

Working Paper 28

Comparing locations: Grouping wards in London, the South East and East of England

By **Said Hirsh** and **Ellie Cooper** (Volterra Consulting)



**Transport
for London**



MAYOR OF LONDON

copyright

**Greater London Authority
January 2008**

Published by

Greater London Authority
City Hall
The Queen's Walk
London SE1 2AA

www.london.gov.uk

enquiries **020 7983 4000**

minicom **020 7983 4458**

ISBN: **978-1-84781-116-5**

Cover photograph

© Fergus Noone

This publication is printed on recycled paper

For more information about this publication, please contact:

GLA Economics

telephone 020 7983 4922

email glaeconomics@london.gov.uk

GLA Economics provides expert advice and analysis on London's economy and the economic issues facing the capital. Data and analysis from GLA Economics form a basis for the policy and investment decisions facing the Mayor of London and the GLA group. The unit is funded by the Greater London Authority, Transport for London and the London Development Agency.

GLA Economics uses a wide range of information and data sourced from third party suppliers within its analysis and reports. GLA Economics cannot be held responsible for the accuracy or timeliness of this information and data.

GLA Economics, the GLA, LDA and TfL will not be liable for any losses suffered or liabilities incurred by a party as a result of that party relying in any way on the information contained in this report.

Report Highlights

- This report analyses the changes in *area* and *employment* characteristics of wards in London, the South East and parts of the East government office regions over the period 1991 to 2004. The wards are classified into five *area* groups and three *employment* groups.
- In terms of *area* characteristics, we see increases in both population and employment densities, but there is a less even increase in employment density where substantial increases occur in London and in and around other urban centres with no or little change in many rural locations.
- Another variable experiencing some changes is the proportion of ethnic minorities where we observe increases across the region, with wards both in and outside of London seeing a small but evenly spread increase.
- In terms of *employment* characteristics, employment in Retail & Leisure and Business Support Services sees the fastest growth; as a result, wards classified as Commercial dominate the landscape and the number of wards classified as Industrial or Public Services falls.
- Deprivation levels fall across the board but the sources of deprivation in wards in London and those outside of London are found to be different. In London, the sources of deprivation are housing, crime and the living environment, while education and health emerge as the main sources of deprivation in wards outside of London.
- Wards that are classified as Commercial and those in areas with higher population and employment densities benefited from faster growth in employment over the period. Furthermore, the source of employment growth in less dense and Industrial is employment in Retail & Leisure and Business Services.
- Finally, employment growth was very slow, or even negative, in wards classified as Public Sector. This reflects government expenditure in the most deprived areas.

Contents

Executive Summary	1
1. Introduction	5
2. Methodology	6
3. The Data	6
4. Group Results	16
5. Conclusions	29
Appendix I: Technical Background	30
Appendix II: Industrial Grouping	32
Appendix III: Industrial Variable Maps	35
Appendix IV: Group Maps	46
Appendix V: Degrees of Membership	54

Executive Summary

Introduction

The purpose of this analysis is to identify wards which have sufficient similarity to be identified as a group. In particular, we want to see the extent to which there is an overlap between places in London and other surrounding regions. We look at the employment and demographic structure of wards in London, the South East and part of the East government office regions. Geographical analysis is usually based on administrative boundaries as these provide the basis of many official data sources. However, ward boundaries are defined to try and ensure that the same number of electors is represented in each ward and these boundaries do not necessarily represent anything useful about the economy and society in which we live. However, as wards are the smallest possible available geographical data set, we use these as the unit of measurement and examine its characteristics.

The technique we adopt is a modern extension of cluster analysis called fuzzy clustering. In this technique, we do not need to make assumptions and pre-define locations; instead, we use the data to determine groupings of areas that are similar to each other with respect to employment and demographic characteristics.

Changes in Group Memberships, 1991 to 2004

We use data on the employment and demographic structure of 2723 wards in London, the South East and parts of the East of England to create groupings. The result is a powerful classification system made up of three employment groups and five area groups which are also combined to form 15 employment-area categories. We group the wards based on data from two time periods - 1991 and 2004. We look at how both the characteristics and the ward memberships of the groups have changed over this period. This allows us to get a good indication for how the study area has changed through time.

The employment grouping results in three groups which we entitle the Public Sector, Commercial and Industrial groups, due to relative domination of these sectors of employment in each case. Each ward has some degree of membership of each of the groups and is assigned to the group in which it has the highest membership. The area grouping results in five groups with group 1 containing wards which on average are larger and the least dense and group 5 the smallest and densest. Table one provides a description of each of the area groups with an example of a London ward belonging to that group. Each ward has some degree of membership of each of the groups.

Table 1: Description of area based groups - examples taken from 2004 results

Area Group	Description
1	Wards in this group have, on average, an area of 39 sq. km., a population density of 193 people per sq. km., employment density of 203 people per sq. km., an age of 41 years and 3 per cent unemployment - an example is Darwin in the London Borough of Bromley
2	Wards in this group have, on average, an area of 8 sq. km., a population density of 1,174 people per sq. km., employment density of 634 people per sq. km., an age of 41 years and 3 per cent unemployment - an example is East Sheen in the London Borough of Richmond upon Thames
3	Wards in this group have, on average, an area of 4 sq. km., a population density of 2,925 people per sq. km., employment density of 1,434 people per sq. km., an age of 38 years and 3 per cent unemployment - an example is Edgware in the London Borough of Barnet
4	Wards in this group have, on average, an area of 3 sq. km., a population density of 4,891 people per sq. km., employment density of 1,984 people per sq. km., an age of 36 years and 5 per cent unemployment - an example is Highgate in the London Borough of Camden
5	Wards in this group have, on average, an area of 1 sq. km., a population density of 10,518 people per sq. km., employment density of 4,831 people per sq. km., an age of 34 years and 9 per cent unemployment - an example is Whitechapel in the London Borough of Tower Hamlets

In terms of area data, changes happen slowly. The demographic characteristics of the wards have changed less than the employment pattern has. Nonetheless there has been some trend to a concentration of density in London, especially inner London, with some spreading of suburbs into rural locations.

For employment data, changes are much more apparent, with a fall in the membership of both Public Sector and Industrial groups. This trend is confirmed when we consider overall employment in the industrial and public services sectors in the regions considered; the first falling nearly five per cent overall while the latter increases only fourteen per cent over the period, approximately a third of the growth seen in service-based employment. Furthermore, the employment results seem to capture the trend of service consolidation in the public sector, with wards belonging to the Public Sector group becoming better defined as such despite the fall in the number of wards belonging to that group. The geographical pattern of employment is much more mixed than that of the residential attributes.

Cross-Classification of Groups

Here we consider the cross classification of wards to employment and area groups. For example in 2004, there are 152 wards which have the highest memberships of the Commercial employment group and the least dense area group 1 and only 30 wards which have the highest membership of the Industrial employment group and the smallest area group 5; table two shows the cross classification of wards in 2004.

Table 2: Cross-Classification of Groups, 2004

	Public Sector	Commercial	Industrial	Total
Group 1	69	152	215	436
Group 2	224	288	267	779
Group 3	241	267	204	712
Group 4	186	219	139	544
Group 5	91	131	30	252
Total	811	1057	855	2723

We see that membership of the Public Sector and Commercial employment groups decreased in the first area group, while membership of the Industrial group fell in all area groups except for the first group. When considering the first area group, we find a large proportion of wards are also members of the Industrial employment group; in fact, when compared to the cross classification results in 1991, this number has increased. This suggests that group 1 wards have become more industrial.

Indices of Deprivation

The DCLG Indices of Multiple Deprivation are measures of seven different types of deprivation (eg Housing, Health, Education etc) which are combined together to give an overall measure of deprivation for each Super Output Area in England. When aggregated to a ward level and compared with the results of our classification system, the separate deprivation indices follow an expected trend. The first area group is on average consistently the least deprived across the seven separate deprivation indices. The only index which does not follow the trend is housing deprivation, where the first area group shows similar deprivation levels to those in the fifth group, the densest and smallest. The least deprived group in terms of housing is group 3. The employment groups do not show any consistent trends and show no clear relationship with the indices of deprivation.

Ethnic Minorities

Over the period 1991 to 2004, the proportion of ethnic minorities has increased in all area and employment groups.

London Wards

Average membership of the fourth and fifth groups increased substantially in London wards, while average membership of the other area groups has fallen, mainly due to the increase in population and employment densities across the capital. For wards outside London, average membership increased substantially in the third area group and slightly in the fifth; it fell in the three other area groups. This result suggests that population and employment density grew faster in wards in London than they did outside of London.

The sources of deprivation in London are housing, crime and the living environment. The average indices for all three categories are significantly higher in London wards than in out of London wards, both for the area and employment groups and overall.

Employment Growth

We find that the wards most associated with employment growth are those in the Commercial employment group; overall employment in these wards grew at an average of nearly 3 per cent per annum over the period 1991 to 2004. This compared to an annual average of only 0.3 per cent for Public Sector wards and 1.6 per cent for Industrial wards.

We also find that the source of growth in Public Sector and Industrial wards is service-based employment.

1. Introduction

Geographical analysis is usually based on administrative boundaries as these provide the basis of data sources. However, these boundaries are defined to try and ensure that the same number of electors is represented in each ward and do not necessarily represent anything useful about the economy and society in which we live. The purpose of this analysis is to identify groups of locations which have sufficient similarity to be identified as such. In particular, we want to see the extent to which there is an overlap between places in London and other surrounding regions.

To do this, we take the smallest possible available geographical data set at ward level and examine its characteristics. The aim of this study is to analyse the similarities and differences between the wards in London, the South East and parts of the East government office regions. A total of 2723 wards were used and the analysis was done on the 2004 ward boundaries. All the data used was adjusted for the 2004 boundaries where necessary.

We use modern statistical techniques of cluster analysis, specifically fuzzy clustering. Clustering, in general, enables us to categorise the wards into groups; the wards within a group are similar to each other but the average characteristics of each group are different. Fuzzy clustering, in particular, is a modern extension of clustering which is described in section 2.

In this study, we examine groupings on two different sets of characteristics:

- the type of *area* the ward is - this includes its population density, employment density, average age of the population, unemployment, index of multiple deprivation (IMD), the proportion of ethnic minorities and area in square kilometres
- the distribution of *employment* in the ward - this includes classifying the area's employment profile, i.e. the distribution of employment by specific sectors

Section 2 gives a detailed description of the methodology used to undertake this study. Section 3 discusses the data and section 4 sets out the results. Maps of the wards plotting details of the clustering in colour, breakdown of the employment groups and a formal representation of the techniques used are available in the appendices.

This paper updates and extends our analysis published in February 2003 as *Working paper 2: The Economic and Social Structure of London and the South East*.

2. Methodology

The task of forming groups or clusters is well known in applied statistics: it usually goes by the name of cluster analysis. An essential part of our classification system is the use of a technique called fuzzy clustering rather than standard methods of cluster analysis.

The aim of any clustering exercise is to form groups of areas so that the areas within any one group have similar characteristics but the typical characteristics of each of the groups are sufficiently different from those of the others. These patterns provide a way of understanding or describing the economic properties of areas.

In standard methods of clustering each area is allocated unequivocally to a single group. This procedure has a significant drawback as there is no way of distinguishing between those on the margin of any particular group from those in the centre. Fuzzy clustering offers a way around this drawback. It combines the ideas of standard clustering methods and fuzzy logic. Instead of forcing each area into a single cluster, it is given some degree of membership of every one of the clusters. In this way, an area can have some aspect of each of the clusters. This is not only a natural extension of the usual clustering techniques but also turns out to be a very powerful way of understanding patterns of local economies.

We set out the formal mathematics of the technique in Appendix I.

3. The Data

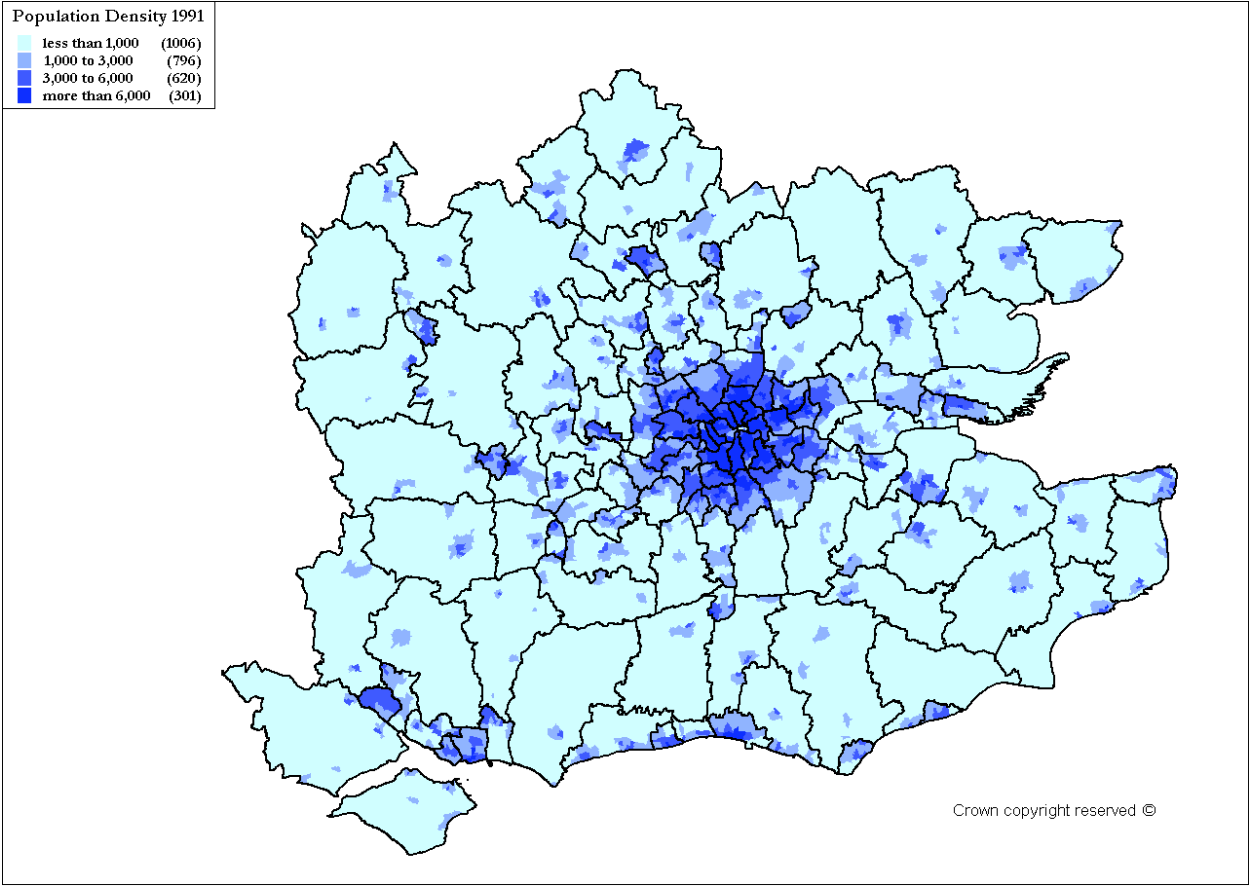
3.1 Area Variables

To describe the type of area a ward is we start by mapping a number of variables in order to get an understanding of which variables show interesting variations across the wards. The variables considered were:

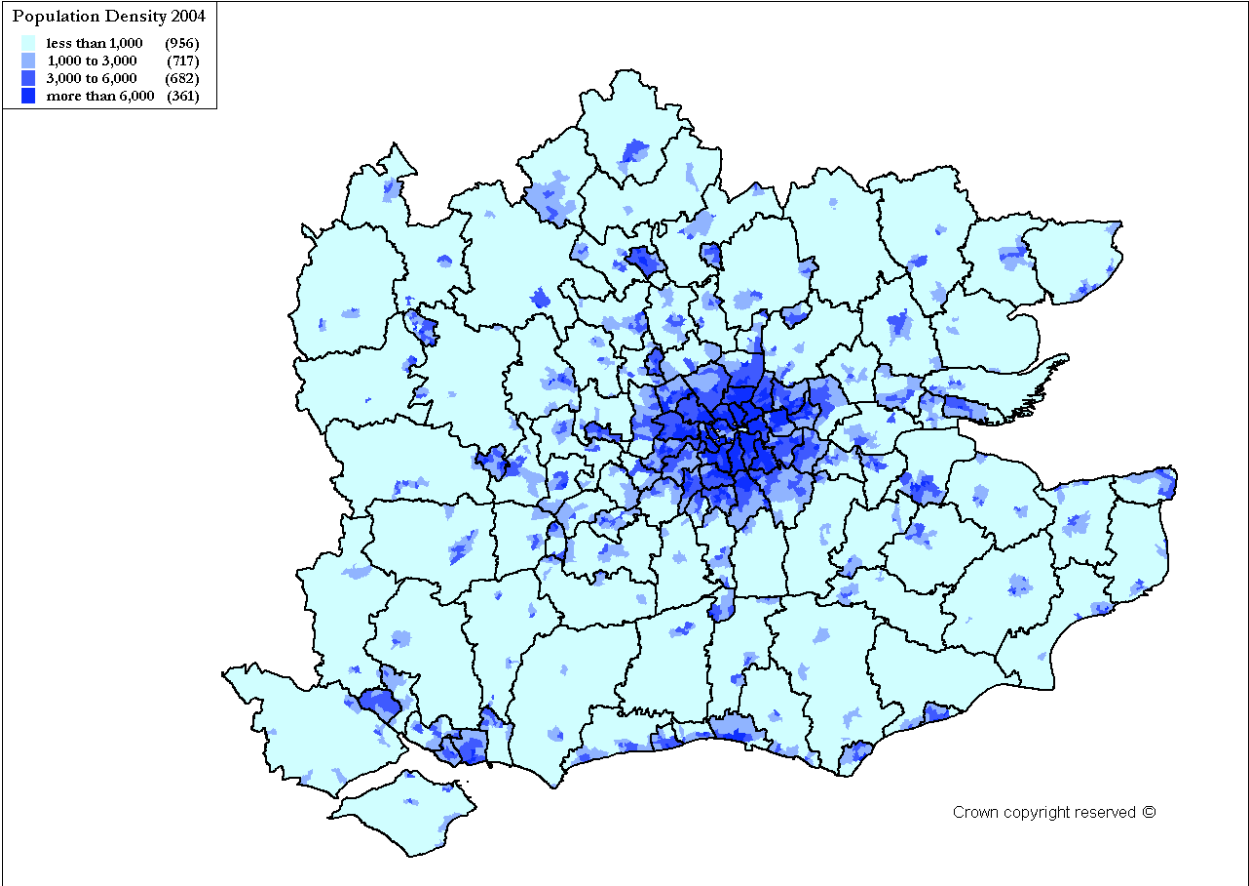
- population density (in people per square kilometre)
- employment density (in employees per square kilometre)
- average age
- unemployment (as a percentage of economically active people)
- proportion of ethnic minorities
- index of multiple deprivation (IMD); and
- size (in square kilometres).

The maps for the individual variables at both dates are shown in Appendix III and table three shows summary statistics for all the variables at both dates where possible. For population density, the results show that there was a large increase in the average population density across the wards; this is the result of an increase in overall population. There also seems to be an increase in the spread of population as indicated by the standard deviation. The results shown in the maps below confirm the trend of rising population densities with the two highest-ranking ranges seeing an increase in the number of wards. Employment density shows similar trends to those of population density; an increase in the average and also the number of wards moving up the ranges as shown in the maps. It is also clear, as shown in maps one and two, that the increase in population density is seen across the board, albeit to varying degrees. Growth in employment density displays similar trends; this is supported by the fall in unemployment across the regions considered during the period 1991 to 2004.

