definitions and variables	Polish Census – definitions and variables	Comments
Base	Usual resident population	Population de facto living in Poland Resident population Temporal residents	Different
Immigrant	Migrant A migrant is a person with a different address one year before the Census to that on Census day. Immigrant – a resident in the area on Census day but whose usual address one year before Census was outside the area	Migrant – changing ones place of residence (permanent or temporal) connected with crossing an administration border (the smallest registry units are communes *gminas*) for longer than 2 months **Internal migration**: Population *de facto* living in Poland People living in a given place from birth People that arrived to a place of residence before 1989 People that arrived to a place of residence in 1989–2002 Resident population People living in a given place from birth People that arrived to a place of residence before 1989 People that arrived to a place of residence in 1989–2002 Immigrants temporarily staying in a given place (that were not included in population de facto living in Poland) over 2 months): Short-term immigrants (2-12 months): Immigrants staying in a given place for at least 2 months up to 6 months Immigrants staying in a given place for at least 6 months up to one year Long-term immigrants (>12 months) **International migration** Population de facto living in Poland who arrived in Poland before 1988 (by year, country of previous residence and citizenship) who arrived in Poland in 1989–2002 (by year, country of previous residence and citizenship) Resident population who arrived in Poland before 1988 (by year, country of previous residence and citizenship) who arrived in Poland in 1989–2002 (by year, country of birth, country of previous residence and citizenship) Immigrants temporarily staying in Poland (that were not included in population de facto living in Poland) over 2 months (by citizenship, country of previous residence, country of birth): Short-term immigrants (212 months): Immigrants staying in Poland for at least 2 months up to 6 months Immigrants staying in Poland for at least 6 months up to one year Long-term immigrants (>12 months)	In Poland much more migration categories depending on a time a person is living in a given place. In UK – census asks about the usual address 12 months ago. But the person could move into this place 13 months ago, 2 years or more. UK migrant = +/ PL short-term migrants 612 months In case of international migration: Person with migrant origin outside UK = Long-term foreign immigrants
Foreigner	Country of birth, ONS Geography	Foreigner every person without Polish citizenship. Foreigners that were permanent residents were included in population de facto living in Poland. Other were included in resident population (long-term immigrants, living over 12 month in Poland) or in temporal residents (short-term immigrants, living under 12 months in Poland).	In Poland, foreign population is distinguished on a basis of citizenship. Census contains country of birth variable, but it overestimates 'foreign' population. For example, 70 per cent of all foreign-born arrived from three countries: Belarus, Lithuania and Ukraine, and, secondly, 82 per cent of those persons were 60 years of

			age or older. People who were born in pre-war Polish territory, after the war there were displaced to Poland, because that land was annexed to the Soviet Union.
Ethnic / national group	The ethnic group question records each person's perceived ethnic group and cultural background. 16 main categories: White British Irish Any other White background Mixed White and Black Caribbean White and Black African White and Asian Any other Mixed background Asian or Asian British Indian Pakistani Bangladeshi Any other Asian background Black or Black British Caribbean African Any other Black background Chinese or other ethnic group Chinese Any other ethnic group Not stated	The national group – perceived/declaration Polish NonPolish – 99 categories Not stated	Ethnic vs. National (=ethnic) In Polish census national group are not divided into categories. To measure diversity indices there should be taken the same number of categories in UK and Poland. High per cent of non response in the Polish census, because only one answer on national identity was possible.
Multiple ethnic identifier (household)	This household classification provides an indicator of multiple ethnic identities within and between family generations. 1 One person household 2 All household members have the same ethnic group 3 Different identities between generations only 4 Different identities within partnerships (whether or not also different identities between generations) 5 Any other combination of multiple ethnic identities X Not applicable	No variable.	No data in the Polish census
Citizenship	No variable.	195 countries Stateless person Non stated	No data in the UK census
Religion	Christian Buddhist Hindu Jewish Muslim Sikh Other religions No religion Not stated	No variable.	In Polish census there was no question on religion
Language used at home	No variable.	Only Polish Polish and another Polish and 1 another Polish and 2 other Only another 1 another 2 other Not stated	No data in the UK census

Table A1.4. Variables in the disability and health dimension of diversity

Characteristics	UK Census – definitions and variables	Polish Census – definitions and variables	Comments
Base	Usual resident population	Population de facto living in Poland Resident population Temporal residents	Different.
General health	This question asks for an assessment of each person's general health over the last year. Variables: 1 Good 2 Fairly good 3 Not good	No variable.	In Polish census there was no question on general health
Disability	Limiting long-term illness, health problem or disability This variable records whether a person perceives that they have a limiting long-term illness, health problem or disability which limits their daily activities or the work they can do, including problems that are due to old age. Variables 1 Has limiting long-term illness, health problem or disability 2 No limiting long-term illness, health problem or disability	**Legal disability** Adults (aged over 16): with light degree of disability with light degree of disability with light degree of disability not stated Children (aged 015) **Biological disability** only, noticing limitation in ability: full serious	UK 'limiting long-term illness, health problem or disability' is a broader category than Polish 'legally' or 'biologically disable'
Health and disability selected characteristics, household indicator	A household has selected characteristics in the health and disability domain if any person in the household has general health 'not good' or has a limiting long term illness. Variables: 0 Not deprived in health and disability dimension 1 Deprived in health and disability dimension X Not applicable	Against order.	Probably, similar data could be ordered in Poland.

APPENDIX A2. INDEXES OF SEGREGATION AND DIVERSITY

A2.1. A notation

It is important to describe a set of related measures using common and consistent mathematical notations. Throughout this chapter we use the following variables:

- P – represents a count of people or a population
- R – is a probability that a population will be in a particular state (R refers to a row in the data matrix)
- C – is another probability that a population will be in a particular state (C refers to a column in the data matrix)

The other variables we attach superscripts or subscripts that define the particular state or states which the population is in:

- i – indicates (is an index for) a zone of residence (Community Areas in Leeds and Urban Region in Warsaw)
- j – also indicates a zone of residence, used if we need to refer to two types of zones in a measure
- e – represents (is an index for) a social group
- f – also represents a social group, when we refer to more than one group
- $+$ – is used to indicate, where it replaces an index, that the variable has been summed over that indicator.

It is useful to arrange these variables in the form of data matrix as showed in Figure A2.1.

Residential zone	Social groups: 1	...	e	...	M	Total
1	P_1^1	...	P_e^1	...	P_M^1	P_+^1
⋮	⋮		⋮		⋮	⋮
i	P_1^i	...	P_e^i	...	P_M^i	P_+^i
⋮	⋮		⋮		⋮	⋮
N	P_1^N	...	P_e^N	...	P_M^N	P_+^N
Total	P_1^+	...	P_e^+	...	P_M^+	P_+^+

Figure A2.1. A data matrix for residential zones and social groups

A2.2. Diversity and locational indexes for a single group

We frequently want to compute the concentration of a group in a zone and to compare it with the concentration of other social groups. To do this we compute the probability, R_e^i, that a person living in a zone is a member of a particular group:

$$R_e^i = \left(P_e^i \;\; P_+^i\right) \tag{A2.1}$$

where
P_e^i = population in zone i of social group e
$P_+^i = \sum_e P_e^i$ = sum of overall social groups e of the population in zone i, the total population of zone i.

Often we represent the probability as a percentage by multiplying by 100 to produce a more comprehensible number.

We can map the R measure and compare maps for different groups. This index shows where a group in concentrated or underrepresented. We can also define further measures of relative concentration such as location quotient, which is computed thus:

$$LQ_e^i = \left(R_e^i \;\; R_e^+\right) \tag{A2.2}$$

The location quotient assesses whether a social group is more or less concentrated in a zone than in the study area as a whole.

Also useful on occasion to compute how concentrated a group is in a zone relative to the land extent of that zone using the following index:

$$T_e^i = \left(C_e^i \;\; C_a^i\right) \tag{A2.3}$$

where T_e^i is the territorial concentration of group *e* in zone *i*, and C_e^i is given by

$$C_e^i = \left(P_e^i \;\; P_e^+\right) \tag{A2.4}$$

and

$$C_a^i = \left(A^i \;\; A^+\right) \tag{A2.5}$$

where A^i is the land area of zone *i* and A^+ is the land area of the study region.

A2.3. Indexes that compare two social groups

There is a long history of invention and application of indexes that compare the distribution of two or more social groups across a set of reference categories such as geographical units. The indexes have been a subject of detailed metareview about once a decade (Duncan, Duncan 1955; Goodman, Kruskal 1954, 1959; White 1983, 1986; Massey, Denton 1988; Peach 1996b, Peach 2009). The measures are summary indexes which compare the spatial distributions of two groups across the whole study area.