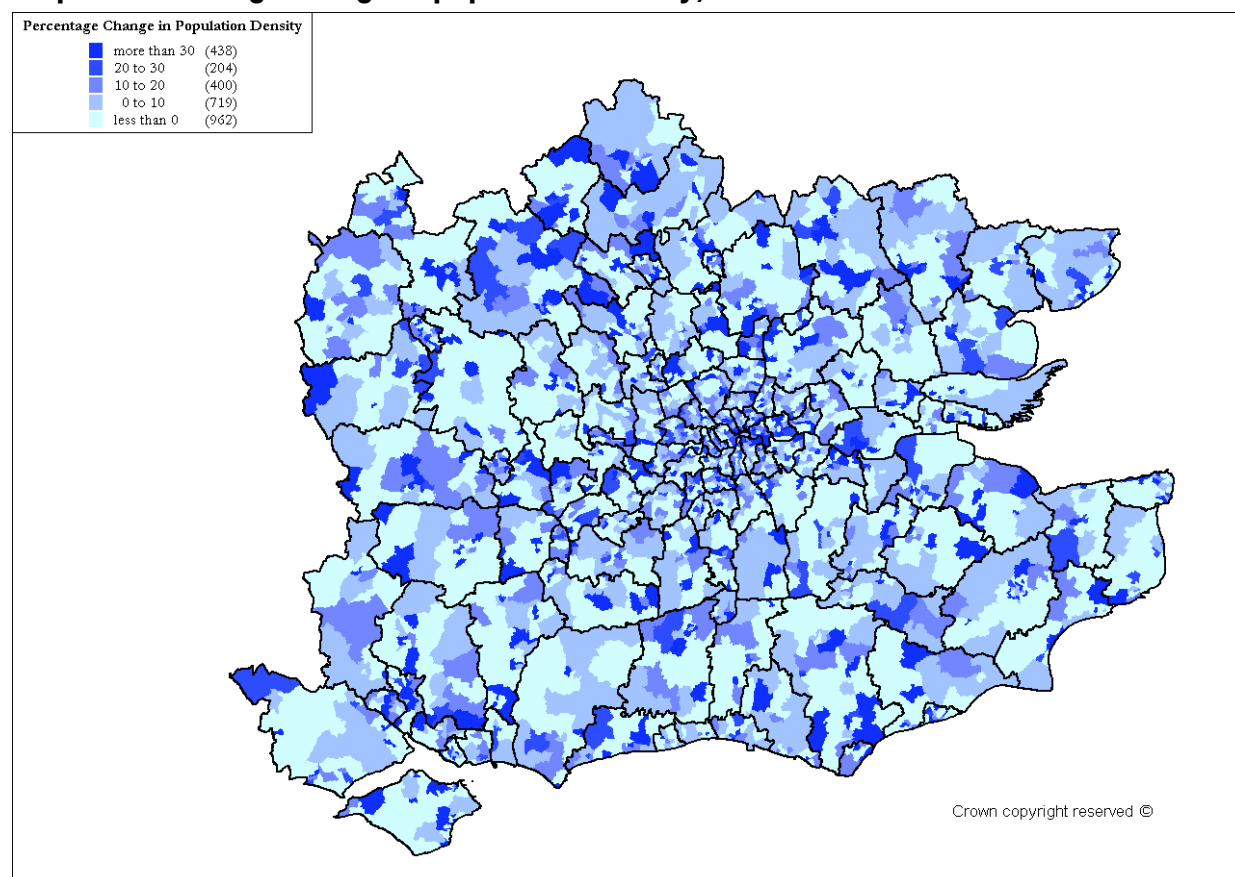
Map 1: Population Density, 1991



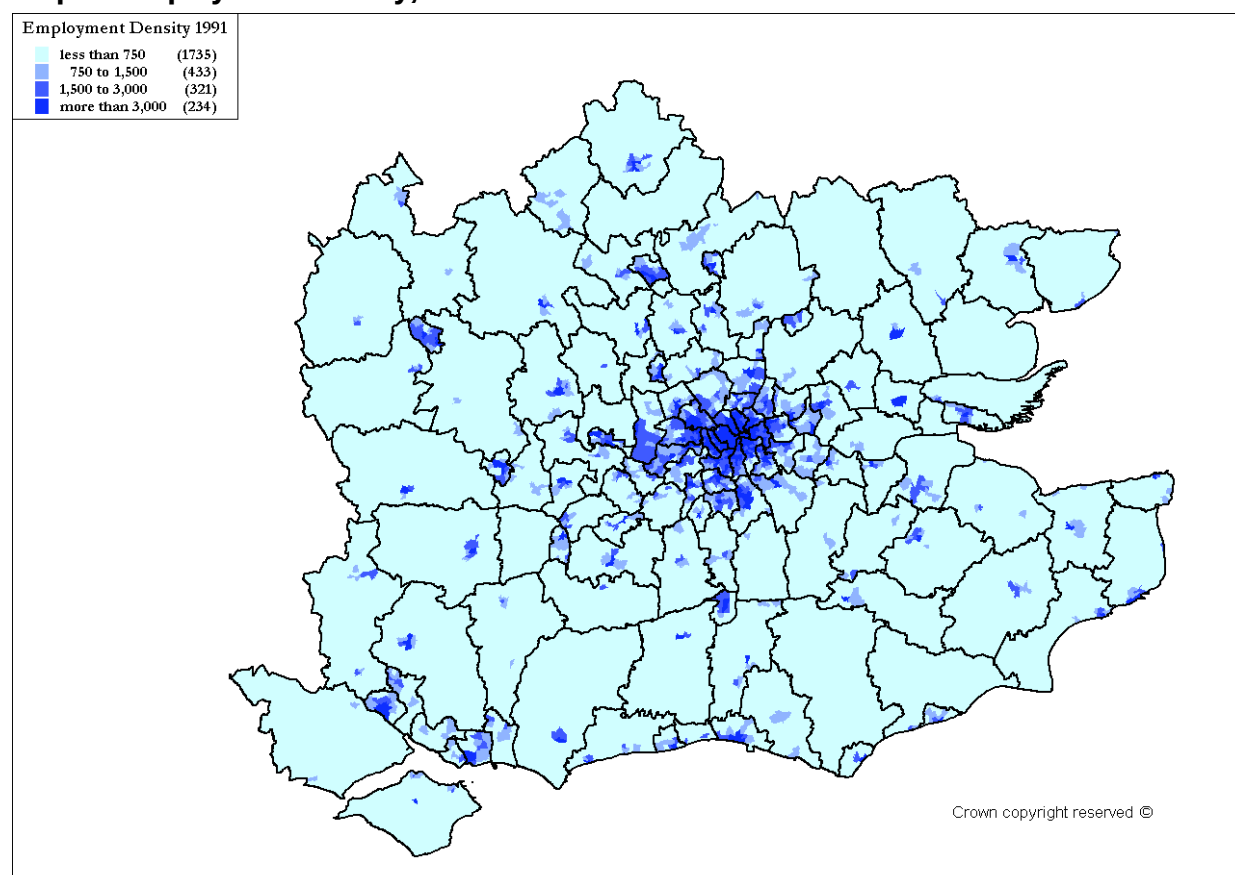
Map 2: Population density, 2004



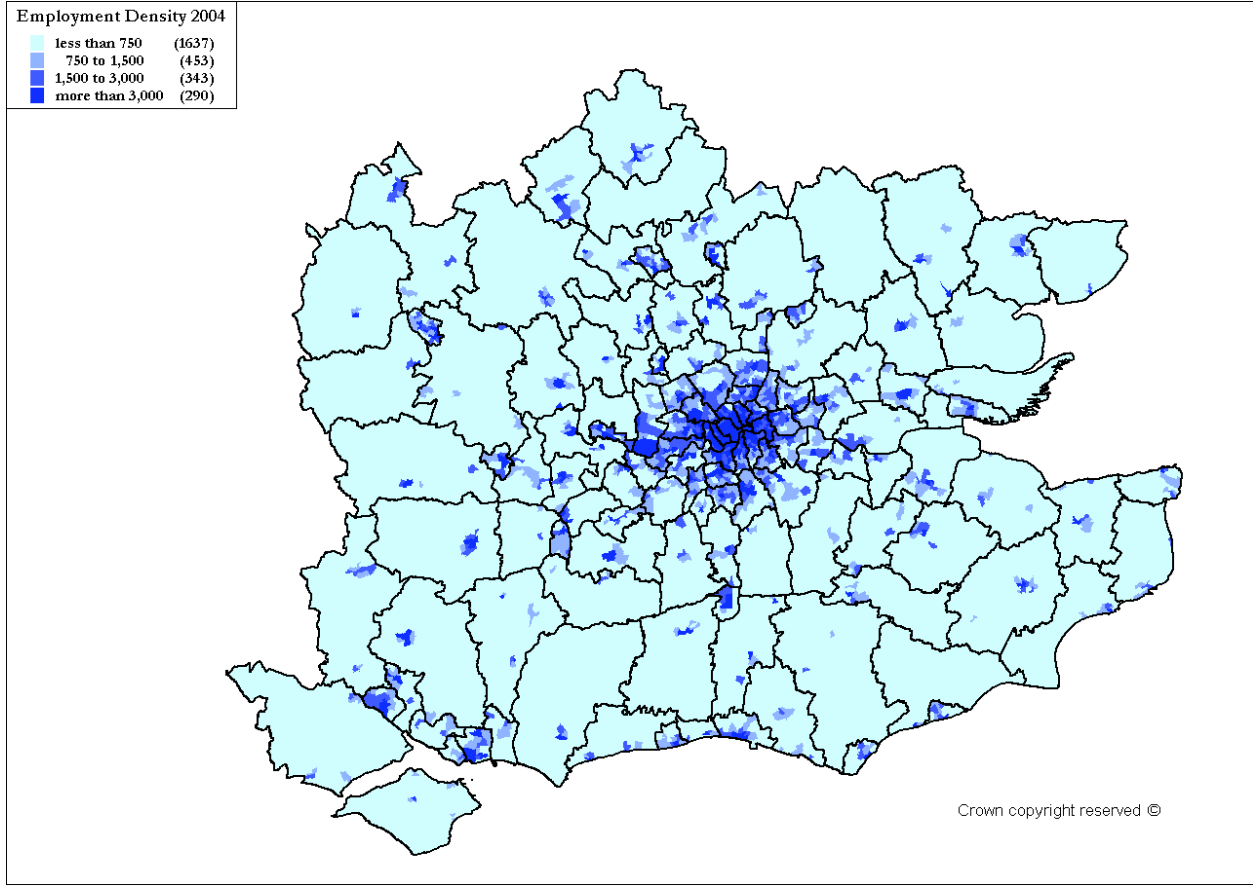
Map 3: Percentage change in population density, 1991 to 2004



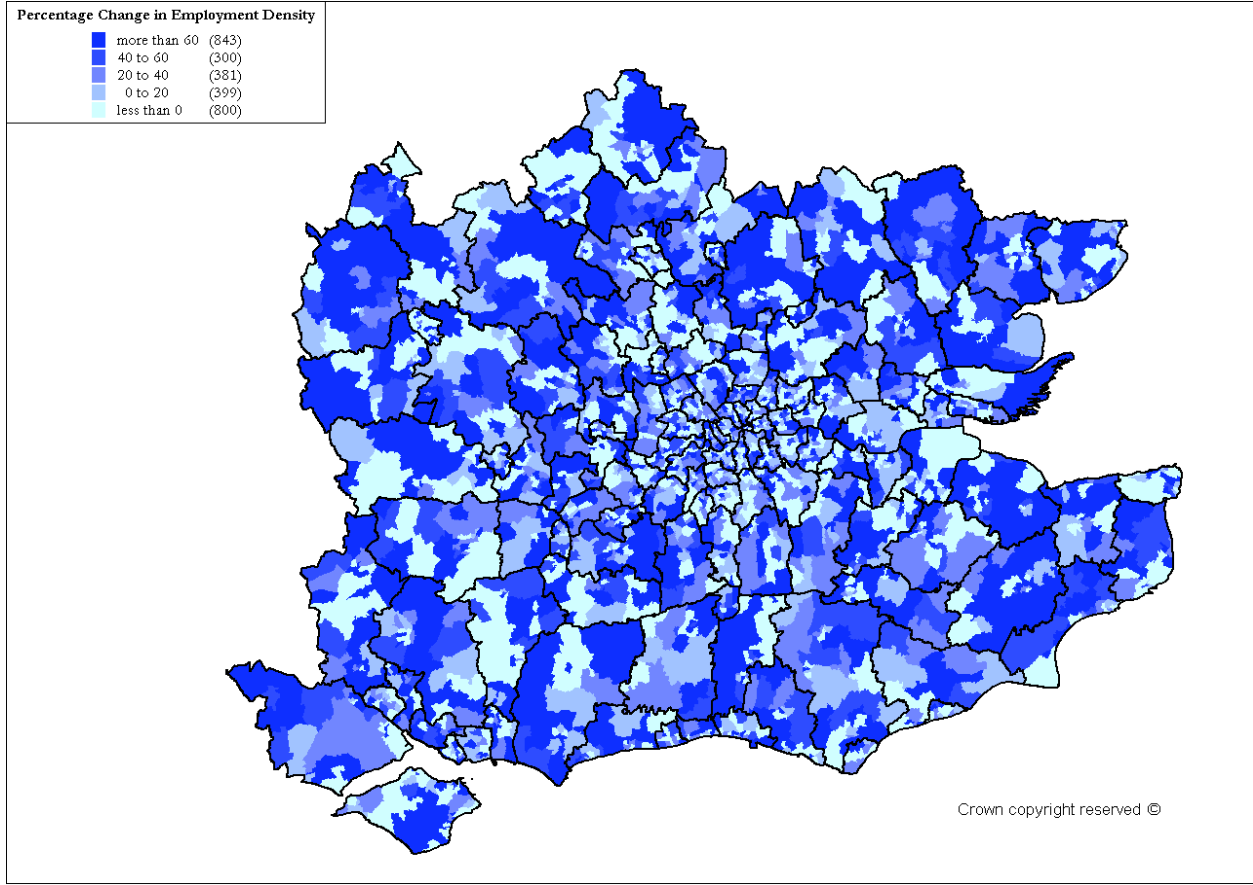
Map 4: Employment density, 1991



Map 5: Employment density, 2004



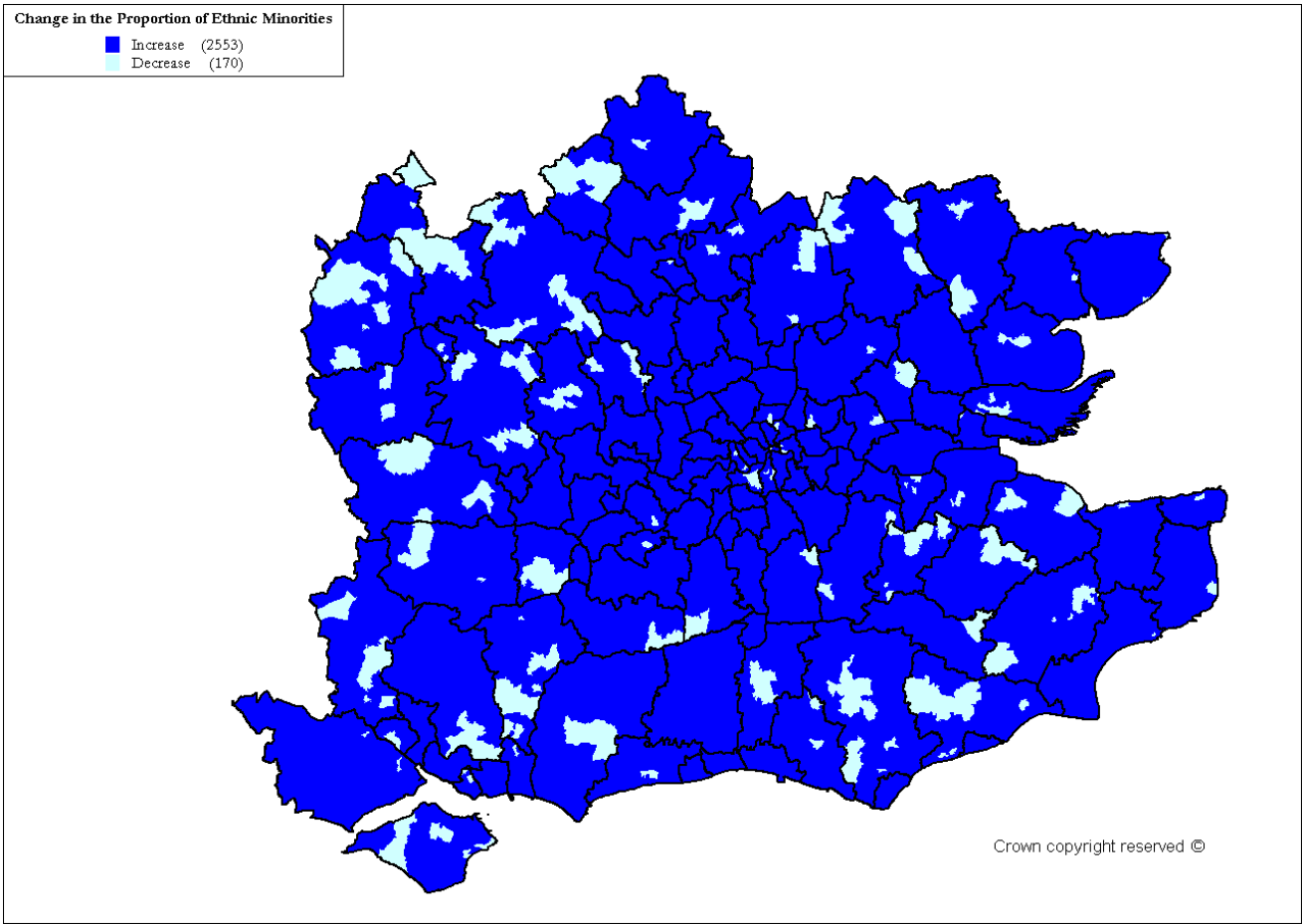
Map 6: Percentage change in employment density, 1991 to 2004



The other variables that show an increase are age and the proportion of ethnic minorities. It is expected that age will increase across the wards following from national trends. The increase in the proportion of ethnic minorities raises the issue of how are they spread geographically and whether the increase is simply an increase in wards where the proportion of ethnic minorities is already very high. Map seven shows that the proportion of ethnic minorities has increased across the regions considered; most consistently in wards in and around London. The majority of wards outside of London, however, saw an increase in the proportion of ethnic minorities.

These changes are reflected in the summary statistics shown in table three, where the proportion of ethnic minorities rose over the period from an average of 6.8 per cent in 1991 to an average of 9.8 per cent in 2004.

Map 7: Change in the proportion of ethnic minorities, 1991 to 2001



The only variable where we see a drop is unemployment. This is expected as the UK as a whole has experienced low levels of unemployment over the past few years.

Table 3: Summary statistics - Area Variables

	Average	Standard Deviation	Inter-quartile Range
Population Density (People per sq. km.)			
1991	2683.6	2964.4	3408.2
2004	3032.9	3341.1	3907.8
Employment Density (Employees per sq. km.)			
1991	1393.3	5146.9	1155.4
2004	1713.6	6363.2	1245.4
Average Age (Number of years)			
1991	35.2	2.8	3.3
2004	38.6	4.5	5.1
Unemployment (Proportion of economically active people)			
1991	7.8	4.1	4.3
2004	4.0	2.5	2.4
Proportion of Ethnic Minorities (Proportion of total population)			
1991	6.8	10.8	5.5
2004	9.8	14.0	8.7
Index of Multiple Deprivation (Index value)			
2004	15.2	11.8	12.9
Area (In sq. km.)			
2004	10.2	16.1	8.7

We now move to consider which variables to use for the area grouping. It is important to mention that we are interested in both how the characteristics of the wards have changed over the period 1991 to 2004 and how the characteristics vary across the study area. A variable showing no significant change or changes similar to other variables would not add value to the analysis.