A2.3.1. The Dissimilarity Index

The Dissimilarity Index (extensively used by Taeuber and Taeuber 1965) is the most commonly used of these two group comparisons. This index is defined as follows:

$$D_{ef} = {}^1\!/_2 \; 100 \sum_{i=1}^{i=n} \left| C_e^i - C_f^i \right| \tag{A2.6}$$

where C_e^i is the probability that a person belonging to social group *e* is living in residential zone *i* (equation 4.5), and C_f^i is the equivalent variable for social group *f*. Substituting the RH side of equation (4.5) into the C variables in equation (4.6) we obtain:

$$D_{ef} = {}^1\!/_2 \; 100 \sum_{i=1}^{i=n} \left| \left(P_e^i \;\; P_e^+\right) - \left(P_f^i / P_f^+\right) \right| \tag{A2.7}$$

The constant of 100 'converts probabilities' into percentages and the ½ ensures the index ranges between 0 and 100, not 0 and 200. Absolute values of the difference between the two group percentages for a zones are summed for all zones in the study area. Note that we could use the differences to describe the separation of groups zone by zone, although this is rarely done.

The index measures the percentage of the first group that would need to move to have exactly the same geographic distribution (in relative terms) as the second group. The index is symmetric – that means we get the same index value even if we swap the two groups being compared. The Dissimilarity index can also be derived as the maximum vertical difference between the line of equality and the Lorenz curve in a Lorenz diagram (Lorenz 1905; Duncan, Duncan 1955; Massey, Denton 1988). A Lorenz curve is a curve joining points plotted in two dimensional space, with dimensions representing the cumulative percentage distributions of two variables. The line of equality is simply one special Lorenz curve, to which other curves can be compared (see http://mathworld.wolfram.com/LorenzCurve.html for demonstrations of the Lorenz curve).

A2.3.2. The Gini Coefficient

The Gini Index was invented by Italian statistician Corrado Gini (Gini 1912, 1921) and has been intensively used ever since to measure inequalities in income, where the differences between the distribution of households classified by decile and the income received by each household decile are measured.

We can define the method for computing the Gini Index as follows. Order the zones from the most concentrated to the least concentrated for group e such that $R_e^i < R_e^{i+1}$. Compute cumulative probabilities for the groups being compared:

$$X_e^k = \sum_{i=1}^{i=k} C_e^i \qquad \text{(A2.8)}$$

$$Y_e^k = \sum_{i=1}^{i=k} C_f^i \qquad \text{(A2.9)}$$

Then the discrete approximation of the Gini coefficient, approximated by trapezoids, is:

$$G_{ef} = 1 - \sum_{k=1}^{k=n} \left(X_e^k - X_e^{k-1}\right) \times \left(Y_f^k + Y_f^{k-1}\right) \qquad \text{(A2.10)}$$

The Gini coefficient is represented in the Lorenz diagram as the area between the Lorenz curve and the line of equality divided by the whole area underneath the line of equality. Unlike the Dissimilarity Index it is sensitive to all transfers between areas, not just between areas of under and overrepresentation. A scalar factor of 100 is not usually applied though we do so in the paper to compare values for our two cities against Dissimilarity Index values.

A2.3.3. The Exposure Index

Developed originally by Lieberson (1958), the Exposure Index measures the chances that members of one group have of being exposed to members of anther group. The formula is as follows:

$$E_{ef} = \sum_{i=1}^{i=n} \left(C_i^e \times R_i^f\right) \qquad \text{(A2.11)}$$

Substituting the RH sides of equation (4.4) and (4.1) respectively for C_e^i and , multiplying by 100 to obtain a percentage, we obtain:

$$E_{ef} = 100 \sum_{i=1}^{i=n} \left[\left(P_e^i \;\; P_e^+\right) \times \left(P_f^i \;\; P_+^i\right)\right] \qquad \text{(A2.12)}$$

The index is asymmetric, which means we obtain different results depending on which group is exposed to which. The converse of the Exposure Index is the isolation index, which measures the extent to which members of one group are exposed to members of the same group only.

A2.4. Indexes that measure the diversity of more than two social groups

The concept of diversity extends beyond just two groups. The more groups live in an area the more diverse is its population. The more evenly the population is spread across the groups, the more diverse we consider that population to be (Reardon, Firebaugh 2002).

A2.4.1. The Simpson/Diversity Index

Simpson (1949) proposed methods to measure diversity. The one most commonly used in social geography is the Index of Diversity, which is, when applied to a population and sample respectively:

$$S = 1 - \sum \left({}^{n}/_{N}\right) \tag{A2.13}$$

$$S = 1 - \sum \left(n\,(n-1)\;\; N(N-1)\right) \tag{A2.14}$$

Using our common notation for population groups in geographical areas, the Index of Diversity becomes

$$V_i = 1 - \sum_{e=1}^{e=M} \left(R_e^i\right)^2 \tag{A2.15}$$

where the summation is over *M* ethnic groups. It can be expressed in terms of the population counts as follows:

$$V_i = 1 - \sum_{e=1}^{e=M} \left(P_e^i \;\; P_+^i\right)^2 \tag{A2.16}.$$

The minimum value of V is 0 when there is only one group. The maximum value depends on the number of groups used and occurs when the population is distributed evenly between the groups. With a very large number of groups across which the population is evenly distributed, the maximum approaches 1 – with a 100 groups the maximum is 0.99.

A2.4.2. The Entropy Index

The state of entropy, derived from statistical mechanics used in the physical sciences, occurs when all state probabilities are the same. This is conceived of as the most chaotic state of a system. The concept of entropy was applied to Information Theory by Shannon (1948), who proposed the Entropy index (Shannon Index) as a measure of information, which is useful for measuring diversity:

$$H_i = -1 \times \sum_e \left(P_{ie} \times \ln(P_{ie})\right) \quad \ln(n) \tag{A2.17}$$

where ln is the natural logarithm function. The entropy index takes into account the number of groups and the evenness of groups. The index increases when there are more groups and they have more uniform sizes. Its value is 0 when there is only 1 group in the population and 1 when all groups are equally present. The disadvantage of the index is that it cannot be computed when one or more groups have zero populations. In practice, a very small constant can be introduced. The correlation between Simpson's Diversity Index and Shannon's Entropy Index when applied to the same body of population data is very high. For convenience, we prefer to use Simpson's V.