The process of elimination involves a combination of statistical techniques and expert judgment to decide which variables to include. We pair each variable with every other variable in a two way scatter plot; this shows the relationship between each of the variables and is similar to a correlation coefficient. The results are shown in figure 1 for the 1991 data and figure 2 for the 2004 data. This shows that unemployment and the IMD are highly correlated with each other, with a correlation coefficient of 0.90 in 1991 and 0.92 in 2004. We decide to include unemployment since we have data at both dates while the data for the IMD is only available for 2004; the high correlation between the two variables, especially in 2004, means that unemployment will act as a proxy for deprivation. However, the IMD will not be ignored and we will study its variation across the groups formed by the clustering.

Another variable that is highly correlated with unemployment is the proportion of ethnic minorities with a correlation coefficient of 0.69 in 1991 and 0.69 in 2004. Including the proportion of ethnic minorities in the clustering as well as unemployment will therefore not add significant value and we decide to exclude it from the clustering but as with the IMD we will address its variations post-clustering.

Figure 1: Scatter Plots for Area Variables - 1991

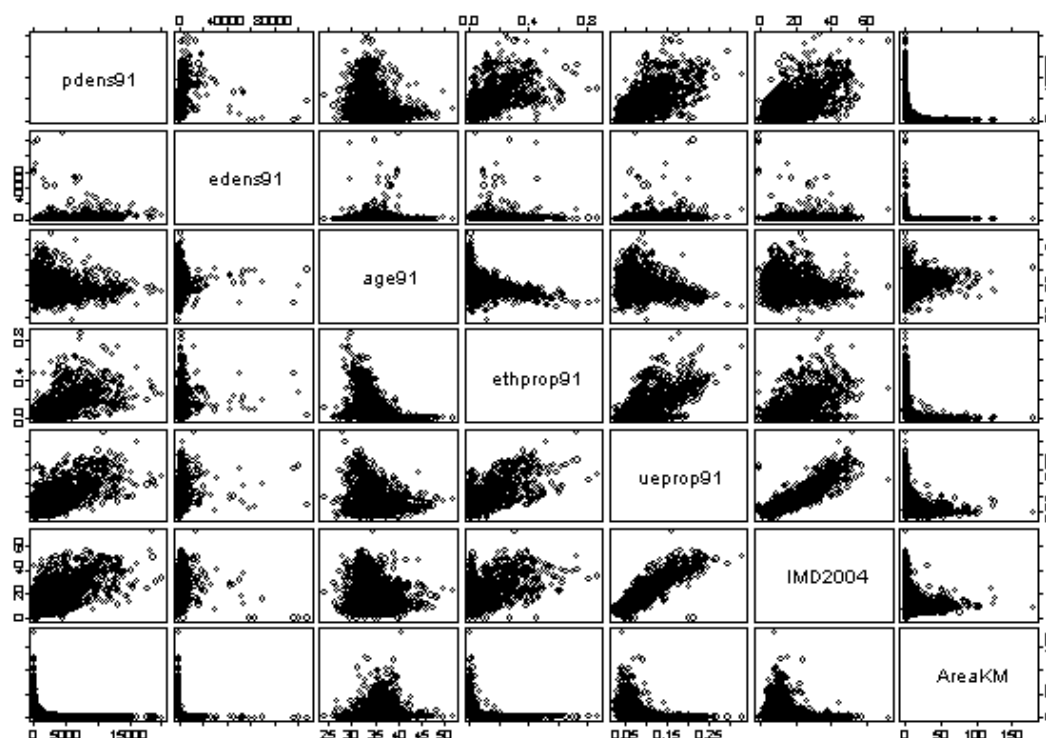


Figure 2: Scatter Plots for Area Variables - 2004

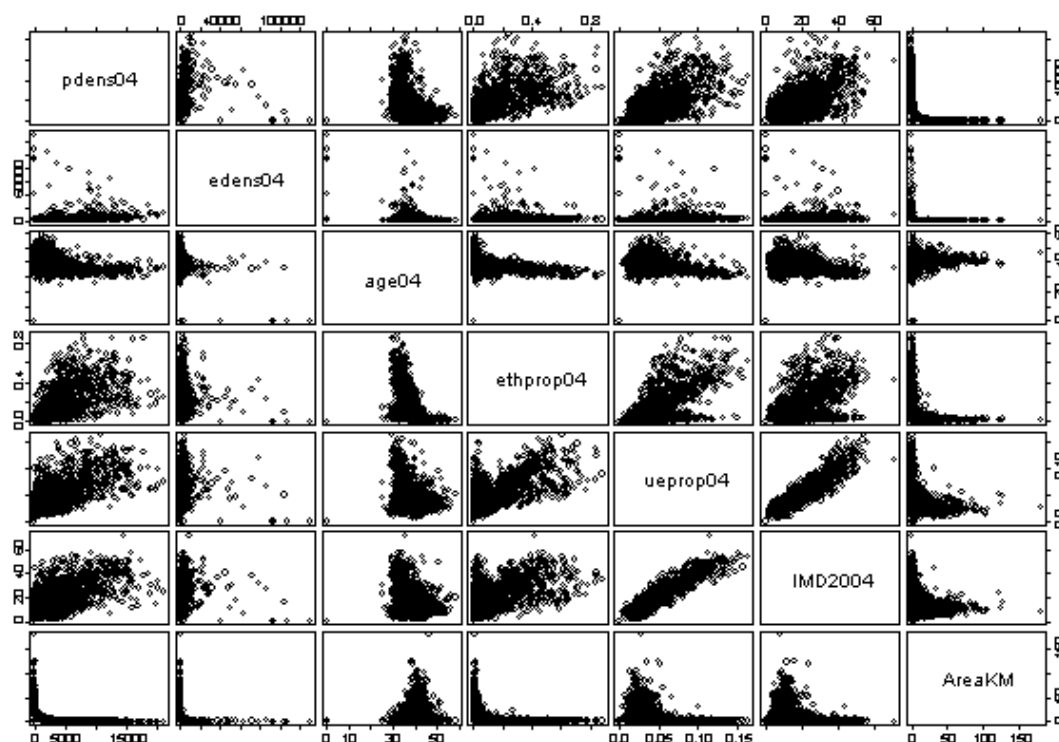


Table 4: Correlation Matrix for Area Variables - 1991

	Population Density	Employment Density	Age	Ethnic Minorities	Unemployed	IMD2004	Size
Population Density	1.00						
Employment Density	0.24	1.00					
Age	-0.27	-0.01	1.00				
Ethnic Minorities	0.66	0.24	-0.37	1.00			
Unemployed	0.69	0.26	-0.29	0.69	1.00		
IMD2004	0.64	0.15	-0.24	0.58	0.90	1.00	
Size	-0.44	-0.14	0.17	-0.27	-0.30	-0.25	1.00

Table 5: Correlation Matrix for Area Variables - 2004

	Population Density	Employment Density	Age	Ethnic Minorities	Unemployed	IMD2004	Size
Population Density	1.00						
Employment Density	0.25	1.00					
Age	-0.39	-0.32	1.00				
Ethnic Minorities	0.66	0.17	-0.43	1.00			
Unemployed	0.67	0.15	-0.30	0.69	1.00		
IMD2004	0.63	0.15	-0.31	0.60	0.92	1.00	
Size	-0.45	-0.14	0.25	-0.28	-0.29	-0.25	1.00

Finally, therefore, to describe the type of *area* a ward is we use four variables:

- population density (in people per square kilometre)
- employment density (in employees per square kilometre)
- average age
- unemployment and
- size (area in square kilometres)

For the population density, average age and unemployment variables, the data is available at two points, 1991 and 2001; from the 1991 and 2001 censuses. For the employment density the data is available at two points, 1991 and 2004; from the Annual Business Inquiry. For area, we use one point, 2004; the size does not change since all the data is adjusted for 2004 boundaries. It is worth noting that there were twelve wards with missing data in some area variables; eight wards had missing data for population density 2001 and age 2001, and four wards had missing data for age 1991. For these wards, we use the average in the district the ward belongs to for all missing variables except for the population density where we work out the population as a proportion of the local district using the ward's area proportion.

3.2 Employment Variables

For the type of *employment* data, we begin by grouping the two-digit SIC classifications, based on the 1992 definitions, into sixteen broad industrial sectors.¹ The data is obtained from the Annual Business Inquiry and grouping the two-digit SIC classifications into broader sectors is concerned with the geographic proximity of each SIC classification in relation to the rest; as opposed to a conventional approach of grouping these in relation to their function.

We then apply a technique called hierarchical clustering² to decide which of these broad industrial sectors could be grouped to form an even broader employment structure consisting of six employment groups. Again, this approach is location-specific rather than function-specific. The final six employment groups are:

Industrial	Manufacturing, construction and other retail
Public Services	Education, health and public administration
Retail and Leisure and recreation	Retail trade, hotels and restaurants, real estate activities
Business Support Services business activities	Financial, computing, research and development and other
Transport and Utilities electricity and other utilities	Air transport, land transport, supporting auxiliary transport,
Other	All other areas of employment not otherwise covered

In the employment analysis we use the proportions of employment in each ward rather than the absolute levels. This puts all wards on the same footing, reducing the influence of the few central wards with particularly high employment levels.

As for the *area* variables, we map each employment group to see if there are any clear groups on the map. The maps are shown in Appendix III. Consider, for instance, the map for the Transport and Utilities employment group, this shows a high proportion of people employed in this group in the Heathrow Village ward and the surrounding wards; a similar group can be seen around the Gatwick area. The Business Support Services group, on the other hand, is spread widely across the map; this is not surprising and we can see how groups form around specific areas.

The category 'Other' is made up of sectors of employment that do not fall neatly into the first five categories. This means that we would not necessarily expect types of employment in 'Other' to be located near to each other. For this reason, and for the fact that we cannot group on variables that sum to one, we only cluster on the first five categories. It is also worth noting that employment in this group forms no more than 5 per cent of overall employment.

Table six provides summary statistics for the employment variables. We find that the proportion of people working in the Industrial, Public Services and Transport and Utilities employment groups fell between 1991 and 2004, while the proportion of people working in the Retail and Leisure and Business Support Services increased in the same period.

¹ Details are provided in Appendix II

² Technical details of this technique are provided in Appendix I

Table seven provides a summary of changes in total employment and for each employment group between 1991 and 2004. These results show that growth in Retail and Leisure and Business Support Services employment groups was twice that for total employment and the rest of the employment groups grew less than the growth in total employment with the employment in the Industrial group falling. We also find that the proportion of people in the Retail and Leisure and Business Support Services employment groups increased in 69 per cent and 71 per cent of the wards respectively.

Whilst the proportion of employment in Public Services fell between 1991 and 2004 the results in table seven show that the total number of people employed in Public Services has increased by 21 per cent. The fall in the proportion of people employed in Public Services is due to the growth in Retail and Leisure and Business Support Services employment groups outstripping all other employment areas and does not represent actual falls in levels of Public Sector employment.

Table 6: Summary Statistics - All Variables

	Average	Standard Deviation	Inter-quartile Range
<i>Employment Variables - 1991</i>			
Industrial (%)	26.1	17.3	23.6
Public Services (%)	26.9	16.9	22.1
Retail and Leisure (%)	19.8	13.0	17.4
Business Support Services (%)	12.1	12.2	13.2
Transport and Utilities (%)	4.7	6.6	5.0
<i>Employment Variables - 2004</i>			
Industrial (%)	21.3	14.2	20.0
Public Services (%)	27.9	17.8	23.0
Retail and Leisure (%)	25.9	13.1	18.0
Business Support Services (%)	17.0	12.1	13.0
Transport and Utilities (%)	3.9	6.4	3.0

Table 7: Employment Changes, 1991 and 2004

	1991	2004	% change
Total Employment	7,151,742	8,910,161	24.6
Industrial	1,564,552	1,521,695	-2.7
Public Services	1,714,378	2,075,262	21.1
Retail and Leisure	1,409,277	2,291,683	62.6
Business Support Services	1,317,224	2,113,912	60.5
Transport and Utilities	418,057	460,759	10.2

4. Group Results

4.1 Area Groups

We start by deciding on the optimum number of groups for the area data by using qualitative judgment and a statistic known as Dunn's coefficient. This coefficient tells us how well a set of data is classified into different groups. The higher the value, the better the clustering is. In figure 3, the highest Dunn's coefficient comes from a clustering with three groups. However, it is not until we get up to five groups that the Dunn's coefficient starts to fall significantly. We also experiment with the data using different values and find that the optimum number of groups is five.

Figure 3: Standardised Dunn's Coefficient for 2004 Area Dataset

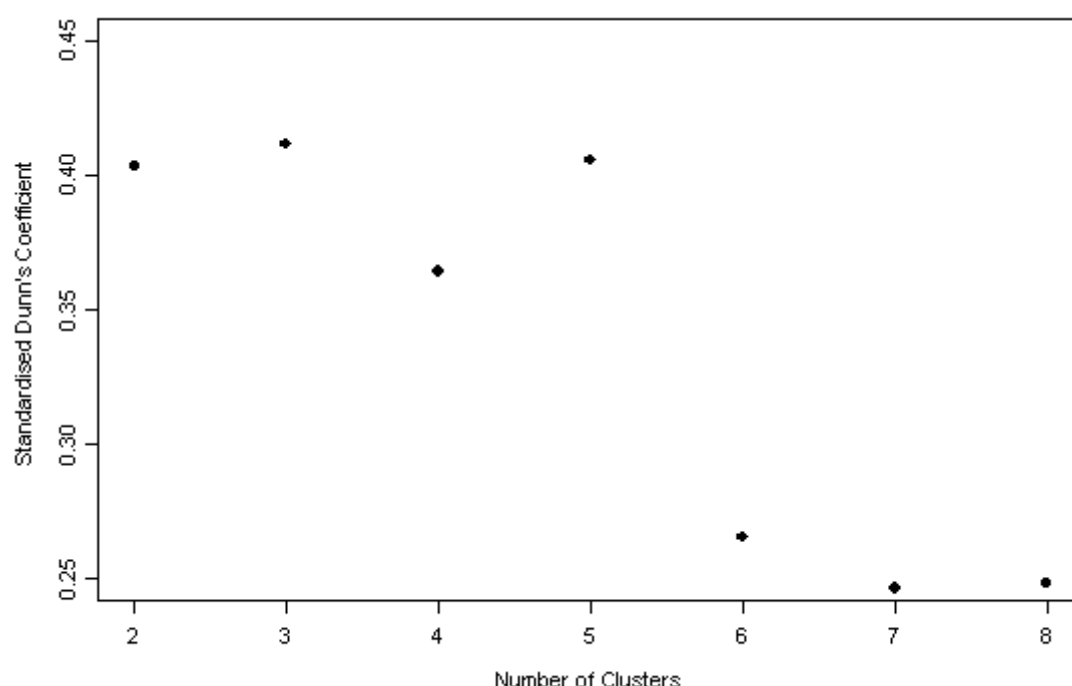


Table eight shows group centres for fuzzy clustering with the 1991 data and the relevant maps are shown in Appendix IV. The group centres represent an average of the ward variable values weighted according to the associated memberships. The process of choosing the group centres is an iterative one and we carry out fifty runs to confirm that the values obtained are stable. It is important to note that, since there are over 200 wards in each group, not all wards in any one group will have the same characteristics in all the variables as the ward used to describe the group. Clear examples of this are the wards within the City of London. The characteristics of these wards are considerably different from the rest of the wards and they have nearly equivalent degrees of membership in each of the groups, however, since there are only nine of these wards, it is not possible for these to form their own group.

As expected (from the rationale behind defining ward boundaries) there is a clear and negative relationship between area size and population density. The rest of the results show that high density areas are on average characterised by higher deprivation and a younger population; with unemployment being used as a proxy for deprivation. Group 1 is characterised by low population and employment densities, an older population and low deprivation; the majority of wards belonging to this group are outside the borders of London with only two London wards, Darwin in the South East and Upminster in the East of London.

Group 2 has similar age and deprivation characteristics as group 1 but wards in this group are much smaller on average and therefore have higher population and employment densities; there is a high concentration of these wards in the districts bordering London especially on the Western and South Western borders of London, such as South Buckinghamshire and Elmbridge. Moving to groups 3, 4 and 5, there is a large concentration of wards belonging to these groups in various London boroughs; a number of outer London Boroughs, such as Richmond-upon-Thames, contain a large number of wards belonging to group 3 whilst other, such as Ealing, are mainly dominated by group 4 wards and inner London boroughs such as Islington and Southwark are dominated by the mostly dense and deprived wards belonging to group 5.

Table 8: Group Centres for Fuzzy Clustering with 1991 Area Data

Group	Population Density (people per sq. km)	Employment Density (people per sq. km)	Average Age	Unemployment (%)	Size (sq. km)	Group size
1	200.4	162.2	36.3	5.0	37.9	456
2	1211.5	556.0	36.4	6.0	7.6	872
3	2802.8	1259.2	34.9	7.0	3.7	649
4	4171.5	1674.3	33.8	10.0	2.8	512
5	9613.6	4164.9	32.8	18.0	1.4	234

Table nine shows group centres for fuzzy clustering with the 2004 data. The relationship between the variables is similar to the one we saw earlier for the 1991 data but we see a change in the group centres; we see an increase in the population density in wards belonging to groups 3, 4 and 5 but decrease in groups 1 and 2, and an increase in employment density and average age across all the groups. We also find a change in the memberships of each group, more specifically an increase in the wards belonging to groups 3, 4 and 5 and a decrease in the number of wards belonging to groups 1 and 2. We can see on the associated map in Appendix IV that groups 3, 4 and 5 have spread into wards in inner and outer London boroughs.

Table 9: Group Centres for Fuzzy Clustering with 2004 Area Data

Group	Population density (people per sq. km)	Employment Density (people per sq. km)	Average Age	Unemployment (%)	Size (sq. km)	Group size
1	193.1	202.8	40.9	3.0	38.6	436
2	1174.3	633.9	40.7	3.0	8.5	779
3	2925.1	1433.6	38.4	3.0	3.7	712
4	4890.7	1984.5	36.4	5.0	2.5	544
5	10518.2	4831.1	33.8	9.0	1.3	252

To enable a more comparable analysis we also project the 2004 data on the 1991 group centres and the 1991 data on the 2004 group centres. The results are shown in table ten and the relevant maps in Appendix IV. For the 2004 data projected on the 1991 group centres, we find that, when compared to the 1991 data, there are a number of interesting changes; the first occurs in group 2, with membership increasing substantially in the group, and, second, is a substantial drop in membership of groups 4 and 5. This suggests that a substantial number of wards have moved membership to become part of the least deprived groups; this may be a result of the decrease in unemployment across all the wards.

When looking the opposite way by projecting the 1991 data on the 2004 group centres, we find similar results. Membership of groups 2 and 3 is increasing and that in groups 4 and 5 is falling. Membership of group 1 increases slightly.

Table 10: Size of Membership Using Different Group Centres - Area Data

Group	1991 Data	2004 Projected on 1991	1991 Projected on 2004	2004 Data
1	456	458	398	436
2	872	1323	332	779
3	649	646	263	712
4	512	198	1203	544
5	234	98	527	252

The other important aspect of the fuzzy clustering technique is that it allows each ward some degree of membership of each of the groups; for example, a ward may have 30 per cent membership of group 1, 20 per cent in group 2, 10 per cent in group 3, 5 per cent in group 4 and 1 per cent in group 5, the group the ward would be defined as belonging to is the one which it has the highest degree of membership of, in this example it would be group 1.