APPENDIX A3. MAPS OF DIVERSITY IN LEEDS AND WARSAW

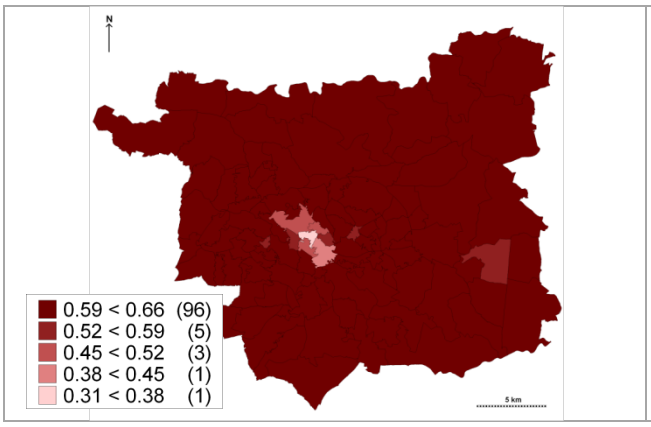


Figure A3.1. Simpson Index – 3 age groups

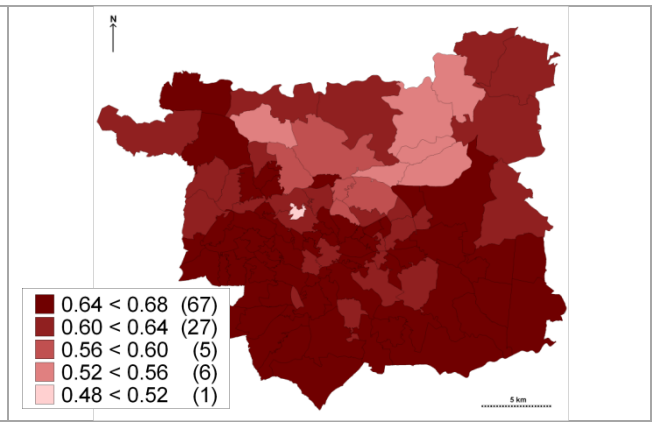


Figure A3.2. Simpson Index – 3 occupation groups

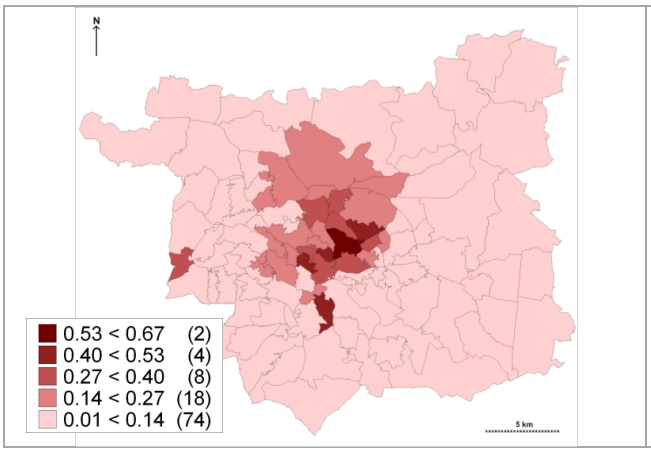


Figure A3.3. Simpson Index – 3 ethnic groups

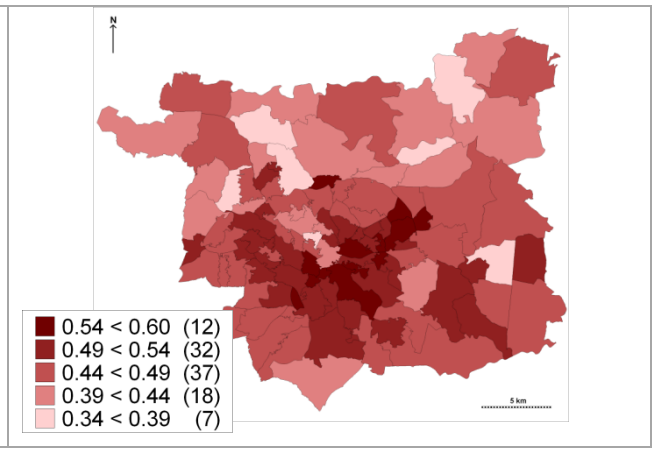


Figure A3.4. Simpson Index – 3 health groups

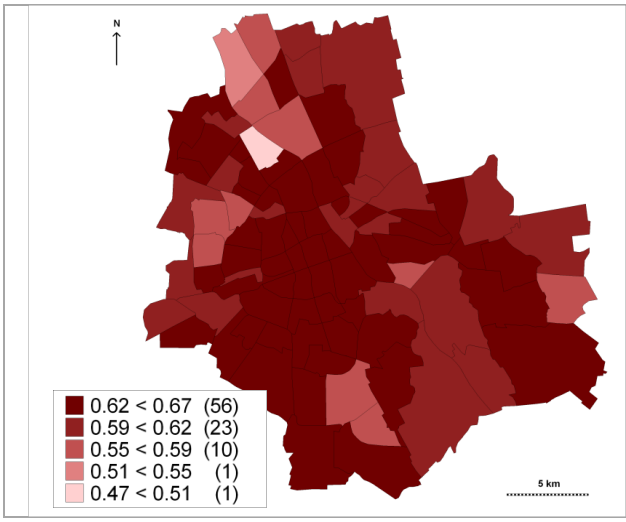


Figure A3.5. Simpson Index – 3 age groups

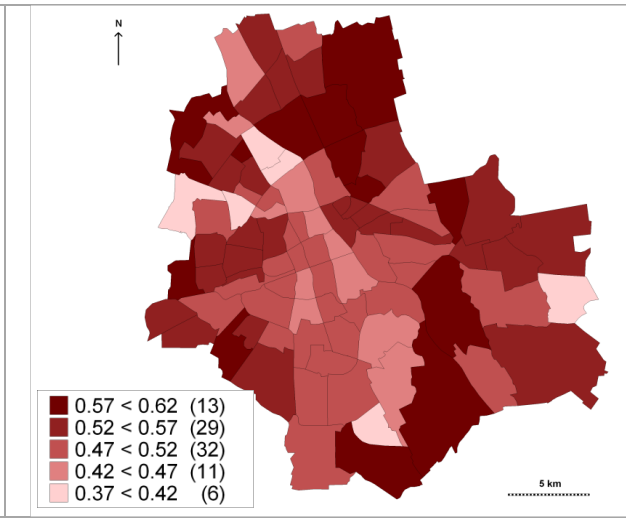


Figure A3.6. Simpson Index – 3 occupation groups

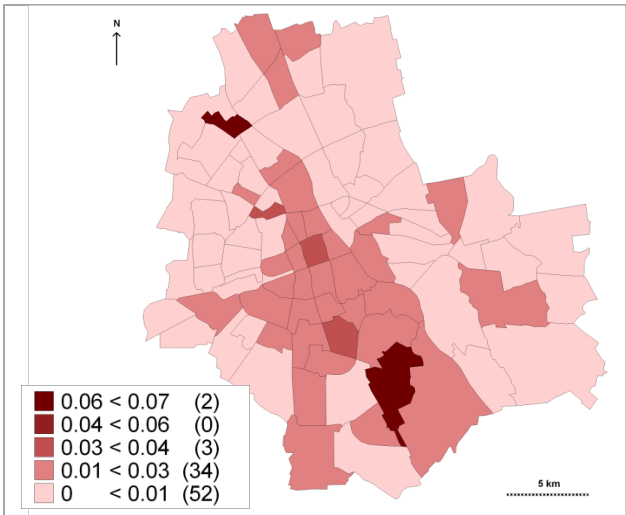


Figure A3.7. Simpson Index – 3 citizenship groups

APPENDIX A4. CHARACTERISTIC OF CLUSTERS IN LEEDS – SIMPLE ANALYSIS

Table A4.1. Some selected statistic for eight clusters in Leeds (simple analysis)

Cluster solution	Number of CAs	Population	*Percentage of city population*	of whom percentage of			
				people aged 65 and more	population with manual occupation*	non White British	population with LLTI and Disabled
1	19	151,289	*21.2*	14.2	49.9	4.5	16.9
2	16	112,289	*15.7*	13.1	63.3	10.7	20.3
3	23	119,439	*16.7*	17.5	31.6	6.8	14.8
4	7	56,577	*7.9*	8.3	43.2	19.3	12.3
5	3	23,825	*3.3*	10.6	57.8	59.6	18.9
6	18	127,999	*17.9*	18.5	49.5	5.1	19.2
7	7	63,047	*8.8*	17.7	32.9	21.6	17.6
8	13	60,701	*8.5*	17.2	66.5	8.6	25.2
All clusters	**106**	**715,166**	***100.0***	**15.5**	**47.5**	**10.8**	**18.0**

* Among economically active population aged 16-74.

Clusters description and classification of Community Areas:

Cluster 1 – White British, Typical, Suburban Communities with Average Socioeconomic Profiles
Community Areas: 4 Ardsley East/West, 15 Bramley, 22 Churwell, 29 Drighlington, 33 Farnley, 34 Farsley & Rodley, 56 Kippax, 57 Kirkstall, 59 Lofthouse & Robin Hood, 63 Micklefield, 67 Morley North, 68 Morley South, 69 New Farnley, 74 Oulton & Woodlesford, 77 Pudsey Lowtown, 88 Stanningley, 93 Upper Armley, 97 Whinmoor, 100 Yeadon.

Cluster 2 – Working Class, Younger, City Communities with Above Average Disability
Community Areas: 2 Allerton Bywater & Gt and Lt Preston, 5 Armley, 10 Beeston Hill, 11 Belle Isle, 39 Gipton North, 43 Halton Moor, 44 Harehills, 47 Hawksworth, 49 Holbeck, 64 Middleton, 79 Richmond Hill, 82 Sandfords, Ganners & Moorside, 84 Scott Hall & Miles Hill, 91 Swinnow & Fairfields, 94 Upper Wortley, 99 Wythers

Cluster 3 – White British, Middle Class, Suburban Communities with Below Average Disability
Community Areas: 1 Adel, 3 Alwoodley & Wigton Moor, 6 Arthington & Pool, 7 Bardsey & East Keswick, 12 Boston Spa, 13 Bramham, 14 Bramhope, 19 Calverley, 24 Collingham & Linton, 25 Colton, 26 Cookridge, 32 Far Headingley, 41 Guiseley, 46 Harewood and District, 51 Horsforth, 52 Horsforth Newlaithes & Woodsid, 53 Horsforth West End, 78 Rawdon, 83 Thorner, 87 Shadwell, 96 Wetherby, 102 Aberford, 103 Scarcroft.

Cluster 4 – Young Adult, City Communities with Middle Class and Multiethnic Co-Residents
Community Areas: 16 Burley, 17 Burley Lodge & Little Woodhouse, 23 City Centre, 36 Garforth East, 48 Headingley, 55 Hyde Park, 58 Woodhouse.

Cluster 5 – Multiethnic, Young, City Communities with Working Class and Disabled Co-Residents
Community Areas: 21 – Chapeltown, 45 Harehills Triangle, 105 Little London.

Cluster 6 – Older, White British, Suburban Communities with Average Socioeconomic Profiles
Community Areas: 8 Barwick & Scholes, Beeston, 31 Fairbank, 35 Fearnville, 37 Garforth West, 38 Gildersome, 42 Halton / Whitkirk, 50 Holt Park, 60 Manston, 62 Methley, 73 Otley, 76 Pudsey, 80 Rothwell, 90 Swillington, 92 Tinshill, 98 Wortley, 101 Ledston & Ledsham, 106 Ireland Wood.

Cluster 7 – Multiethnic, Middle Class, Suburban Communities with Below Average Disability
Community Areas: 20 Chapel Allerton, 61 Meanwood, 66 Moortown, 71 Oakwood, 75 Priesthorpe, 81 Roundhay, 95 West Park

Cluster 8 – Working Class, Older, City Communities with Above Average Disability
Community Areas: 18 Burmantofts, Lincoln Green & Ebor Gardens, 27 Cottingley, 28 Crossgates, 30 East Bank, 40 Gipton South, 54 Hunslet / Stourton, 65 Moor Allerton, 70 New Wortley, 72 Osmondthorpe, 85 Seacroft North, 86 Seacroft South, 89 Swarcliffe, 104 Cross Green.

Table A4.2. Some selected statistic for eight clusters in Warsaw (simple analysis)

Cluster solution	Number of URs	Population	*Percentage of city population*	of whom percentage of		
				people aged 65 and more	population with manual occupation*	foreign immigrants
1	11	316,224	*18.7*	20.2	37.9	0.4
2	21	520,599	*30.8*	24.4	27.8	0.9
3	15	193,387	*11.5*	9.6	36.6	0.3
4	2	12,936	*0.8*	13.3	26.4	2.6
5	13	200,592	*11.9*	14.9	28.4	0.8
6	10	269,479	*16.0*	5.8	24.2	0.5
7	4	3,089	*0.2*	12.5	73.1	0.2
8	15	171,996	*10.2*	13.0	47.8	0.3
All clusters	**91**	**1,688,302**	***100.0***	**16.5**	**32.0**	**0.6**

* Among economically active population aged 16-74.

Clusters description and classification of Urban Regions:

Cluster 1 – Polish, Older Regions with Above Average Working Class (inferior areas)
Urban Regions: 5 Bielany, 7 Brzeziny, 16 Falenica, 20 Grochów Północ, 21 Grochów Południe, 28 Koło Północ, 29 Koło Południe, 38 Młynów, 61 Praga II i III, 77 Ursus Skorosze, 85 Wygoda.

Cluster 2 – Older, Middle Class, City Centre Regions with Foreign Co-Residents
Urban Regions: 8 Centrum Północ, 9 Centrum Południe, 31 Las Kabacki, 35 Mirów (Środm. Zach.), 39 Mokotów Centrum, 40 Mokotów Stary, 41 Mokotów Wschód, 42 Muranów Wschodni, 43 Muranów Zachodni, 47 Ochota Centrum, 49 Okęcie Nowe, 56 Pole Mokotowskie, 58 PowiŚle Północ, 59 PowiŚle Południe, 65 Saska Kępa, 67 Sielce, 70 Stare i Nowe Miasto, 72 Szosa Krakowska, 83 Wierzbno, 89 Żoliborz Centralny, 91 Żoliborz Zachodni.

Cluster 3 – Polish, Young Regions with Average Socioeconomic Profiles
Urban Regions: 4 Białołęka Dworska, 17 FSO, 24 Jelonki, 25 Julianów, 51 Okęcie-Lotnisko, 53 Park Sobieskiego, 64 Rembertów, 73 Tarchomin, 74 Targówek Mieszkaniowy, 76 Ursus Gołąbki, 80 Wawrzyszew, 81 Wesoła Północ, 86 Zacisze, 88 Żerań Zachodni, 90 Żoliborz Przemysłowy.

Cluster 4 – Multiethnic, Middle Class, Young Regions (scattered in the city)
Urban Regions: 13 Czerniaków Wilanów and 36 Młociny.

Cluster 5 – Typical Traits Regions outside City Centre Regions, with Below Average Working Class and Some Foreigners
Urban Regions: 10 Choszczówka, 12 Czerniaków, 14 Czyste, 15 Dąbrówka, 22 Henryków, 34 Marymont, 54 Piaski, 63 Pyry, 66 Siekierki, 68 Służew, 69 Służewiec Południowy, 79 Wawer, 84 Włochy.

Cluster 6 – Young, Middle Class, Polish Regions (new estates outside City Centre)
Urban Regions: 1 Babice, 2 Bemowo Wschód, 3 Bemowo Zachód, 19 Górce, 30 Las Bielański, 32 Lotnisko, 44 Natolin, 46 Nowodwory, 78 Ursynów, 82 Wesoła Południe.

Cluster 7 – Hyper-Working Class, Polish, Younger Regions Post-manufacturing regions outside City Centre
Urban Regions: 23 Huta Warszawa, 48 Odolany, 55 PKP Wola, 75 Targówek Przemysłowy.

Cluster 8 – Working Class, Polish, Younger Regions outside City Centre
Urban Regions: 6 Bródno, 11 Chrzanów, 18 Gocław, 26 Kamionek, 27 Kawęczyn, 33 Mańki Brzeziny, 37 Młociny Las, 45 Nowa Praga, 50 Okęcie Przemysłowe, 52 Okęcie Opacz, 57 Postojowa, 60 Powsin, 62 Praga Centrum, 71 Szmulowizna, 87 Żerań Wschodni.

APPENDIX A5. CLUSTERS DISTRIBUTION FOR 216 CLUSTER SOLUTIONS

Table A5.1. Cases in 216 cluster solutions for 106 CA Leeds, simple analysis (Manhattan distance, 200 replicates)

Number of cluster	Cluster solution														
	2	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
1	63	38	27	19	18	7	16	17	12	14	7	7	7	7	7
2	43	39	11	37	28	9	3	8	7	7	15	6	11	14	6
3		29	37	7	19	14	13	13	8	7	6	13	5	3	5
4			31	15	20	23	7	23	18	12	7	7	7	7	6
5				28	13	15	7	13	11	13	9	9	7	3	5
6					8	19	22	3	13	8	13	13	6	7	3
7						19	19	7	16	7	9	7	6	9	8
8							19	7	7	17	5	6	12	6	3
9								15	3	3	13	7	13	6	9
10									11	11	7	11	8	7	14
11										7	3	5	7	6	7
12											12	12	3	5	3
13												3	11	9	9
14													3	5	10
15														12	6
16															5

Table A5.2. Cases in 216 cluster solutions for 91 UR Warsaw, simple analysis (Manhattan distance, 200 replicates)

Number of cluster	Cluster solution														
	2	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
1	59	45	27	21	15	11	4	12	12	10	11	9	7	7	4
2	32	23	23	21	21	15	15	21	9	8	10	2	3	3	5
3		23	24	15	19	4	21	9	7	9	7	5	9	3	7
4			17	15	10	15	10	8	2	13	7	4	5	7	5
5				19	15	21	15	2	10	12	8	7	4	13	7
6					11	15	2	9	8	10	5	11	5	5	13
7						10	11	11	4	5	13	5	7	5	5
8							13	4	5	11	4	7	2	7	3
9								15	13	4	9	7	13	4	9
10									21	7	10	5	10	7	7
11										2	5	8	3	7	5
12											2	13	11	9	2
13												8	7	7	4
14													5	2	2
15														5	3
16															10

Table A5.3 Cases in 216 cluster solutions for 106 CA Leeds, complex analysis (Manhattan distance, 200 replicates) [to be completed]

Number	Cluster solution														
of cluster	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	69	64	38	29	29	11	18	7	7	6	6	2	7	9	8
2	37	6	29	11	23	13	6	11	11	7	18	16	11	7	2
3		36	6	23	6	19	3	11	9	7	9	1	14	4	11
4			33	6	23	6	13	6	14	15	9	6	1	9	2
5				37	11	23	7	15	6	5	6	3	9	7	5
6					14	27	27	13	3	8	2	14	7	15	4
7						7	19	13	19	15	13	7	9	3	11
8							13	13	19	17	9	15	5	4	6
9								17	5	3	5	5	11	7	5
10									13	17	13	13	7	7	7
11										6	13	9	5	6	13
12											3	7	13	3	5
13												8	6	3	3
14													1	9	2
15														13	5
16															17