Tables eleven and twelve and the maps in Appendix IV illustrate the results of the fuzzy clustering technique; the maps show the percentage membership of wards in each of the groups and the tables show the average degree of memberships in each group. Table thirteen shows the difference in the membership between 1991 and 2004. There are two interesting observations when looking at these tables; the first we can see when we consider the diagonal in table thirteen, this shows that wards belonging to groups 4 and 5 have increased their degree of membership of that group while, in general, decreased their degree of membership of the other groups; in other words, the characteristics of the wards in groups 4 and 5 have become better defined on average. For example, consider group 4, the average change in membership of that group was an increase of 2.4 percentage points. Similarly, the average change in membership of group 5 is also positive. To varying degrees, wards in all the groups have increased their membership in the group they belong to; however, wards in groups 2 and 3 have also had their average memberships increase in other groups.

The second observation is about the range of group memberships within each group. This also provides a check for us in how much these groups differentiate the characteristics of the wards. If we consider any of the groups we find that the degree of membership of the other groups falls the furthest the groups are away from the group we are considering. This finding can be seen in tables eleven and twelve. For example looking at the first row of table eleven, we can see that the degree of membership falls from left to right. Similarly in row five, it falls from right to left.

Table 11: Degree of Membership (%), 1991

	Percentage Membership of the Groups				
Group	1	2	3	4	5
1	67.7	24.3	7.3	4.6	2.1
2	8.8	54.9	16.1	6.1	1.9
3	6.1	28.2	49.0	19.0	3.9
4	4.7	16.1	27.2	52.6	8.0
5	3.3	7.6	7.8	19.1	60.4

Table 12: Degree of Membership (%), 2004

	Percentage Membership of the Groups				
Group	1	2	3	4	5
1	69.0	16.6	8.4	4.4	1.5
2	12.9	55.7	21.9	7.8	1.8
3	5.0	22.0	51.8	18.6	2.6
4	3.7	9.6	22.0	55.0	9.6
5	3.8	6.4	9.6	18.6	61.7

Table 13: Differences in the Degree of Membership

	Change in Membership				
Group	1	2	3	4	5
1	1.3	-11.4	-2.3	-0.9	1.7
2	7.8	0.8	5.9	3.5	4.5
3	2.3	-6.3	2.8	3.0	5.7
4	-0.3	-8.3	-8.6	2.4	10.6
5	-1.8	-5.8	-5.2	-9.5	1.3

4.2 Employment Groups

The second set of data deals with the employment characteristics of the wards. As before, we used qualitative judgment and Dunn's coefficient to decide on the number of groups to use, the results are shown in figure three. The highest Dunn's coefficient comes from clustering with three groups. We named these groups Public Sector, Commercial and Industrial due to the relative locations of their group centres within the five employment types. The associated map for the 1991 employment clustering can be found in Appendix IV.

Figure 4: Standardised Dunn's Coefficient for 2004 Employment Dataset

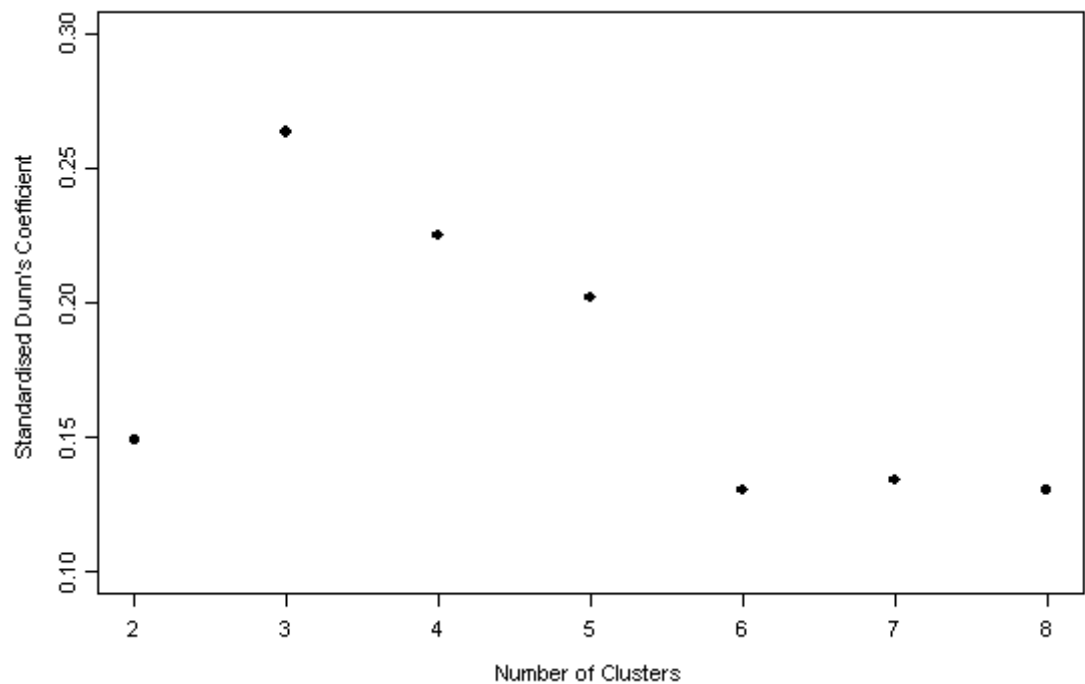


Table fourteen shows the group centres for fuzzy clustering with 1991 employment data. These show a clear distinction in the employment structure of each group. All three groups can be found throughout the whole area which was analysed, but particular concentrations can also be found. The Public Sector group is scattered around the geography of the regions considered and does not dominate any considerable number of districts. The Commercial group dominates the wards in Inner London with strong representation in the Outer London boroughs. This group also dominates the economic make-up in many districts in the South Eastern region; examples include New Forest, Windsor and Maidenhead, Surrey Heath and Guildford. The Industrial group is strongly represented in the Eastern part of the regions considered; a heavy concentration of wards belonging to the Industrial group can be seen in the counties of Essex, to the North East of London, and Kent, to the South East of London.

Table 14: Group Centres for Fuzzy Clustering with 1991 Employment Data - % of employment

Group	Industrial	Public Services	Retail and Leisure	Business Support Services	Transport and Utilities	Group Size
Public Sector	17.0	42.0	19.0	10.0	4.0	836
Commercial	20.0	23.0	26.0	17.0	5.0	950
Industrial	43.0	17.0	15.0	9.0	5.0	937

Table fifteen shows the group centres for fuzzy clustering with 2004 employment data. From these, we can see a change in the group centres. These changes reflect those already displayed in the summary statistics earlier; a fall in the proportion of industrial employment across all the groups, an increase in the proportion of Retail and Leisure and Business Support Services employment across all groups and no significant changes in the proportion of Public Services employment. We also see an increase in the membership of the Commercial group with a decrease in the membership of the other two groups.

With regards to the geography, wards belonging to the Public Sector group remain scattered without any clear concentrations. Wards belonging to the Commercial group become more concentrated in the Western part of the regions considered, while wards in London remain dominated by those belonging to this group. The Industrial group becomes more concentrated in the Eastern part of the regions considered with memberships growing in that part.

Table 15: Group Centres for Fuzzy Clustering with 2004 Employment Data - % of employment

Group	Industrial	Public Services	Retail and Leisure	Business Support Services	Transport and Utilities	Group Size
Public Sector	14.0	47.0	21.0	12.0	3.0	811
Commercial	16.0	22.0	35.0	20.0	3.0	1057
Industrial	35.0	17.0	21.0	17.0	5.0	855

As with the area data, we project the 2004 employment data using the 1991 group centres and vice versa. The results can be found in table sixteen and Appendix IV. Comparing the 2004 employment data projected onto the 1991 group centres, we can see the decline in membership of both the Public Sector and the Industrial groups, with membership of the Commercial group growing substantially. These results are mirrored on the map, where the red wards dominate. The Industrial group shows the largest decline over the period, this confirms the results obtained earlier in the summary statistics.

When we project the 1991 employment data using the 2004 group centres, we get similar results. Membership of the Industrial group declined heavily while the Commercial group remains the dominant group and the only one growing between 1991 and 2004.

Table 16: Size of Membership Using Different Group Centres - Employment Data

Group	1991 Data	2004 Projected on 1991	1991 Projected on 2004	2004 Data
Public Sector	836	849	807	811
Commercial	950	1309	694	1057
Industrial	937	565	1222	855

As for the area data, we look at the degree of membership of each group and the changes between 1991 and 2004. Tables seventeen, eighteen and nineteen and the maps in Appendix IV show these results. This shows that the average percentage membership of the wards has increased in the group to which the wards belong in the period 1991 to 2004 for the Public Sector and Commercial groups but has decreased in the Industrial group. Also, on average, wards have a smaller percentage membership of the Industrial group than any other group when they belong to the non-Industrial groups. The high increase in percentage membership of the Public Sector group in wards belonging to that group reflects the recent trend of service consolidation in the public sector; this is also confirmed by the decrease in the number of wards belonging to the Public Sector group.

Table 17: Degree of Membership, 1991

Group	Percentage Membership of the Groups		
	Public Sector	Commercial	Industrial
Public Sector	63.7	19.7	16.6
Commercial	21.1	55.4	23.5
Industrial	18.4	22.5	59.2

Table 18: Degree of Membership, 2004

Group	Percentage Membership of the Groups		
	Public Sector	Commercial	Industrial
Public Sector	68.6	18.0	13.4
Commercial	19.0	59.8	21.2
Industrial	14.5	21.7	63.8

Table 19: Differences in the Degree of Membership

Group	Change in Membership		
	Public Sector	Commercial	Industrial
Public Sector	4.9	-1.6	-3.2
Commercial	-2.1	4.5	-2.3
Industrial	-3.9	-0.7	4.6

4.3 Combining Employment and Area Memberships

In this section we combine the results of the separate employment and area classifications to obtain a combined level of description of each ward's economic and demographic characteristics.

Wards are classified according to both their employment and area characteristics in time periods, 1991 and 2004. The results are summarised in tables twenty, twenty one and twenty two. These show that, between 1991 and 2004, the fall in Public sector wards is mainly attributable to the decrease in the number of Public sector wards in area groups 1 and 2. The number of wards in the Commercial sector increased across area groups 3, 4 and 5, with the highest increases in group 4. Finally, the number of Industrial wards increased in area group 1 but fell in the all other groups.

Table 20: Cross-Classification of Groups, 1991

	Public Sector	Commercial	Industrial	Total
Group 1	115 (4.22%)	160 (5.88%)	181 (6.65%)	456 (16.75%)
Group 2	293 (10.76%)	290 (10.65%)	289 (10.61%)	872 (32.02%)
Group 3	188 (6.90%)	231 (8.48%)	230 (8.45%)	649 (23.83%)
Group 4	151 (5.55%)	161 (5.91%)	200 (7.34%)	512 (18.80%)
Group 5	89 (3.27%)	108 (3.97%)	37 (1.36%)	234 (8.59%)
Total	836 (30.70%)	950 (34.89%)	937 (34.41%)	2723

Table 21: Cross-Classification of Groups, 2004

	Public Sector	Commercial	Industrial	Total
Group 1	69 (2.50%)	152 (5.62%)	215 (7.90%)	436 (16.01%)
Group 2	224 (8.19%)	288 (10.58%)	267 (9.84%)	779 (28.61%)
Group 3	241 (8.81%)	267 (9.92%)	204 (7.42%)	712 (26.15%)
Group 4	186 (6.90%)	219 (8.04%)	139 (5.03%)	544 (19.98%)
Group 5	91 (3.34%)	131 (4.81%)	30 (1.10%)	252 (9.25%)
Total	811 (29.78%)	1057 (38.82%)	855 (34.41%)	2723

Table 22: Differences in Cross-Classifications

	Public Sector	Commercial	Industrial	Total
Group 1	-46 (-66.67%)	-8 (-5.26%)	34 (+15.81%)	-20 (-4.59%)
Group 2	-69 (-30.80%)	-2 (-0.69%)	-22 (-8.24%)	-93 (-11.94%)
Group 3	53 (+21.99%)	36 (+13.48%)	-26 (-12.75%)	63 (+8.85%)
Group 4	35 (+18.82%)	58 (+26.48%)	-61 (-43.88%)	32 (+5.88%)
Group 5	2 (+2.20%)	23 (+17.56%)	-7 (-23.33%)	18 (+7.14%)
Total	-25 (-3.08%)	107 (+10.12%)	-82 (-9.59%)	

It is very striking that group 1 areas, which are large and less dense, are becoming both less commercial and public sector - maybe office - oriented, with the result that they are more industrial. This is likely to mean that they have shared less in the growth in employment over these years. This may be one reason why there are also fewer areas whose main membership is of these groups. Notice also that the Public sector has grown in more dense and deprived areas - this may reflect the investment in trying to reduce deprivation.

4.4 Indices of Deprivation

In this section we look at the relationship between the various employment and area groups and the separate deprivation indices. First we look at the correlations of each index of deprivation with membership of various employment and area groups; this is shown in table twenty three.

Table 23: Correlation of membership of employment and area groups with indices of deprivation

Area groups	Income	Employment	Health	Education	Housing	Crime	Living Environment
Group 1	0.05	0.15	0.06	0.14	0.00	-0.05	0.00
Group 2	-0.25	-0.31	-0.24	-0.26	-0.01	0.08	-0.08
Group 3	-0.32	-0.29	-0.25	-0.12	-0.47	-0.19	-0.30
Group 4	-0.03	-0.09	0.02	-0.09	0.02	0.11	-0.03
Group 5	0.44	0.33	0.31	0.16	0.24	0.32	0.39
Employment groups							
Public Sector	0.13	0.16	0.18	0.15	-0.05	0.06	0.04
Commercial	-0.03	-0.03	-0.06	-0.17	0.01	0.09	0.16
Industrial	-0.12	-0.15	-0.14	-0.01	0.05	-0.15	-0.19

The results in table twenty three show no strong correlation between any of the indices and membership of any of the employment groups. However, we can see some relationships with the area groups: income and living environment deprivation show a mild positive correlation with the group 5, the most dense and deprived group.

We also consider the average deprivation indices for each employment and area group; the results are shown in table twenty four below.

Table 24: Averages of each deprivation index for wards belonging to each employment or area group - Index Values³

Area Groups	Income	Employment	Health	Education	Housing	Crime	Living Environment
Group 1	0.06	0.04	-1.13	8.41	32.71	-0.80	11.10
Group 2	0.07	0.05	-0.95	11.18	22.70	-0.59	9.91
Group 3	0.08	0.06	-0.70	14.91	19.60	-0.34	13.82
Group 4	0.15	0.09	-0.11	23.21	25.02	0.22	24.42
Group 5	0.26	0.14	0.54	20.53	37.27	0.77	41.69
Employment Groups							
Public Sector	0.12	0.08	-0.43	17.14	24.66	-0.22	17.23
Commercial	0.10	0.07	-0.64	12.80	25.56	-0.19	19.58
Industrial	0.09	0.06	-0.74	15.62	25.60	-0.42	13.47
Ward Minimum	0.01	0.01	-2.67	0.21	3.87	-2.58	0.29
Ward Maximum	0.59	0.32	2.05	77.95	55.90	1.69	66.64
Ward Mean	0.10	0.07	-0.61	15.02	25.38	-0.27	17.01
Ward Standard Deviation	0.08	0.04	0.74	11.50	9.33	0.69	12.46

The results in table twenty four show that there is some noticeable variation in deprivation indices according to area group. The variation across the deprivation indices is not surprising showing consistent trends except for housing deprivation. This is highest in group 5, the most dense and deprived, followed very closely with the group 1, the least dense and deprived. In terms of employment groups, we do not observe any consistent trends.

4.5 Proportion of Ethnic Minorities

In this section, we consider the changes in the proportion of ethnic minorities in each employment and area group. The results in table twenty five show that, on average, the proportion of ethnic minorities has increased in all the area groups. With regards to employment groups, there has also been an increase in the proportion of ethnic minorities in all the wards.

Table 25: Average proportion of ethnic minorities in each employment and area group

Area Groups	1991	2004
Group 1	1.0	1.7
Group 2	2.2	3.6
Group 3	4.5	6.0
Group 4	12.5	17.2
Group 5	28.5	37.3
Employment Groups		
Public Sector	6.9	11.1
Commercial Sector	7.6	10.6
Industrial Sector	5.7	7.5

³ For a full explanation of the construction of the Indices of Deprivation, refer to: <http://www.communities.gov.uk/documents/communities/pdf/131209>

4.6 London Wards

In this section, we consider differences between wards in London and those in the other regions considered. We begin by looking at the difference in average membership of each of the area and employment groups between wards in London and those in the other regions. We look at the figures in 2004 and compare with the 1991 figures; the 1991 figures used are those projected on the 2004 group centres to enable direct comparison. The maps in Appendix V and tables twenty six and twenty seven show these results.

In London, the results for the area groups show that there are significant changes in average membership in all the groups with average membership in groups 1, 2 and 3 falling while membership in the other two groups increasing. This suggests that population density and deprivation levels have spread out across London wards rather than being only concentrated in a few areas. In wards outside London, there has been a small decrease in average membership in area groups 1, 2 and 4, with membership in groups 3 and 5 increasing, on average.

With regards to employment groups, Public Sector membership fell outside of London but slightly increased in wards in London, while membership of the Commercial group, unsurprisingly, increases in both, though significantly more in London. Industrial membership, on the other hand, fell significantly in London and outside of London. The fall in the number of Industrial group members can be attributed to the overwhelming increase in Commercial membership.

Table 26: Changes in area and employment group memberships, London 1991 to 2004

Area groups	1991	2004	Absolute Change	% Change
Group 1	4.4	3.8	-0.6	-14.2
Group 2	14.9	12.0	-2.9	-19.4
Group 3	24.2	22.1	-2.1	-8.6
Group 4	30.7	34.0	3.3	10.7
Group 5	25.8	28.1	2.3	9.0
Employment groups				
Public Sector	33.1	33.8	0.7	2.1
Commercial	32.7	40.6	7.8	23.9
Industrial	34.2	25.7	-8.5	-24.9

Table 27: Changes in area and employment group memberships, not in London 1991 to 2004

Area groups	1991	2004	Absolute Change	% Change
Group 1	21.8	21.1	-0.7	-3.1
Group 2	32.1	31.3	-0.8	-2.4
Group 3	25.6	27.8	2.2	8.4
Group 4	17.3	16.5	-0.8	-4.8
Group 5	3.2	3.3	0.1	3.4
Employment groups				
Public Sector	32.3	31.9	-0.4	-1.3
Commercial	29.2	33.9	4.7	16.0
Industrial	38.5	34.2	-4.3	-11.1

Next, we look at the differences in deprivation between London and wards outside London. The results are shown in tables twenty eight and twenty nine. For group 1 wards, there are no significant differences between wards in London and outside London except for housing deprivation which is higher in wards outside London. For the other groups, it is clear that the sources of deprivation in London are housing and the living environment while education and health are the main sources of deprivation outside London.

For the employment groups, we can see some differences between London and wards outside London. Again, London wards' deprivation comes from housing, crime and the living environment. London commercial wards are least deprived on all criteria except for housing and the living environment. For wards outside London, the living environment in Commercial wards is the criterion on which these wards are most deprived.

The ward averages for each of the deprivation indices in London and outside confirm the above trends.

Table 28: Averages of each deprivation index for wards in London - Index Values

	Income	Employment	Health	Education	Housing	Crime	Living Environment
Group 1	0.05	0.05	-1.09	8.04	26.48	0.00	10.39
Group 2	0.08	0.06	-0.86	10.57	24.22	-0.12	13.97
Group 3	0.09	0.06	-0.72	11.50	24.07	-0.08	20.81
Group 4	0.15	0.09	-0.19	15.25	29.20	0.26	28.58
Group 5	0.26	0.14	0.53	19.89	38.42	0.79	41.92
Public Sector	0.19	0.11	0.11	18.50	32.21	0.40	30.73
Commercial	0.16	0.09	-0.18	12.88	31.36	0.38	33.33
Industrial	0.18	0.10	-0.06	20.18	29.92	0.29	25.63
Ward Minimum	0.01	0.01	-2.33	1.08	8.13	-2.15	4.74
Ward Maximum	0.59	0.31	2.05	62.20	49.60	1.55	66.64
Ward Mean	0.17	0.10	-0.07	15.98	31.37	0.37	31.12
Ward Standard Deviation	0.10	0.04	0.72	10.01	8.98	0.54	12.73

Table 29: Averages of each deprivation index for wards outside London - Index Values

	Income	Employment	Health	Education	Housing	Crime	Living Environment
Group 1	0.06	0.04	-1.13	8.42	32.74	-0.80	11.10
Group 2	0.07	0.05	-0.96	11.25	22.68	-0.62	9.70
Group 3	0.08	0.06	-0.71	15.66	19.01	-0.39	12.72
Group 4	0.16	0.10	-0.04	29.88	21.52	0.18	20.93
Group 5	0.22	0.16	0.71	27.18	25.27	0.56	39.35
Public Sector	0.09	0.07	-0.61	16.76	22.26	-0.43	12.85
Commercial	0.08	0.06	-0.85	12.88	23.29	-0.44	13.86
Industrial	0.08	0.06	-0.84	14.92	24.94	-0.52	11.62
Ward Minimum	0.01	0.01	-2.67	0.21	3.87	-2.58	0.29
Ward Maximum	0.39	0.32	1.42	77.95	55.90	1.69	65.16
Ward Mean	0.08	0.06	-0.78	14.74	23.58	-0.47	12.77
Ward Standard Deviation	0.06	0.04	0.67	11.87	8.66	0.61	8.68

4.7 Employment Growth

It is interesting to consider which groups were associated with faster growth in employment over the period. We look at the combined area and employment groups and measure the annual growth in overall employment throughout the period 1991 to 2004; the results are shown in table thirty. The first result to note is the steady growth in employment in wards belonging to the Commercial group; wards in all area groups experienced steady growth in overall employment if they were also part of the Commercial group. For the Public Sector group, the results show that there has been either slow growth or even negative growth in overall employment in the wards which belong to this group.

The interesting result is growth in employment in wards belonging to the Industrial group. The results suggest that employment has grown substantially in these wards. We explore this further to try to find the source of this growth. We find that the proportion of people employed in Retail & Leisure and Business Support Services has increased from 13.4 and 7.9 per cent in 1991 to 19.5 and 16.0 per cent respectively in 2004, whilst the proportion of people employed in the industrial sector fell from 43.8 per cent in 1991 to 37.8 per cent in 2004.

Table 30: Annual Growth in Employment, 1991 to 2004 - Percentages

	Public Sector	Commercial	Industrial	Overall
Group 1	-1.1	7.9	57.3	26.5
Group 2	-12.0	7.2	45.8	14.7
Group 3	-5.5	39.9	34.4	16.3
Group 4	33.2	77.2	-21.1	32.5
Group 5	6.8	26.5	42.2	21.8
Overall	3.3	37.9	20.8	24.6

4.8 The City of London

In this section we look at the wards belonging to the City of London. The characteristics of wards in this district are significantly different than the rest of the wards considered. There is a very small number of people living in these wards and therefore it is very difficult to compare the characteristics of these wards with all the other wards considered. The other issue is employment which is much higher in these wards than in the rest of the wards; coupled with their relatively small size, the employment density in these wards is very high.

Table thirty one shows the average membership for each of the City of London wards in all the area groups. These show that, on average, wards in the City of London have very high membership in each of the area groups; this suggests that they share some characteristic with each of the area groups and therefore it is very difficult to assign them membership in only one group.

Table 31: Average area group membership in City of London wards - Percentages, 2004

Ward	Group 1	Group 2	Group 3	Group 4	Group 5
Aldersgate	21.0	22.0	21.0	16.0	20.0
Bishopsgate	21.0	21.0	20.0	19.0	18.0
Cripplegate	19.0	22.0	23.0	14.0	22.0
Farringdon Within	22.0	22.0	21.0	19.0	16.0
Farringdon Without	20.0	23.0	25.0	16.0	17.0
Portoken	16.0	19.0	24.0	13.0	27.0
Queenhithe	21.0	22.0	21.0	19.0	17.0
Tower	21.0	22.0	21.0	19.0	17.0
Walbrook	21.0	22.0	21.0	19.0	18.0

5. Conclusions

In terms of the area characteristics of the wards, we have seen a large increase in population across the denser groups - groups 3, 4 and 5. Groups 1 and 2 also saw an increase in population but a larger increase in employment. Average age across the wards has generally increased, except for the wards belonging to group 5, the most deprived group, where the average age has fallen. The number of wards in groups 3, 4 and 5 has also risen, and these are more focused in London.

However, the most interesting change can be seen when we project the 2004 data onto the 1991 group centres (and vice-versa). These show two interesting observations, firstly, the number of wards moving into the first three groups over the period; the size of membership of groups 1, 2 and 3 increased by 9, 134 and 170 per cent respectively when we consider the 2004 group centres. Secondly, the fact that the number of wards in groups 4 and 5, the most dense and deprived, has fallen 55 and 52 per cent respectively when we consider the 2004 group centres. The main reason for this change is the overall drop in unemployment over the period; the result is that there has been a shift in what is characterised as low or high unemployment.

As for the employment data, it is important to note the change in the group centres between 1991 and 2004. Most importantly the average proportion of people employed in industrial jobs in wards belonging to the Industrial group has dropped from 21.9 per cent in 1991 to 17.1 per cent in 2004. This result is in line with the fact that overall industrial employment in the regions considered has fallen by nearly 5 per cent over the period 1991 to 2004. Employment in Public Services grew more slowly than overall employment and at ward level stayed relatively stable with no interesting changes. Employment in both Retail and Leisure and Business Support Services grew at twice the rate for overall employment, together making up over 95 per cent of total employment growth in the region considered over the period 1991-2004.

Looking at the group centres for both 1991 and 2004 employment data, the proportion of people employed in Retail and Leisure and Business Support Services has grown in all three groups; this shows the extent of the change in the employment structure of the wards in these regions. This is illustrated most when looking at the change in the membership of the Commercial group when we project the 2004 employment data onto the 1991 group centres; membership of the Commercial group grew by over 53 per cent while membership of the Public Sector stayed the same and the Industrial groups' fell by over 30 per cent.

Housing deprivation is a problem in both groups 1 and 5; it is also a problem across all area and employment groups in London wards compared to wards outside London. London wards also suffer from crime and living environment deprivation.

Finally, we found that the fastest growing wards in terms of overall employment are those belonging to both the Commercial employment group and the area group 5. This is not surprising as we found that commercial employment saw the largest increase over the period 1991 to 2004 in the regions considered. This trend is confirmed when looking at all the area groups, as those wards that also belong to the Commercial employment group saw the fastest growth in employment overall. This, however, requires a closer look as it does not address the issues of variability and the impact of similar wards on each other.

Appendix I: Technical Background

Fuzzy Clustering

We start with a data set X consisting of n observations, where each observation is a vector in d dimensions. The aim of fuzzy clustering is to divide the data into c clusters, where c can be between 2 and n . The divisions should be such that within the clusters the data have similar characteristics and the average difference between cluster characteristics is maximised.

$$X = \{x_1, x_2, \dots, x_n\} \quad x_j \in R^d \quad (1)$$

The attributability of observation x_j to cluster k is u_{kj} . With classical clustering u_{kj} can only take the value 0 or 1, but with fuzzy clustering it can take any value between 0 and 1.

$$u_{kj} \in \{0, 1\} \quad \text{Classical clustering} \quad (2)$$

$$u_{kj} \in [0, 1] \quad \text{Fuzzy clustering} \quad (3)$$

However for each type of clustering we still have the condition:

$$\sum_j u_{kj} > 0 \quad \sum_k u_{kj} = 1 \quad (4) \text{ \& } (5)$$

The objective function, whose size is to be minimised for an optimal solution is:

$$J_m(U, v) = \sum_{j=1}^n \sum_{k=1}^c (u_{kj})^m \|x_j - v_k\|^2 \quad 1 \leq m < \infty \quad (6)$$

In this equation we have U , the matrix of memberships and v_k , the centre of cluster k . The variable m determines the type of clustering that is done. When $m = 1$ and $u_{kj} \in \{0, 1\}$ the minimalization of (6) is what is known as ordinary k-means. When m takes a value greater than 1 and $u_{kj} \in [0, 1]$ we have fuzzy clustering. In this case the values of u_{kj} and v_k that minimise (6) are:

$$\hat{u}_{kj} = \left(\sum_{g=1}^c \left(\frac{\|x_j - \hat{v}_k\|}{\|x_j - \hat{v}_g\|} \right)^{\frac{2}{m-1}} \right)^{-1} \quad \forall j, k \quad (7)$$

$$\hat{v}_k = \frac{\sum_{j=1}^n (\hat{u}_{kj})^m x_j}{\sum_{j=1}^n (\hat{u}_{kj})^m} \quad \forall k \quad (8)$$

As the centres of the clusters are not known before the clustering process, the memberships cannot be calculated directly, and an iterative process has to be used. The optimal u_{kj} can be found by repeating the following process:

- (i) m and cluster number c are assumed, and a norm in equation (6) is defined appropriately (for our purposes, the standard Euclidean norm). In addition, an initial value $U^{(0)} \in M_c$ is set for U (where M_c is the space satisfying the above conditions (3), (4) and (5)). The value can be taken at random.

- (ii) The cluster centre $v_k^{(0)}$ is calculated using $U^{(0)}$ and equation (8)
- (iii) $U^{(1)}$ is calculated using $v_k^{(0)}$ and equation (7)
- (iv) Defining an appropriate norm and threshold value ϵ , the preceding steps are repeated until $\|U^{(p)} - U^{(p-1)}\| \leq \epsilon$

When the inequality in step (iv) is satisfied, we are left with the c optimal cluster centres, $v_k^{(p)}$, whose memberships $U^{(p)}$ are given by equation (7).

Hierarchical Clustering

Given a data set X consisting of n observations and an $n \times n$ distance matrix the basic process of hierarchical clustering is the following:

1. Assign each observation to its own cluster so that you have n clusters each containing one observation.
2. Find the closest pair of clusters and merge them into one cluster so that you have one less cluster.
3. Compute the distances between the new cluster and each of the old clusters.
4. Repeat steps 2 and 3 until all items are clustered into one cluster of size n .

There are several methods to calculate the distances in step 3; single-link, complete-link and average-link. In single-link clustering we consider the distance between one cluster and another as the shortest distance between any member of one cluster to any member of the other cluster. In complete-link clustering we consider the longest distance between any two members of the clusters. In average-link clustering we consider the average distance between all members of the clusters.⁴

Information criterion used in selecting the number of clusters

$$D = \frac{1}{n} \sum_{j=1}^n \sum_{k=i}^c (\hat{u}_{kj})^2 \quad \text{Dunn's Coefficient} \quad (9)$$

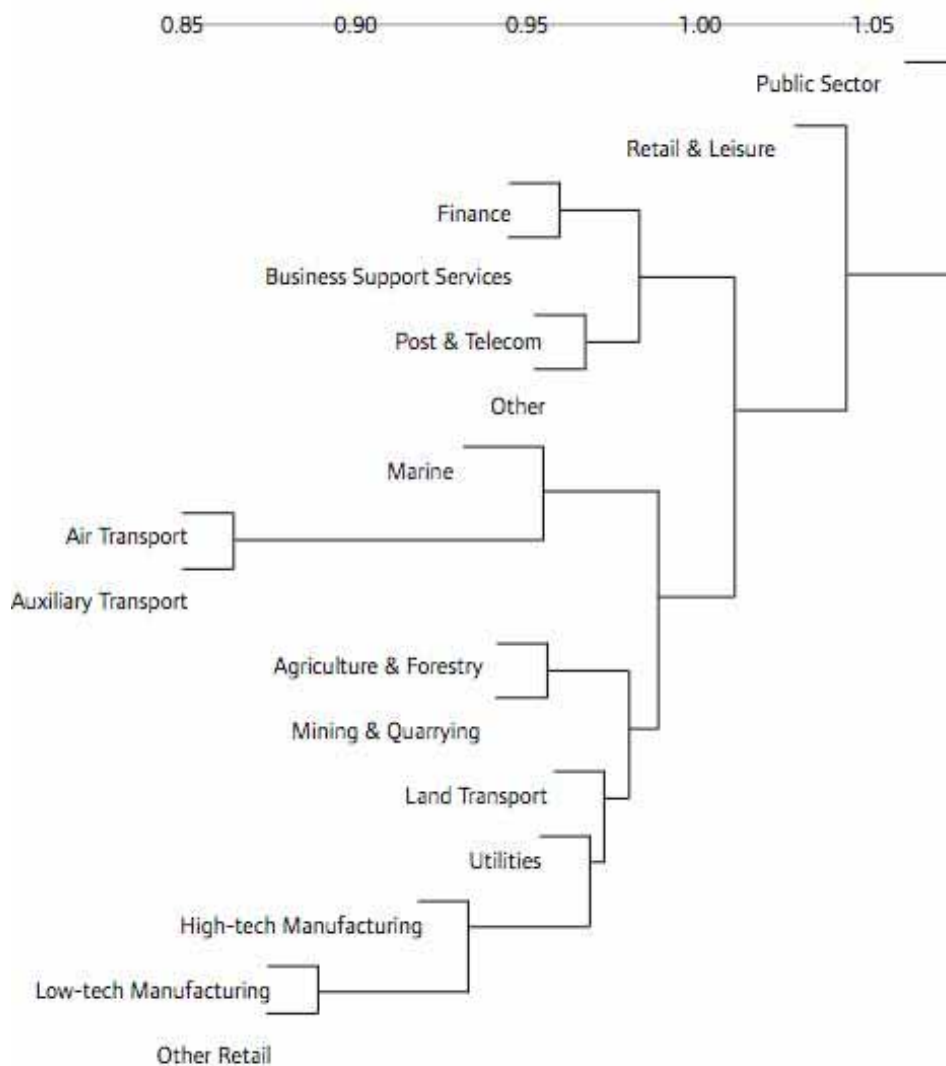
$$D_s = \frac{D - \left(\frac{1}{c}\right)}{1 - \left(\frac{1}{c}\right)} \quad \text{standardised Dunn's Coefficient} \quad (10)$$

⁴ Source: <http://www.analytictech.com/networks/hiclus.htm>
GLA Economics

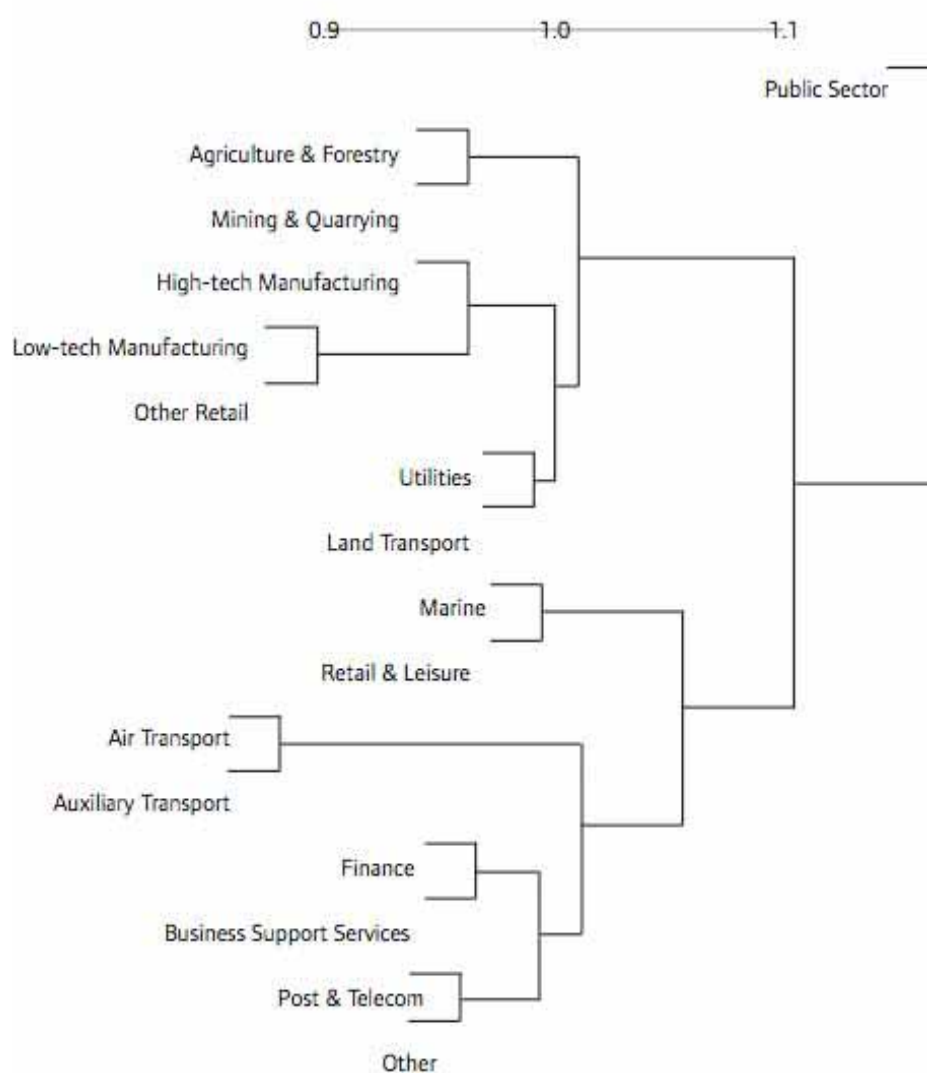
Appendix II: Industrial Grouping

SIC92 - 2 digit	Primary Industrial Group	Broad Industrial Group
01 : Agriculture, hunting, etc	Agriculture & Forestry	Other
02 : Forestry, logging, etc	Agriculture & Forestry	Other
05 : Fishing, operation fish hatcheries/farms	Marine	Other
10 : Mining coal/lignite; extraction of peat	Mining & Quarrying	Other
11 : Extraction crude petroleum/natural gas	Mining & Quarrying	Other
12 : Mining of uranium and thorium ores	Mining & Quarrying	Other
13 : Mining of metal ores	Mining & Quarrying	Other
14 : Other mining and quarrying	Mining & Quarrying	Other
15 : Manuf food products and beverages	Low-tech manufacturing	Industrial
16 : Manuf tobacco products	Low-tech manufacturing	Industrial
17 : Manuf textiles	Low-tech manufacturing	Industrial
18 : Manuf apparel; dressing/dyeing fur	Low-tech manufacturing	Industrial
19 : Tanning/dressing of leather, etc	Low-tech manufacturing	Industrial
20 : Manuf wood/products/cork, etc	Low-tech manufacturing	Industrial
21 : Manuf pulp, paper and paper products	Low-tech manufacturing	Industrial
22 : Publishing, printing, repro recorded media	Other	Other
23 : Manuf coke, refined petroleum products	Low-tech manufacturing	Industrial
24 : Manuf chemicals and chemical products	High-tech manufacturing	Industrial
25 : Manuf rubber and plastic goods	Low-tech manufacturing	Industrial
26 : Manuf other non-metallic products	Low-tech manufacturing	Industrial
27 : Manuf basic metals	Low-tech manufacturing	Industrial
28 : Manuf fabricated metal products, etc	Low-tech manufacturing	Industrial
29 : Manuf machinery and equipment nec	High-tech manufacturing	Industrial
30 : Manuf office machinery and computers	High-tech manufacturing	Industrial
31 : Manuf electrical machinery/apparatus nec	High-tech manufacturing	Industrial
32 : Manuf radio, tv/communications equipment	High-tech manufacturing	Industrial
33 : Manuf medical, precision instruments, etc	High-tech manufacturing	Industrial
34 : Manuf motor vehicles, trailers, etc	High-tech manufacturing	Industrial
35 : Manuf other transport equipment	Low-tech manufacturing	Industrial
36 : Manuf furniture; manufacturing nec	Low-tech manufacturing	Industrial
37 : Recycling	Low-tech manufacturing	Industrial
40 : Electricity, gas, steam/hot water supply	Utilities	Transport and Utilities
41 : Collection, purification/distrib of water	Utilities	Transport and Utilities
45 : Construction	Low-tech manufacturing	Industrial
50 : Sale, maintenance/repair motor vehicles	Other Retail	Industrial
51 : Wholesale trade/commission trade, etc	Other Retail	Industrial
52 : Retail trade, except of motor vehicles	Retail & Leisure	Retail and Leisure
55 : Hotels and restaurants	Retail & Leisure	Retail and Leisure
60 : Land transport; transport via pipelines	Land Transport	Transport and Utilities
61 : Water transport	Marine	Other
62 : Air transport	Air Transport	Transport and Utilities
63 : Supporting/auxiliary transport, etc	Supporting Auxiliary Transport	Transport and Utilities
64 : Post and telecommunications	Post & Telecom	Other
65 : Financial intermediation, etc	Finance	Business Support Services
66 : Insurance and pension funding, etc	Finance	Business Support Services
67 : Act auxiliary financial intermediation	Finance	Business Support Services
70 : Real estate activities	Retail & Leisure	Retail and Leisure
71 : Renting machinery/equipment, etc	Other Retail	Industrial
72 : Computing and related activities	Business Support Services	Business Support Services
73 : Research and development	Business Support Services	Business Support Services
74 : Other business activities	Business Support Services	Business Support Services
75 : Public admin/defence; compulsory SS	Public Sector	Public Services
80 : Education	Public Sector	Public Services
85 : Health and social work	Public Sector	Public Services
90 : Sewage/refuse disposal, sanitation, etc	Utilities	Transport and Utilities
91 : Activities membership organisations nec	Other	Other
92 : Recreational, cultural and sporting	Retail & Leisure	Retail and Leisure
93 : Other service activities	Retail & Leisure	Retail and Leisure
95 : Private households with employed persons	Other	Other
99 : Extra-territorial organisations/bodies	Other	Other