Table A5.4. Cases in 216 cluster solutions for 91 UR Warsaw, complex analysis (Manhattan distance, 200 replicates) [to be completed]

Number	Cluster solution														
of cluster	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	40	30	28	19	11	15	8	9	13	8	9	1	3	8	11
2	51	42	15	11	10	9	19	3	1	9	7	9	1	5	9
3		19	25	25	13	9	15	11	7	15	1	3	7	1	5
4			23	15	21	21	17	18	9	9	17	7	3	18	1
5				21	17	9	7	7	11	7	8	3	9	7	1
6					19	9	9	9	9	3	13	11	6	8	9
7						19	9	7	18	9	1	1	9	11	6
8							7	8	8	11	9	11	11	5	5
9								19	9	18	3	1	1	1	3
10									6	1	9	9	7	6	5
11										1	7	9	9	3	8
12											7	9	5	3	1
13												17	13	1	11
14													7	9	8
15														5	5
16															3

Table A5.5. Cases in 216 cluster solutions for 197 CA Leeds and UR Warsaw, combined analysis (Manhattan distance, 200 replicates)

Number	Cluster solution														
of cluster	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	120	33	67	37	33	35	25	21	15	14	9	9	15	19	15
2	77	64	27	43	51	23	22	33	21	17	13	23	18	19	11
3		100	58	21	19	29	34	30	31	19	17	19	14	3	12
4			45	45	23	26	43	13	20	25	27	14	4	13	13
5			0	51	42	48	16	26	15	13	24	15	11	13	9
6					29	23	21	15	11	11	4	17	17	5	9
7						13	23	15	13	16	21	13	9	14	19
8							13	22	17	21	23	19	22	15	7
9								22	31	28	13	20	15	18	4
10									23	14	10	9	19	10	17
11										19	17	13	19	11	14
12											19	15	9	18	13
13												11	14	3	13
14													11	17	14
15														19	11
16															16

APPENDIX A6. SILHOUETTE PLOTS OF CLUSTER SOLUTIONS IN LEEDS AND WARSAW

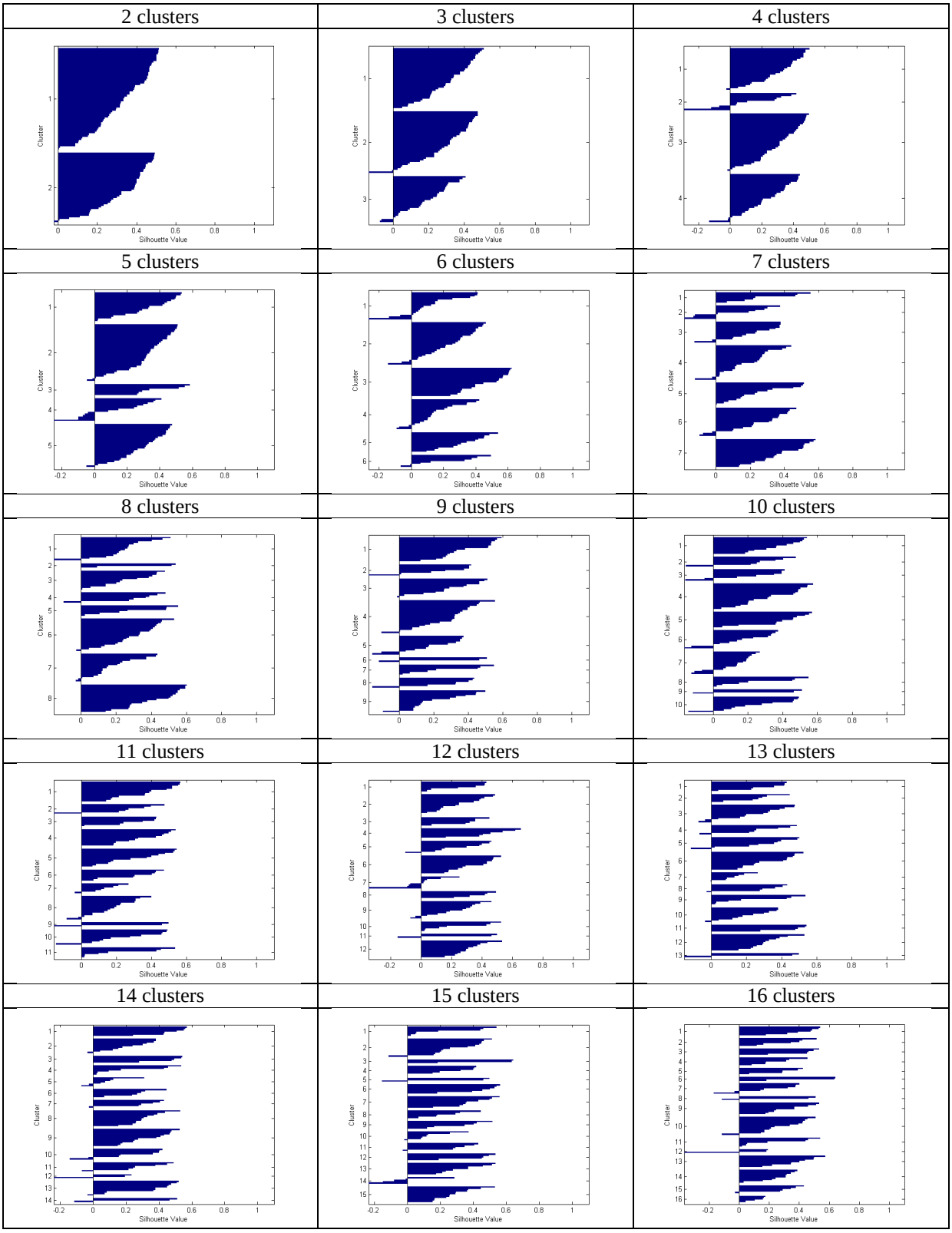


Figure A6.1. Silhouette plots of cluster solutions of 106 CAs in Leeds, simple analysis (k-means, 200 replicates, Manhattan distance)