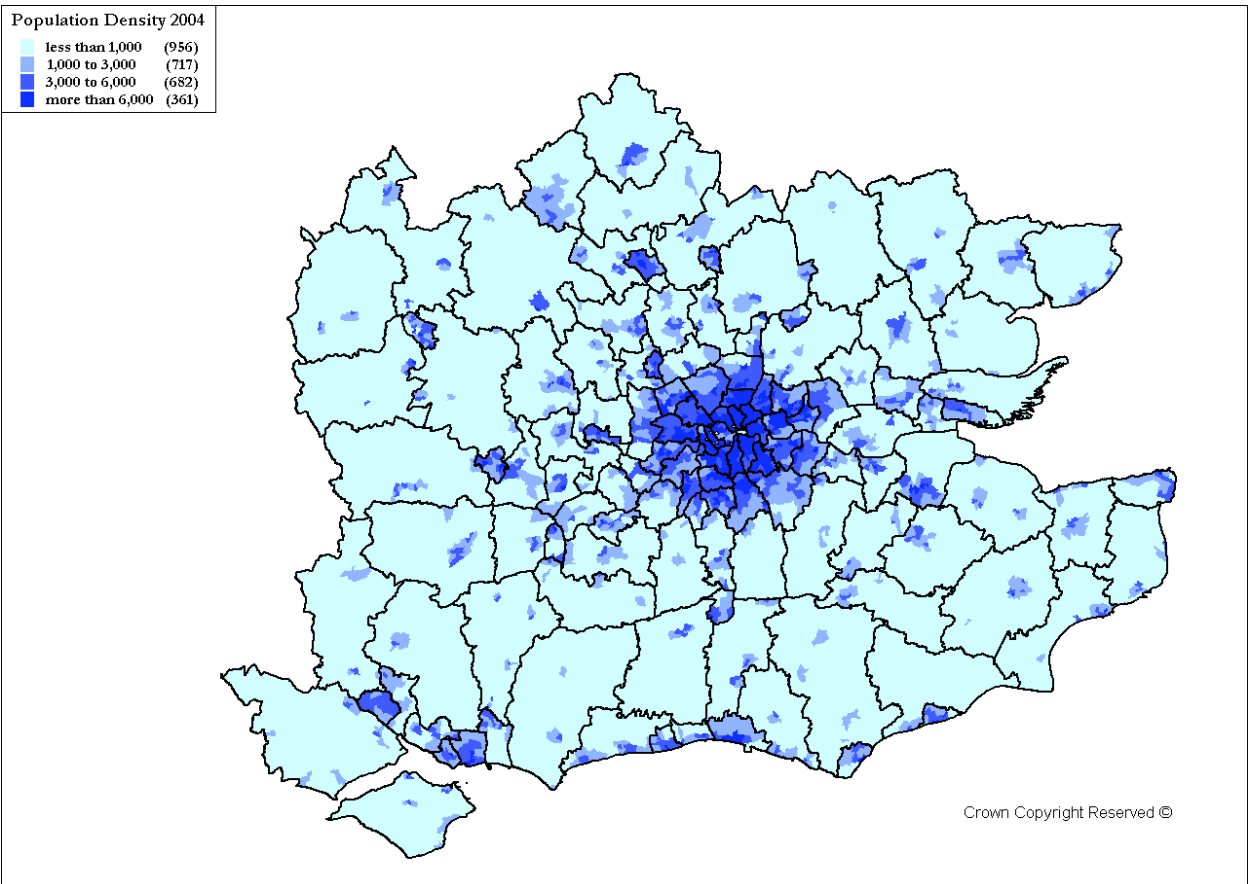
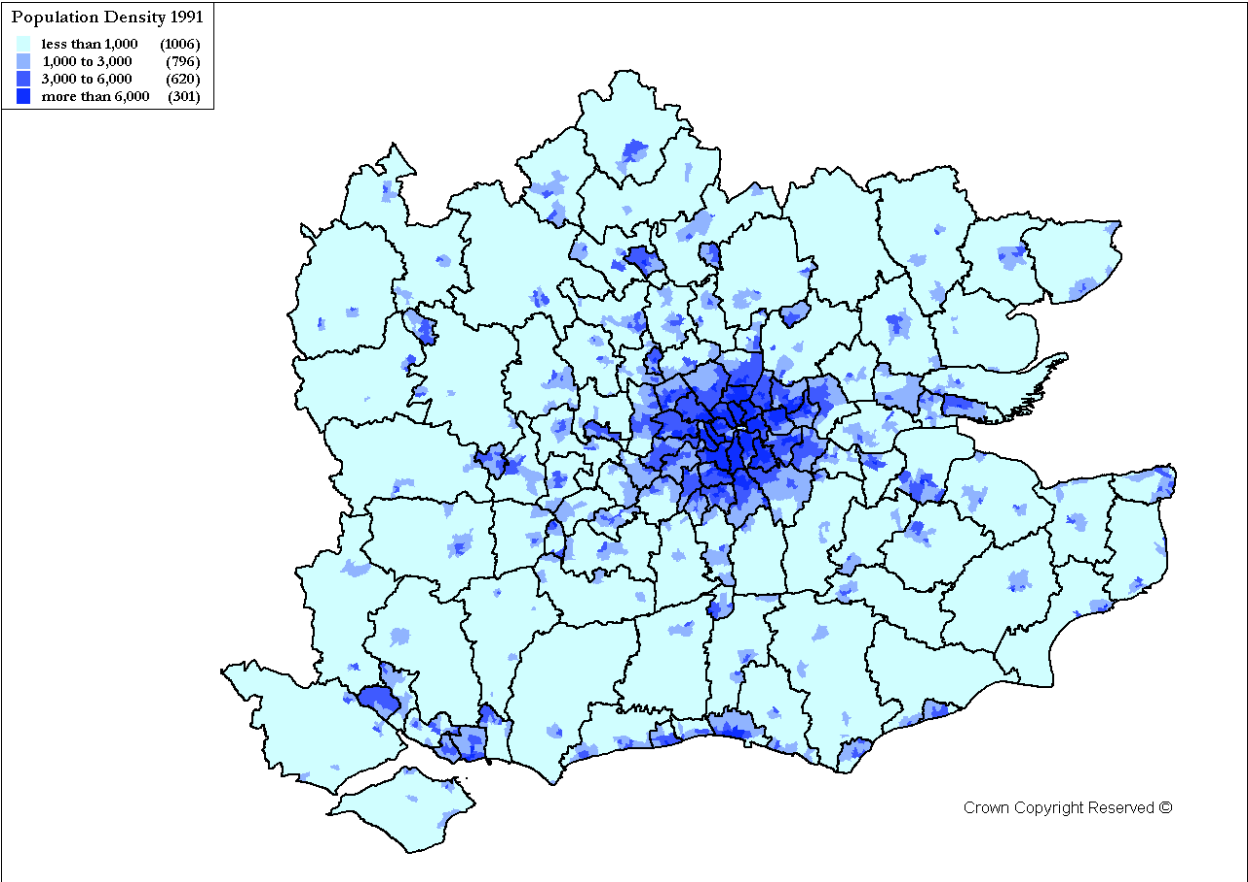
Hierarchical Clustering - Average Linkage

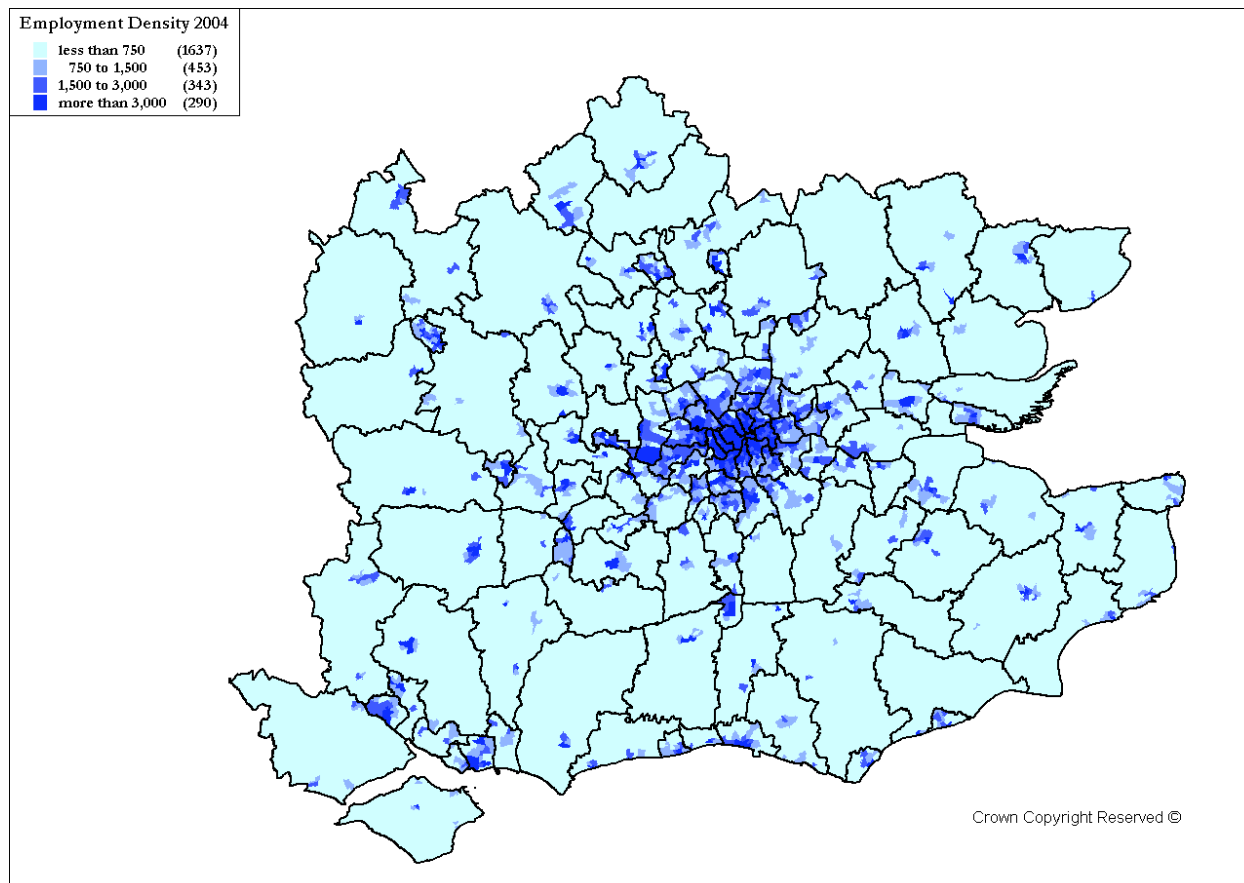
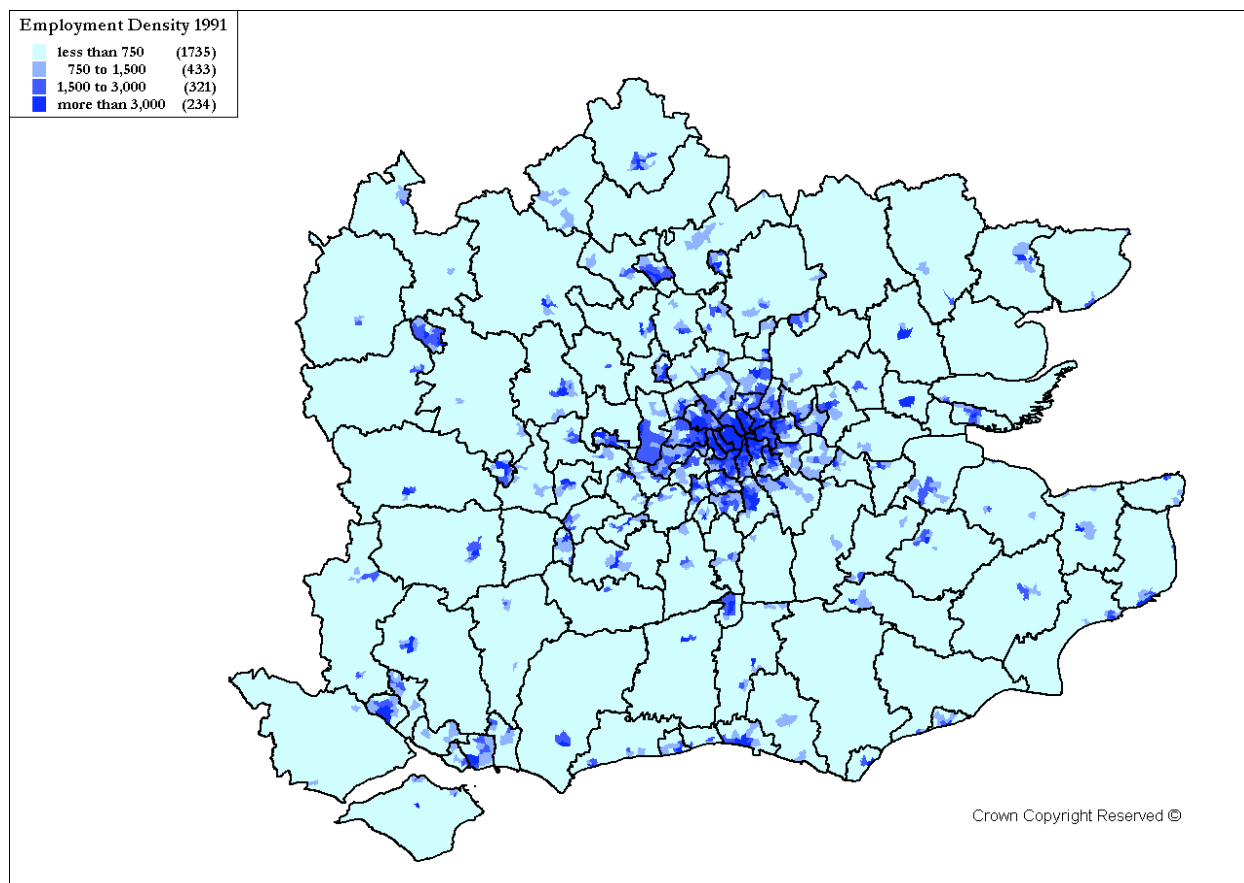


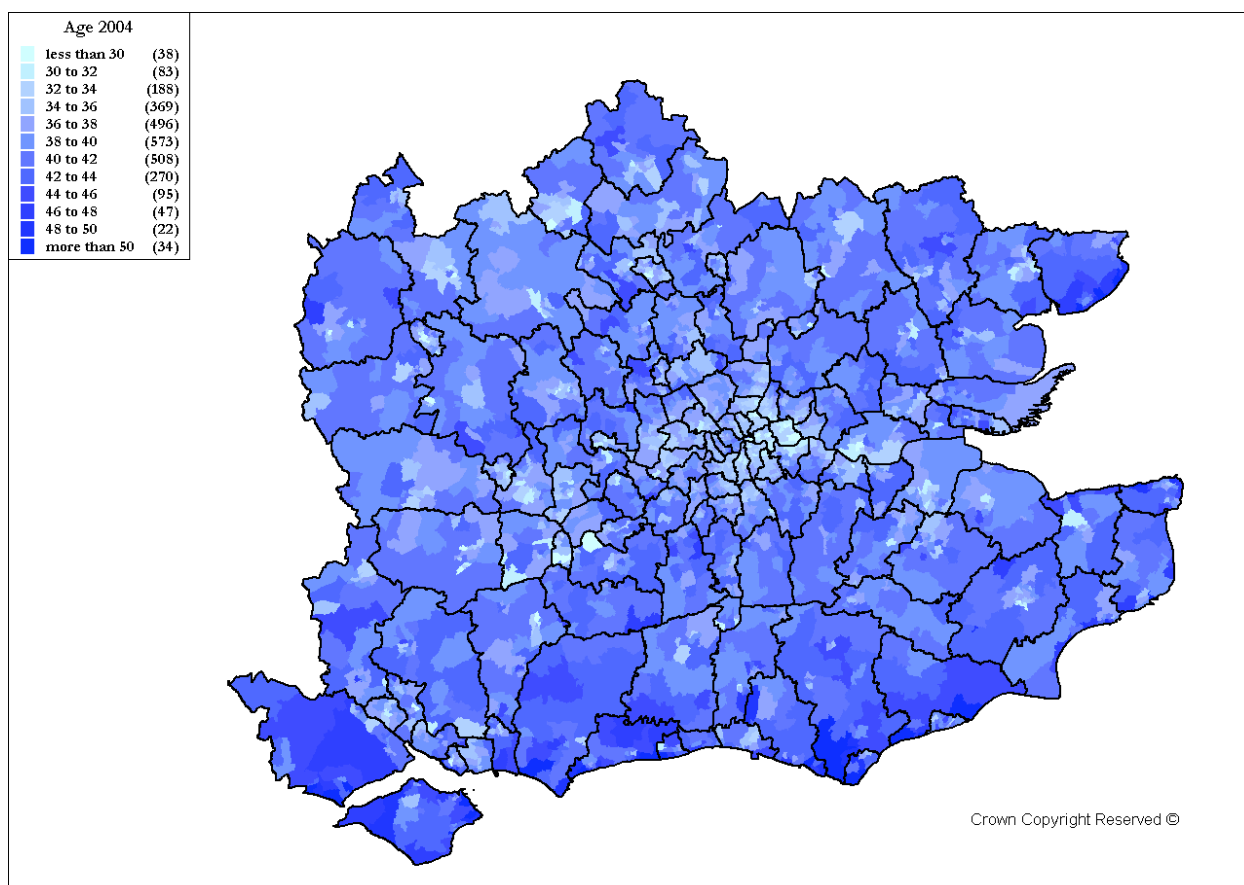
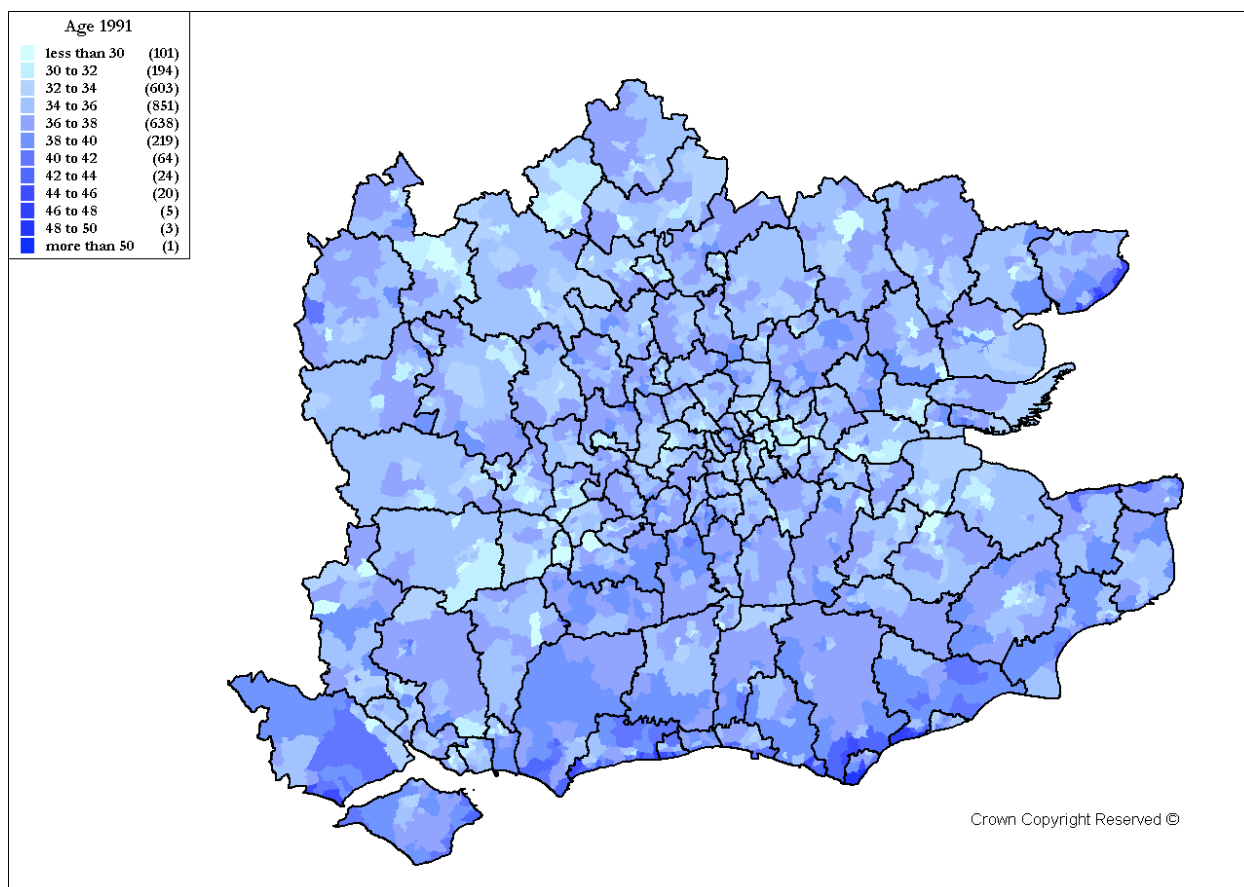
Hierarchical Clustering - Complete Linkage

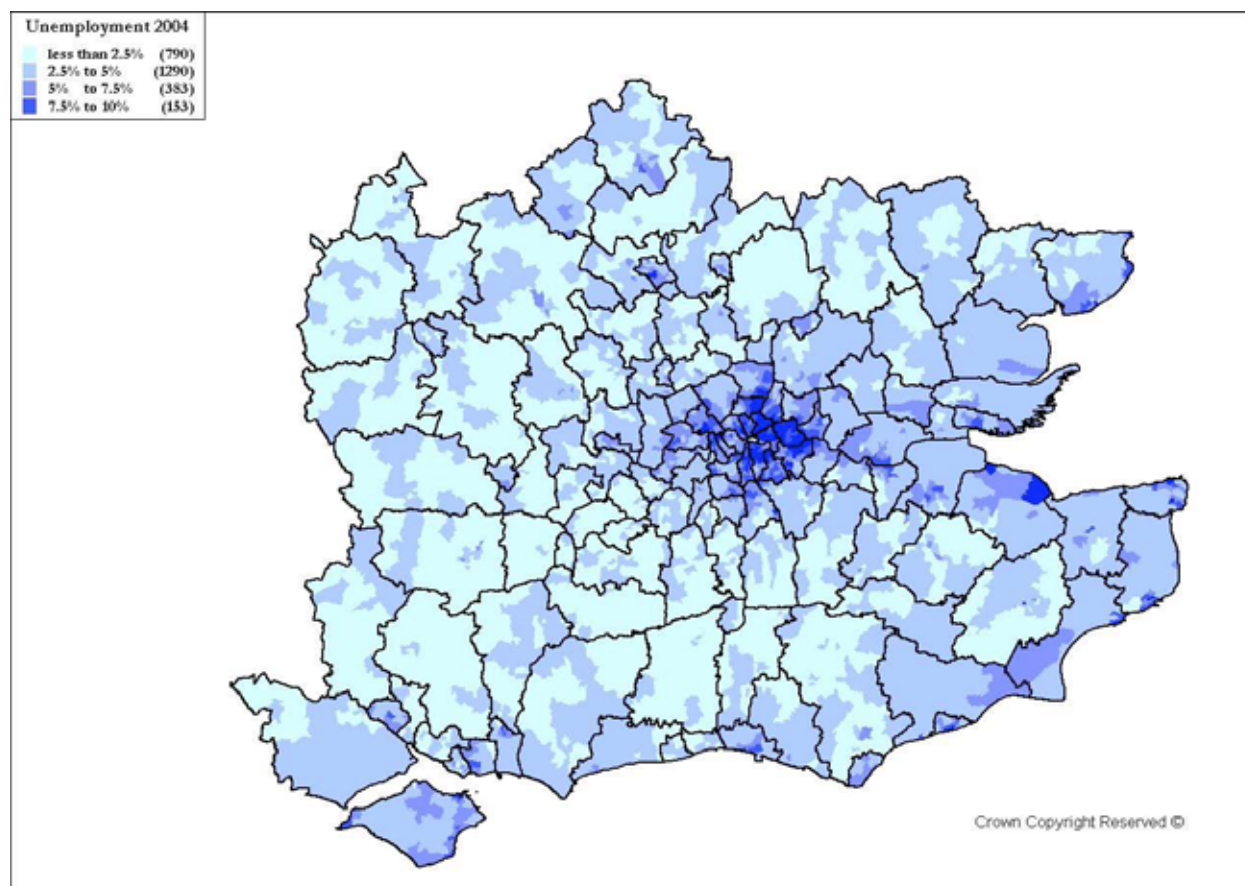
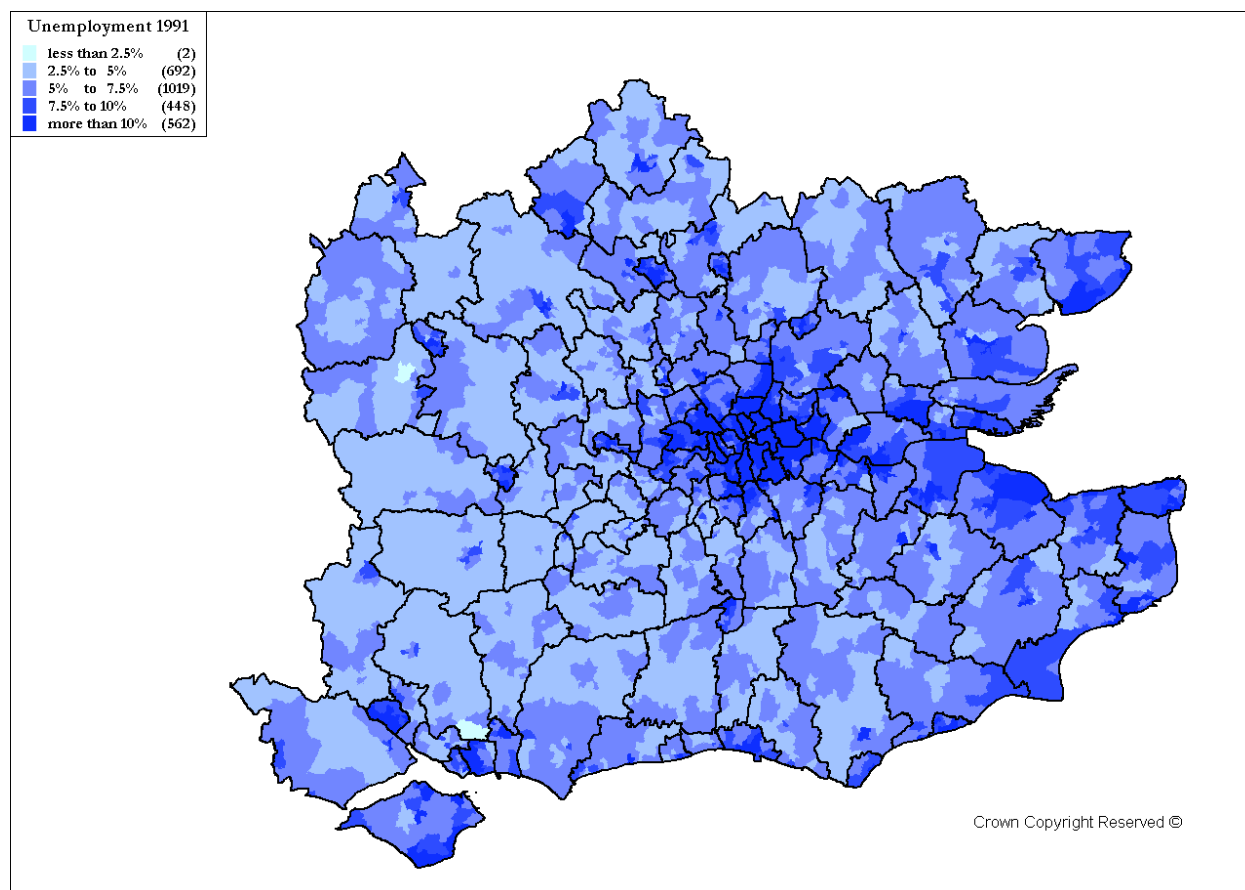


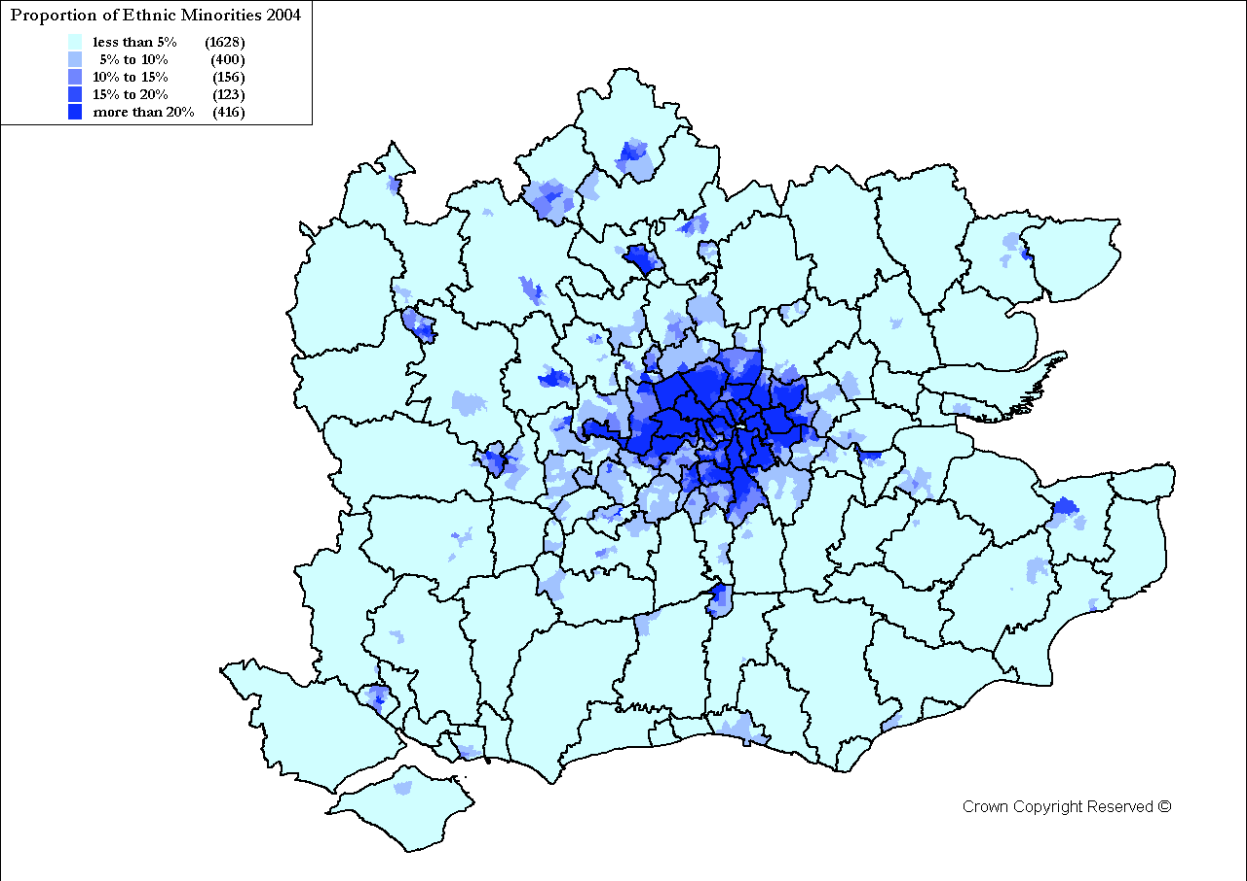
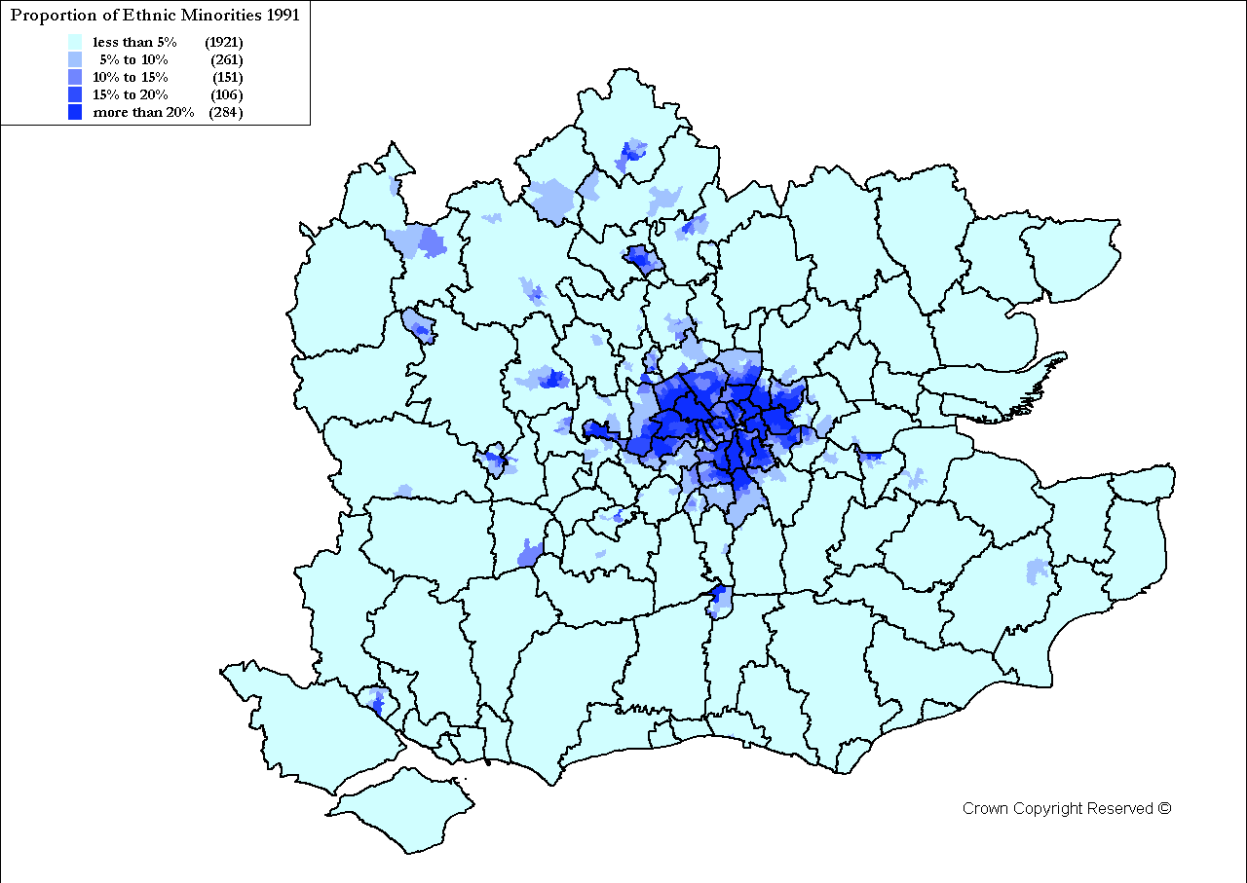
Appendix III: Individual Variable Maps