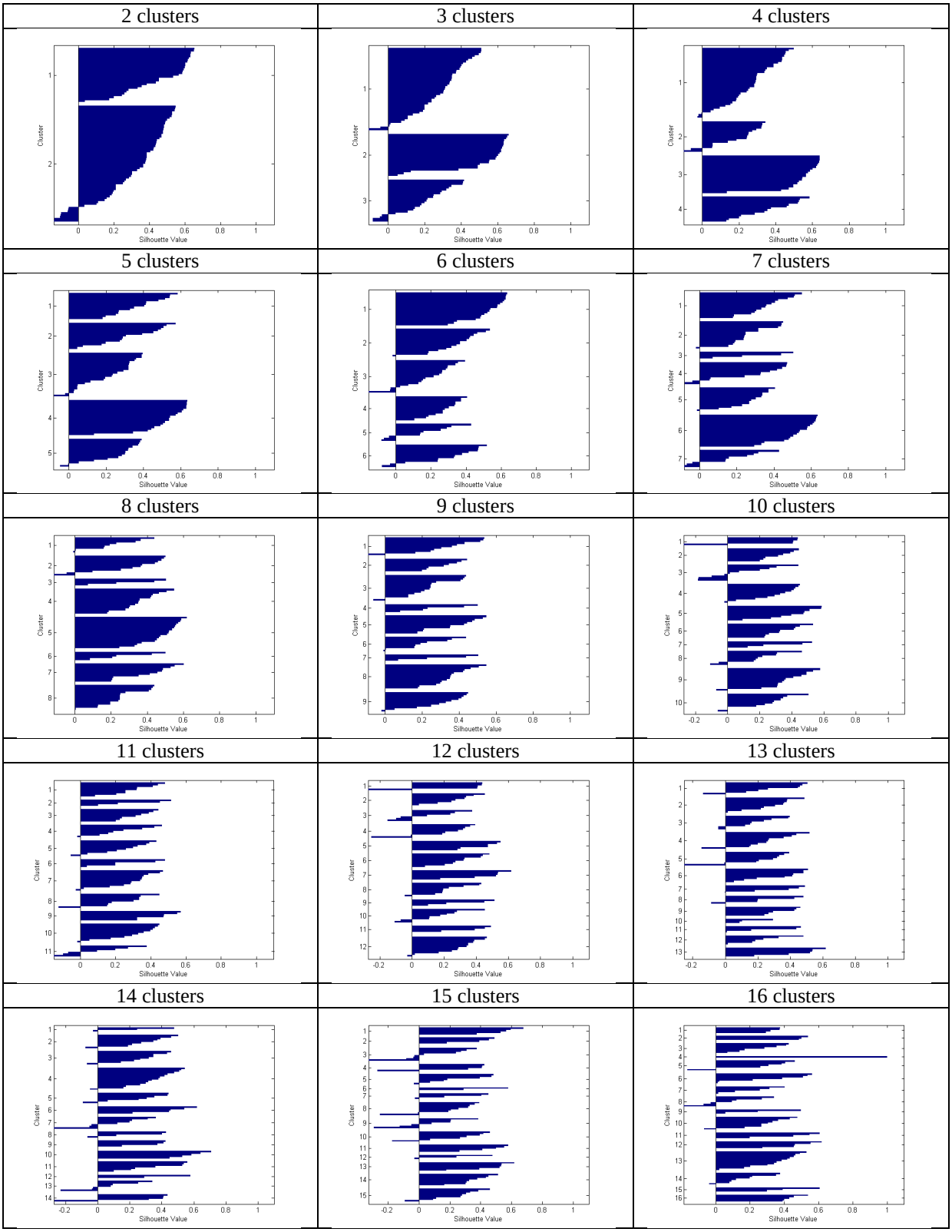


Figure A6.2. Silhouette plots of cluster solutions of 91 URs in Warsaw, simple analysis (k-means, 200 replicates, Manhattan distance)

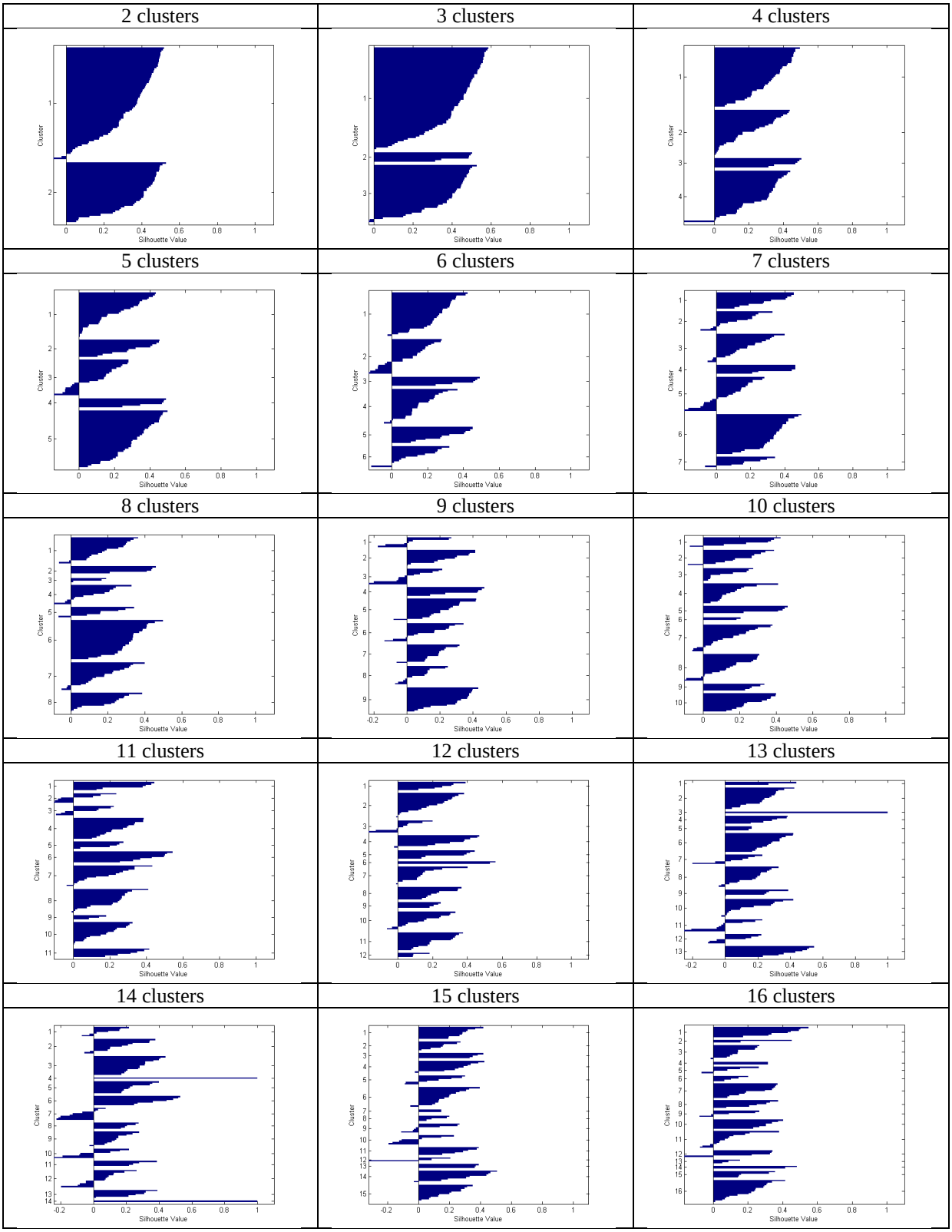


Figure A6.3 Silhouette plots of cluster solutions 106 CA Leeds, complex analysis (k-means, 200 replicates, Manhattan distance)

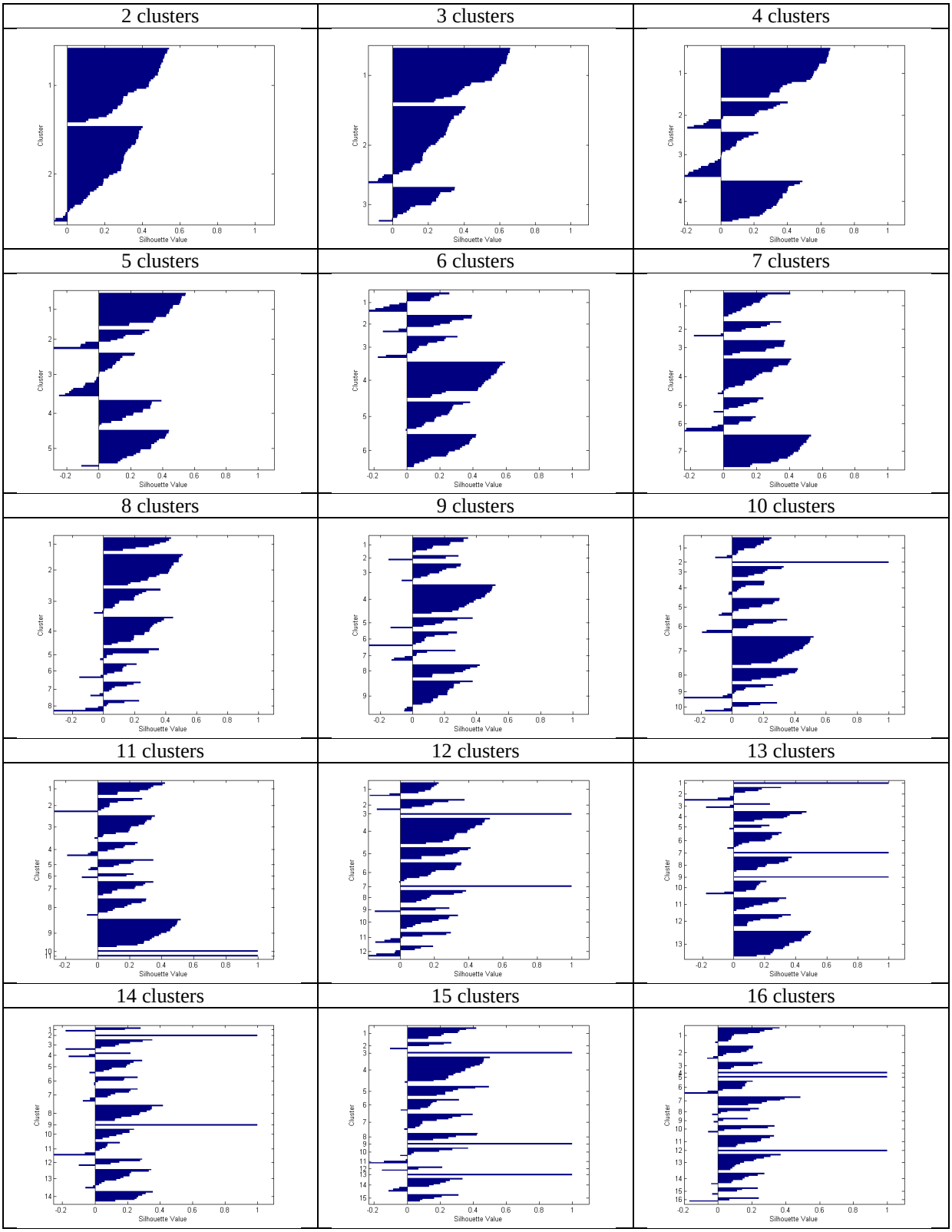


Figure A6.4. Silhouette plots of cluster solutions of 91 UR Warsaw, complex analysis (k-means, 200 replicates, Manhattan distance)

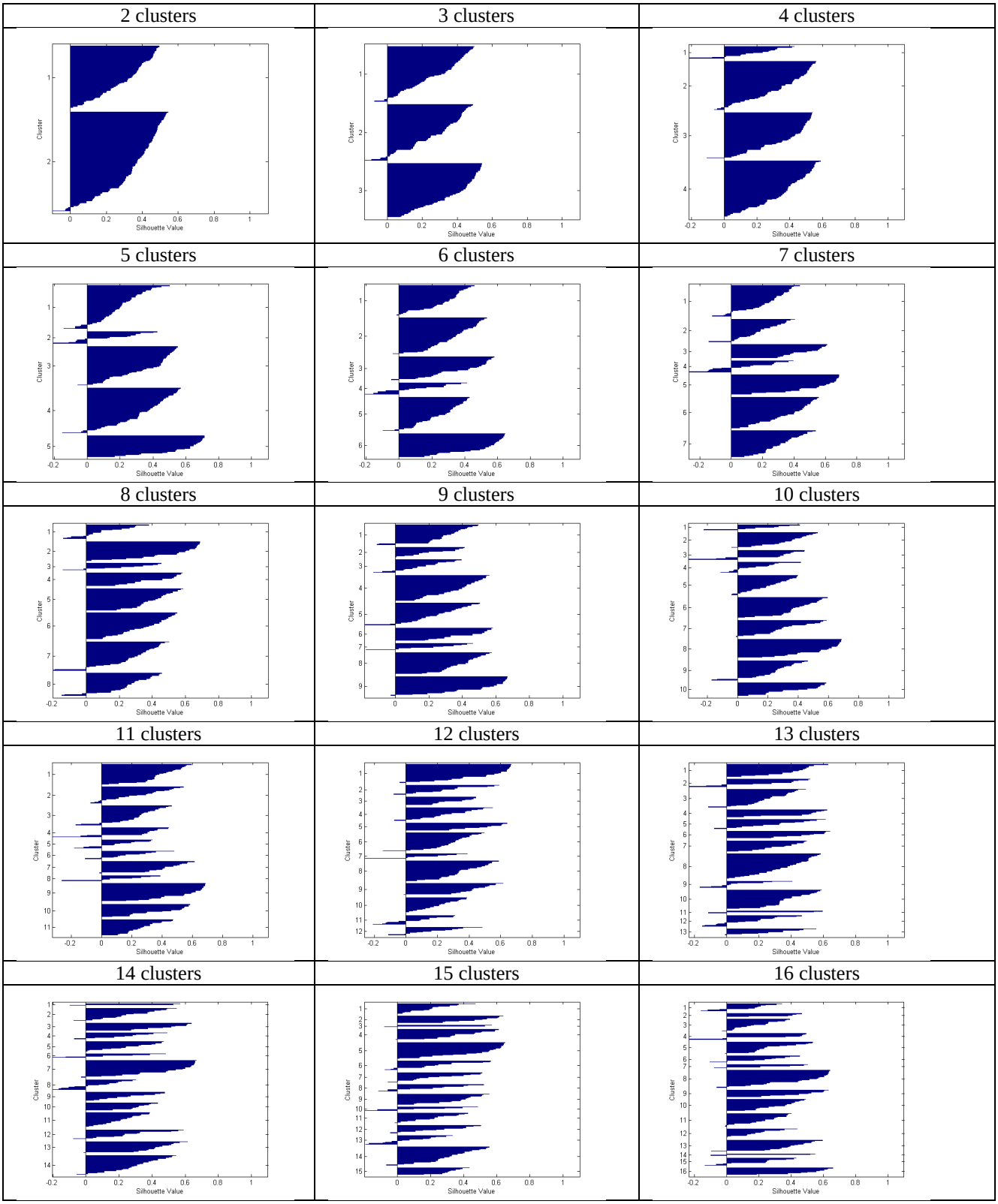


Figure A6.5 Silhouette plots of cluster solutions of 197 CAs and URs in both cities, combined analysis (k-means, 200 replicates, Manhattan distance)

APPENDIX A7. BUBBLE PLOTS OF SIMPSON INDEXES

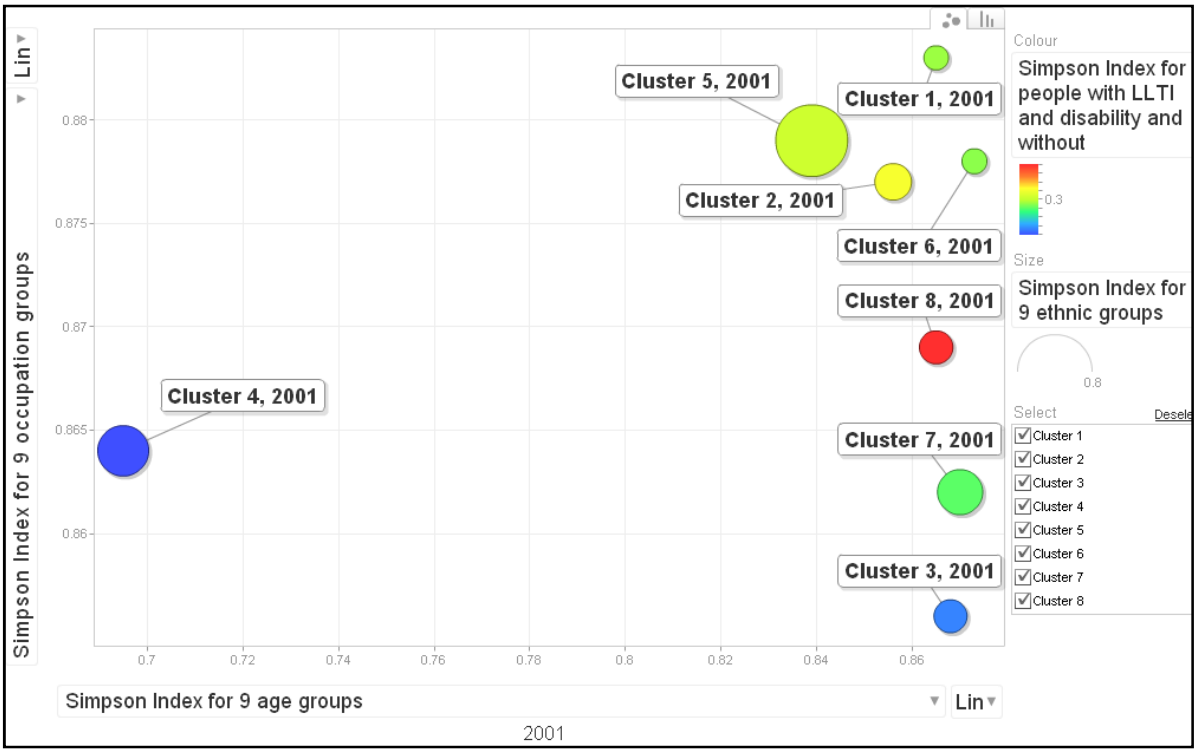


Figure A7.1. Simpson Indexes for 4 diversity dimensions in Leeds

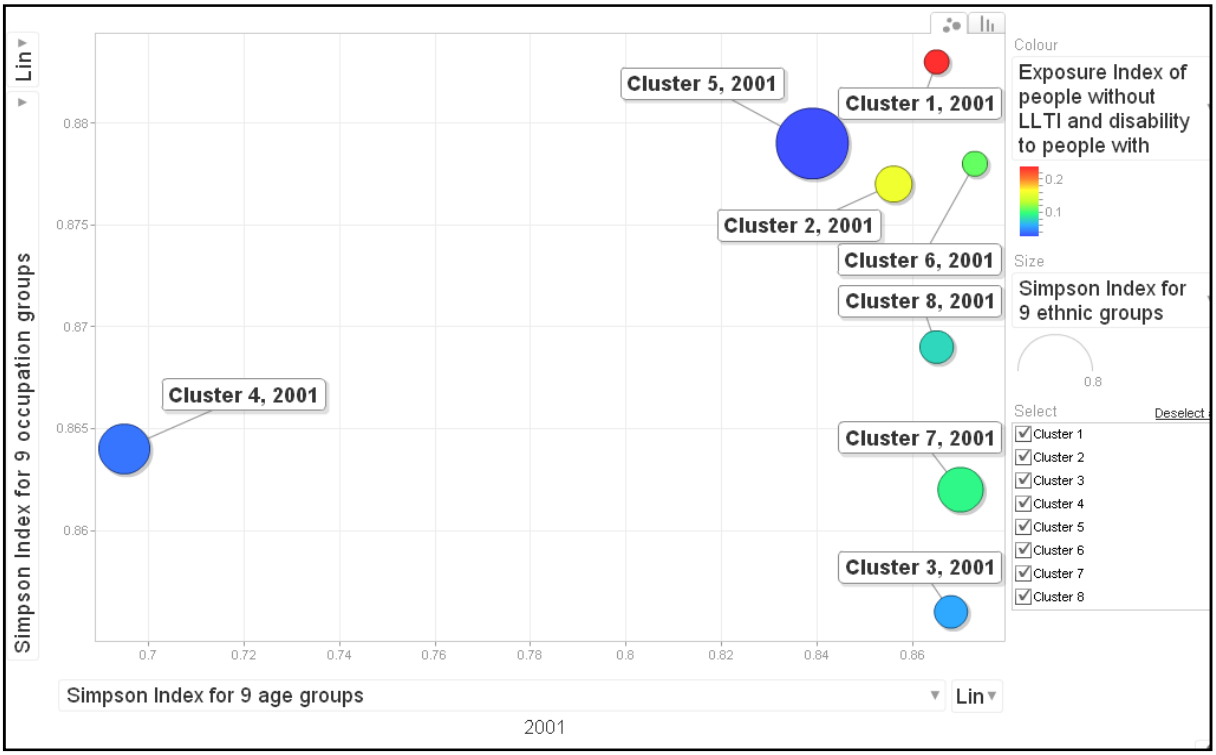


Figure A7.2. Simpson Indexes for 3 diversity dimensions and Exposure Index for 1 diversity dimension in Leeds

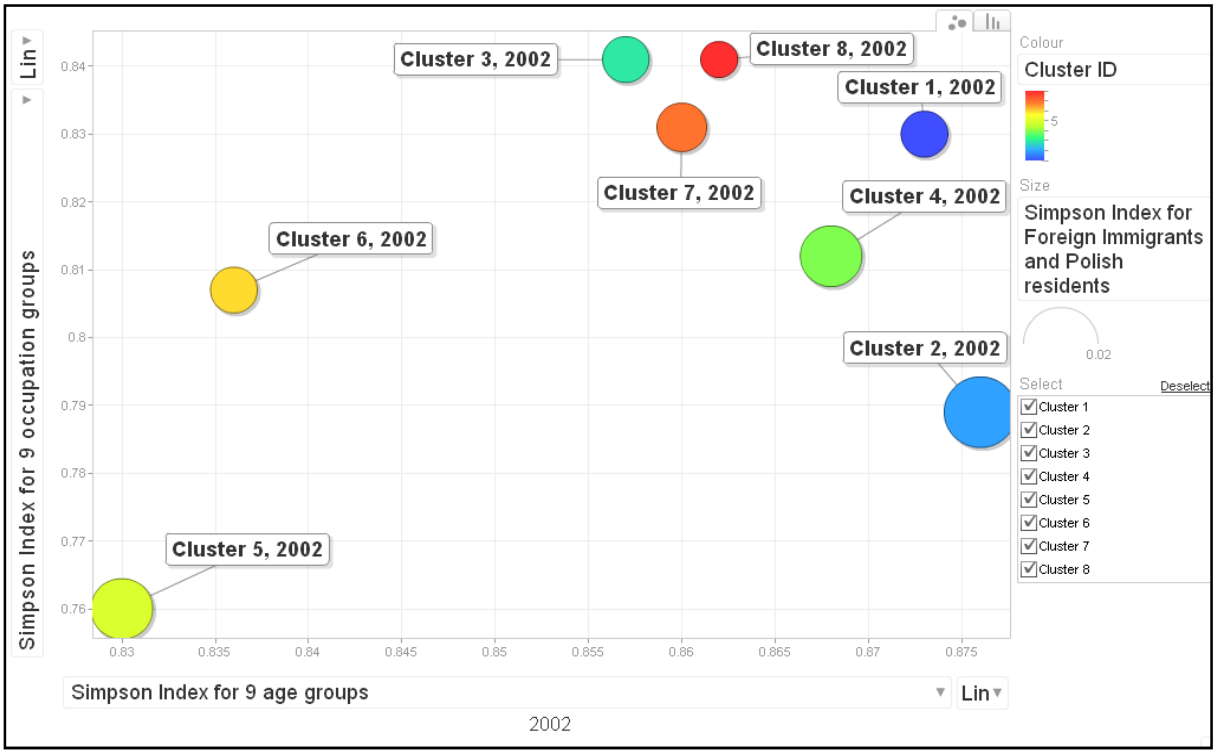


Figure A7.3. Simpson Indexes for 3 diversity dimensions in Warsaw

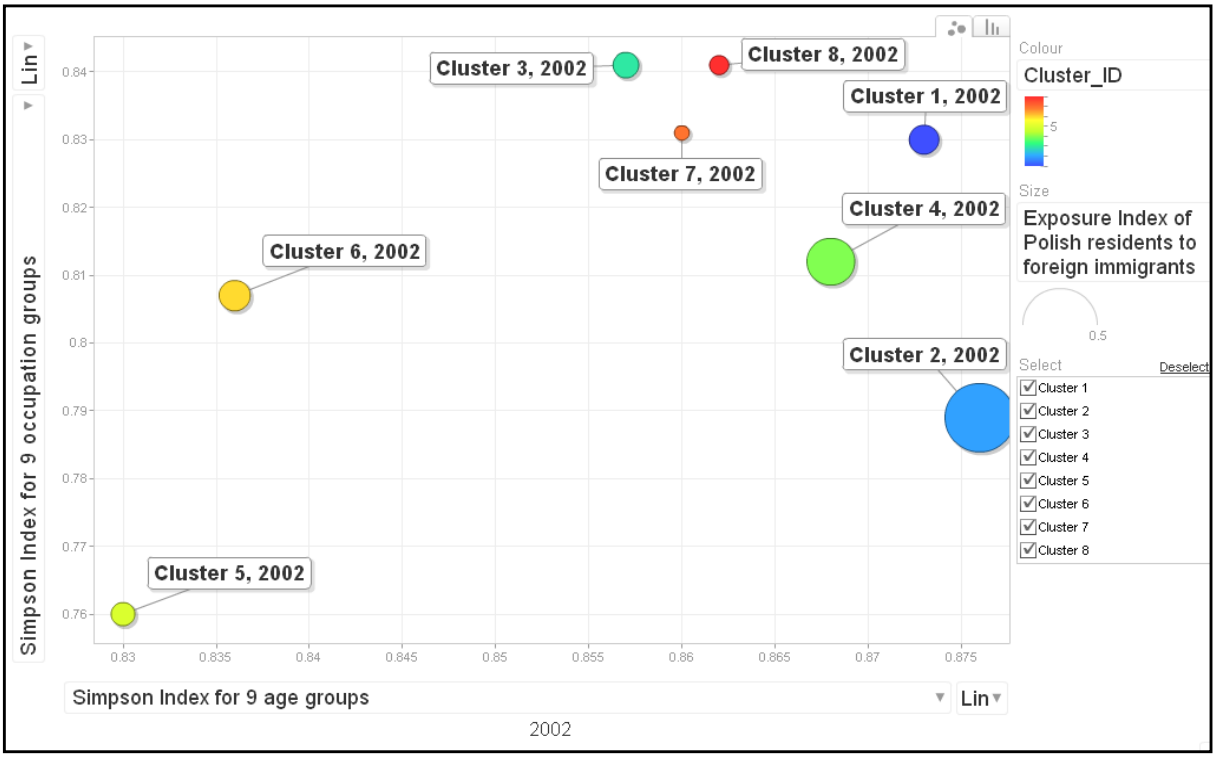


Figure A7.4. Simpson Indexes for 2 diversity dimensions and Exposure Index for 1 diversity dimension in Warsaw