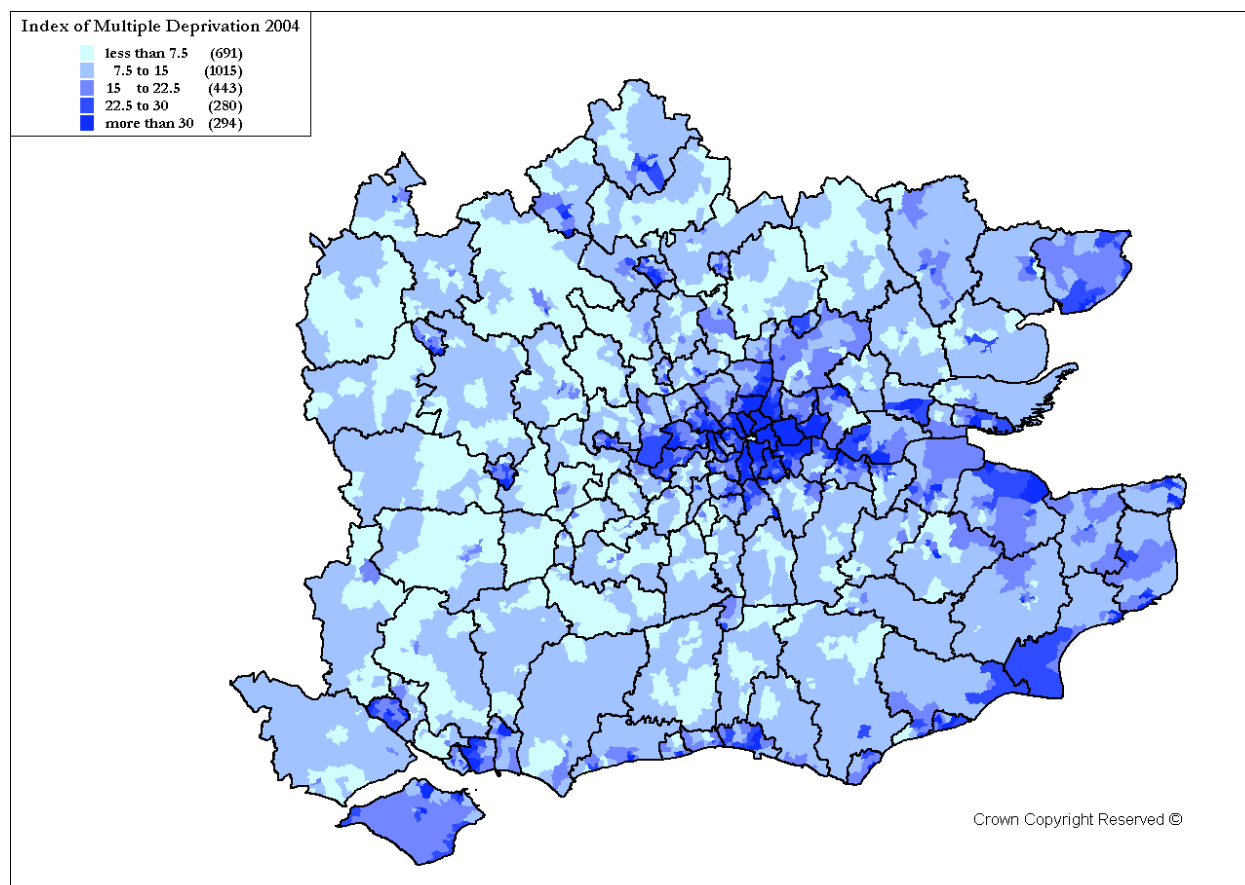


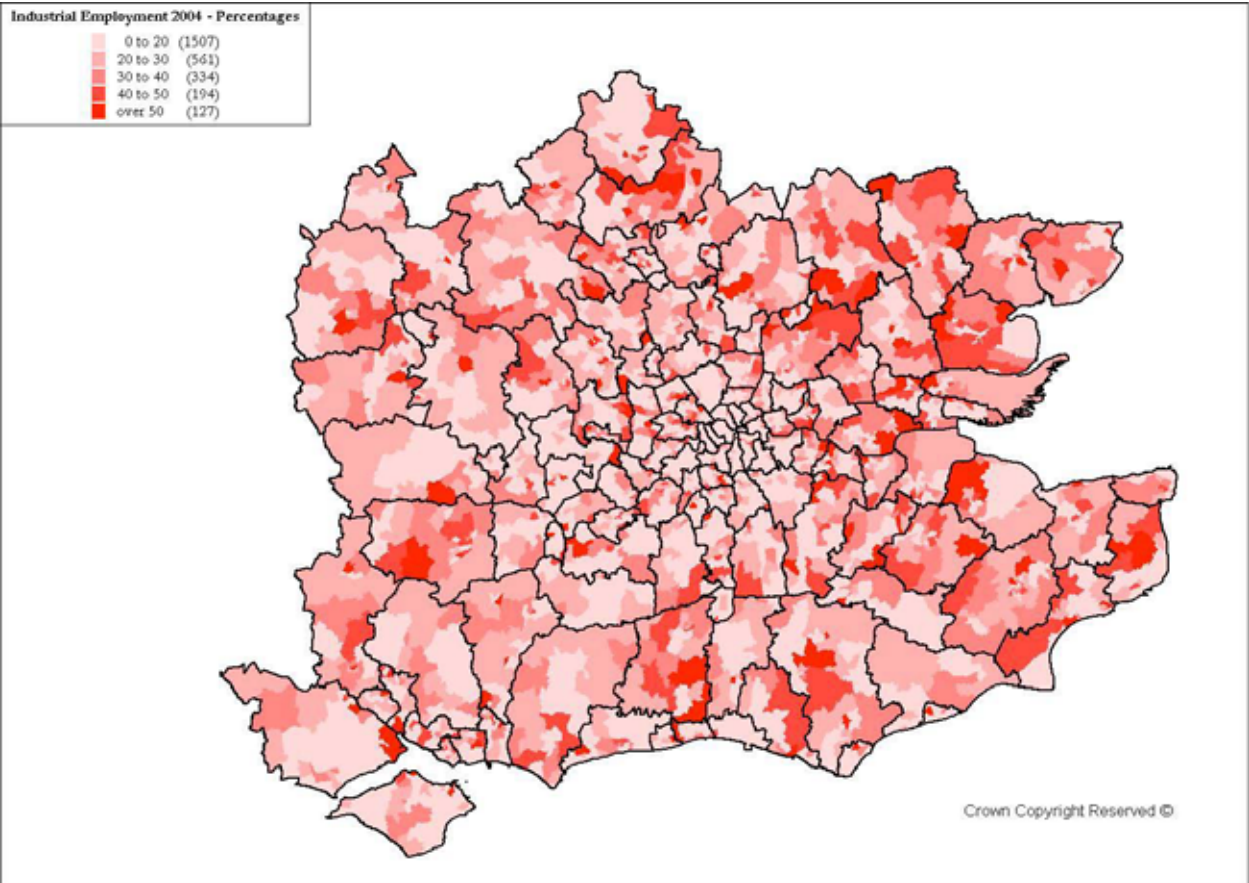
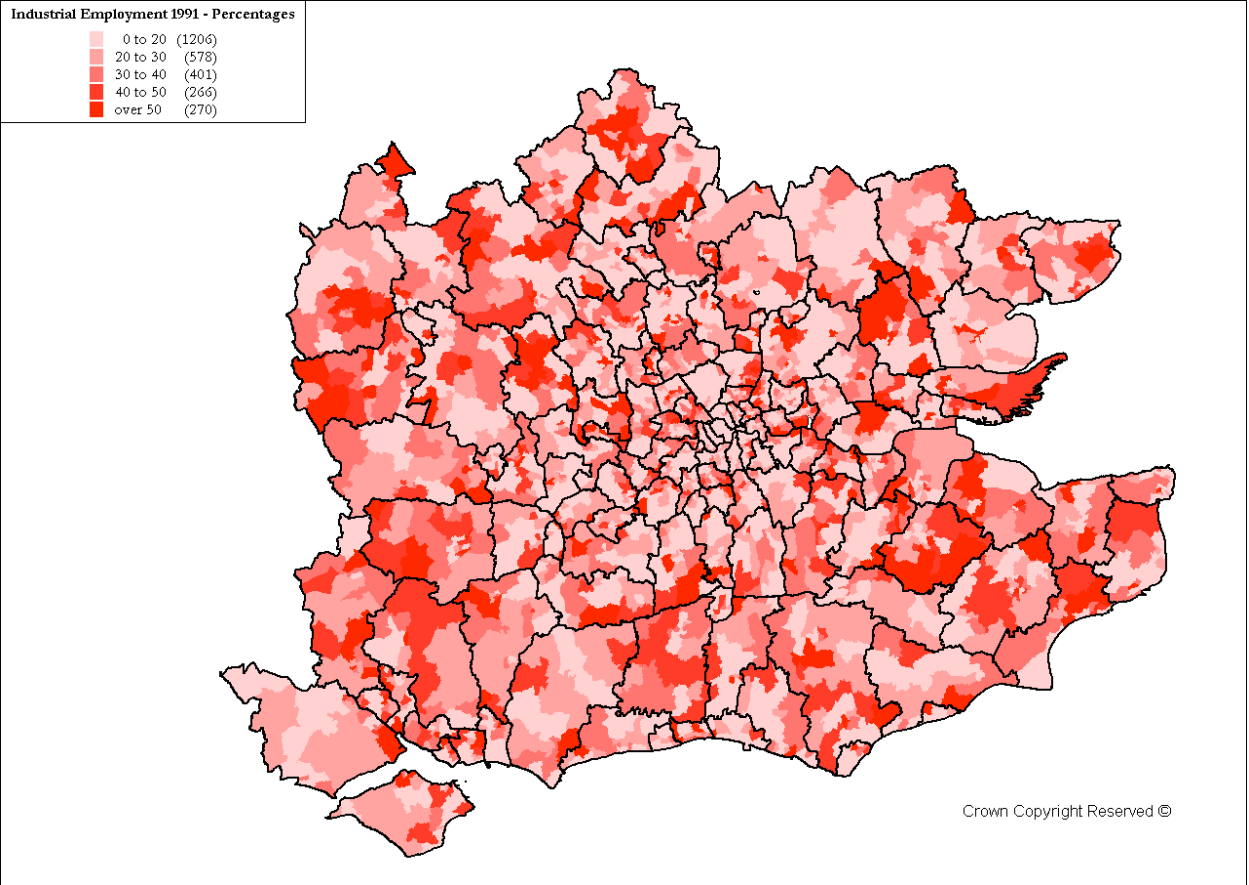


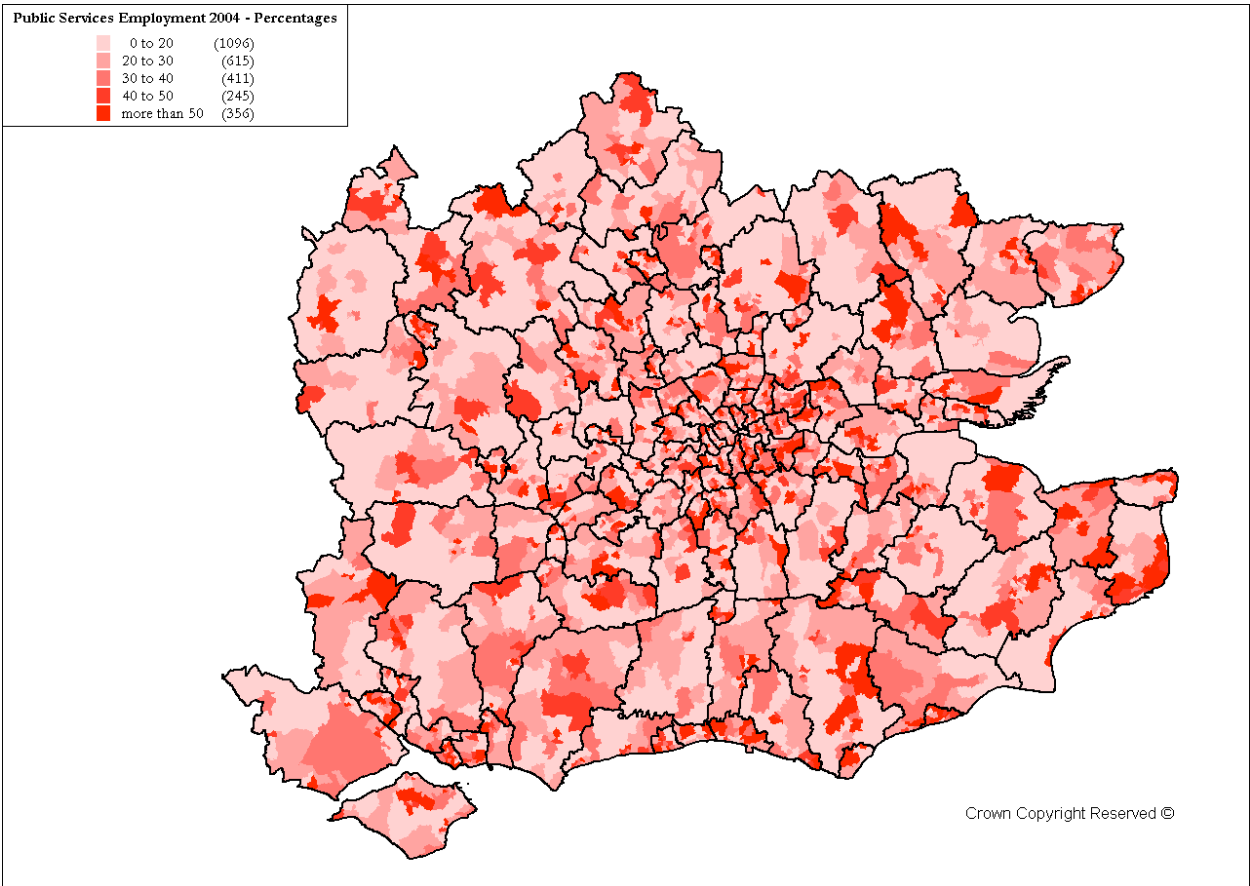
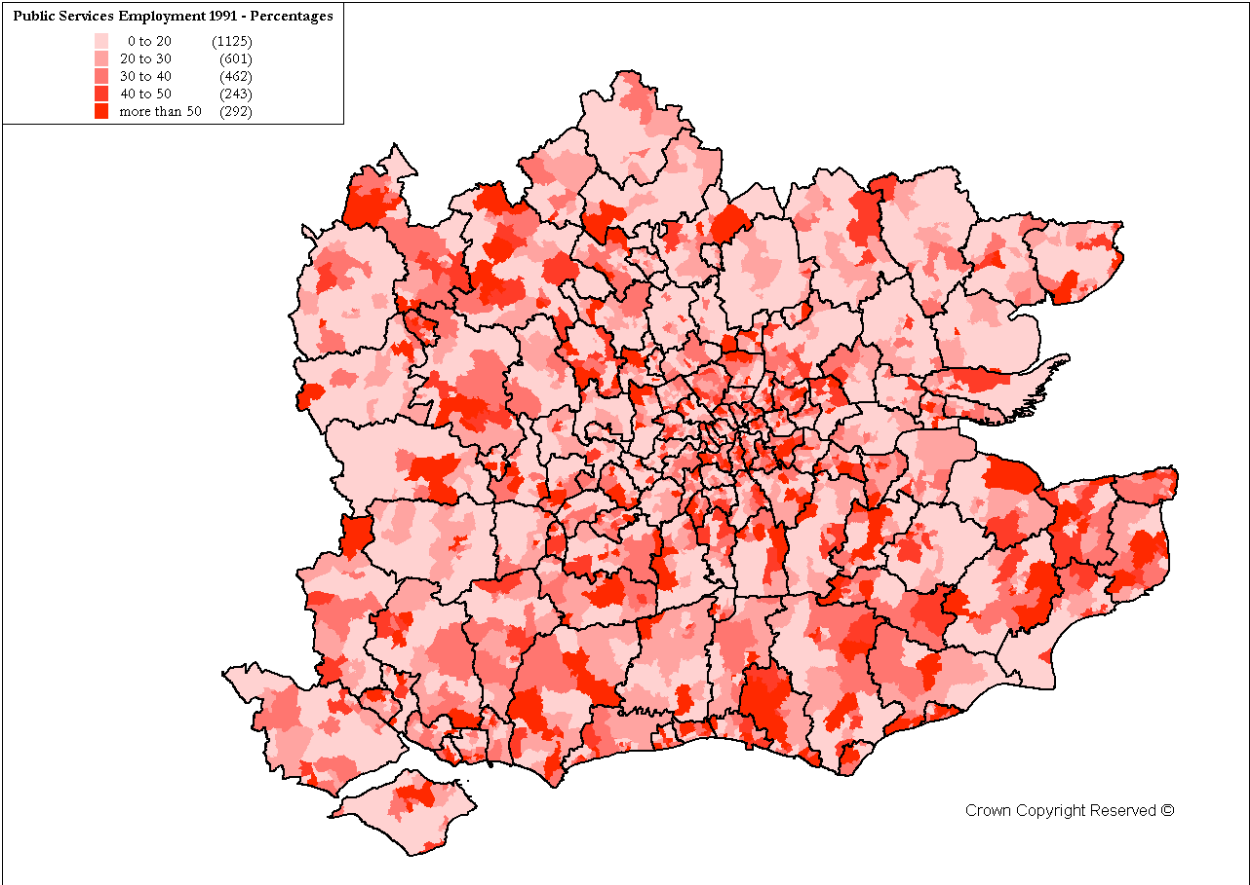


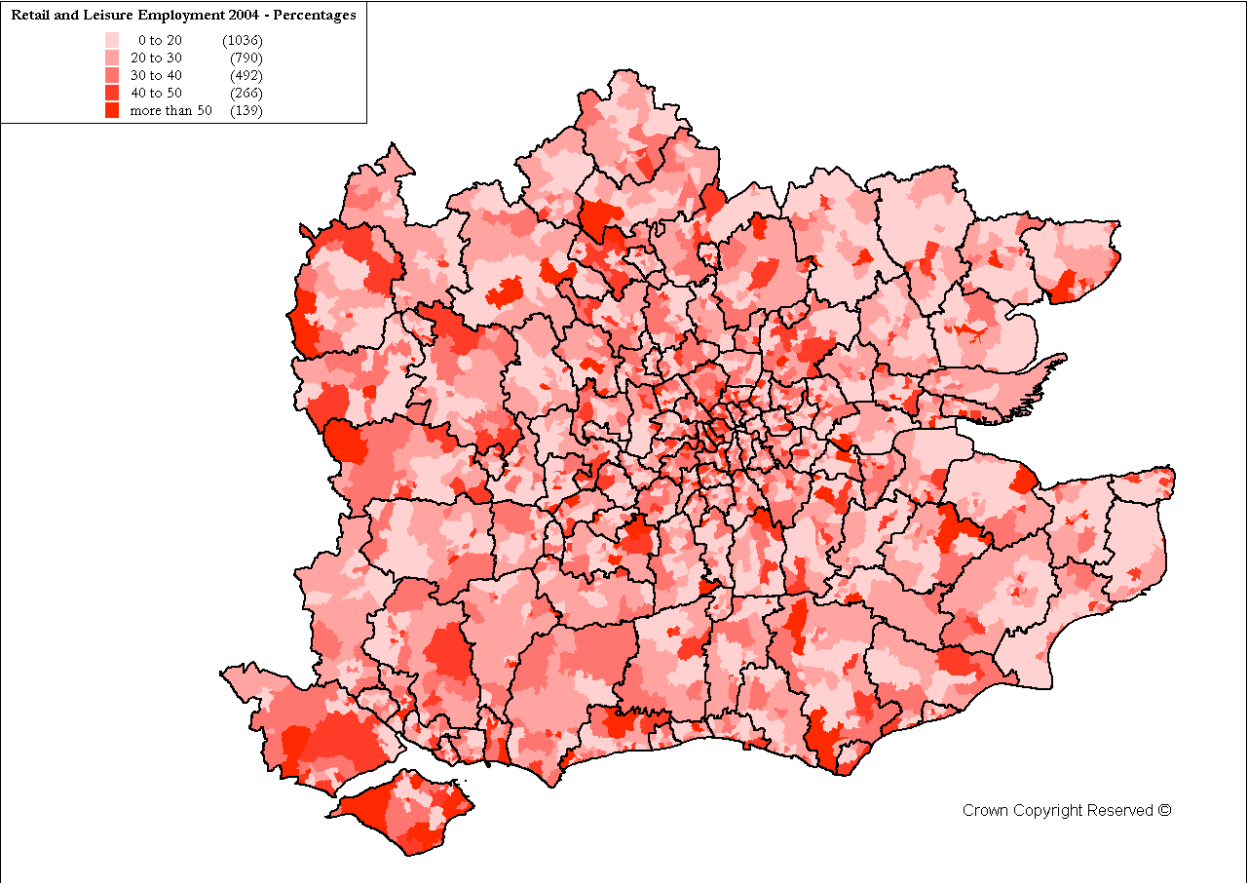
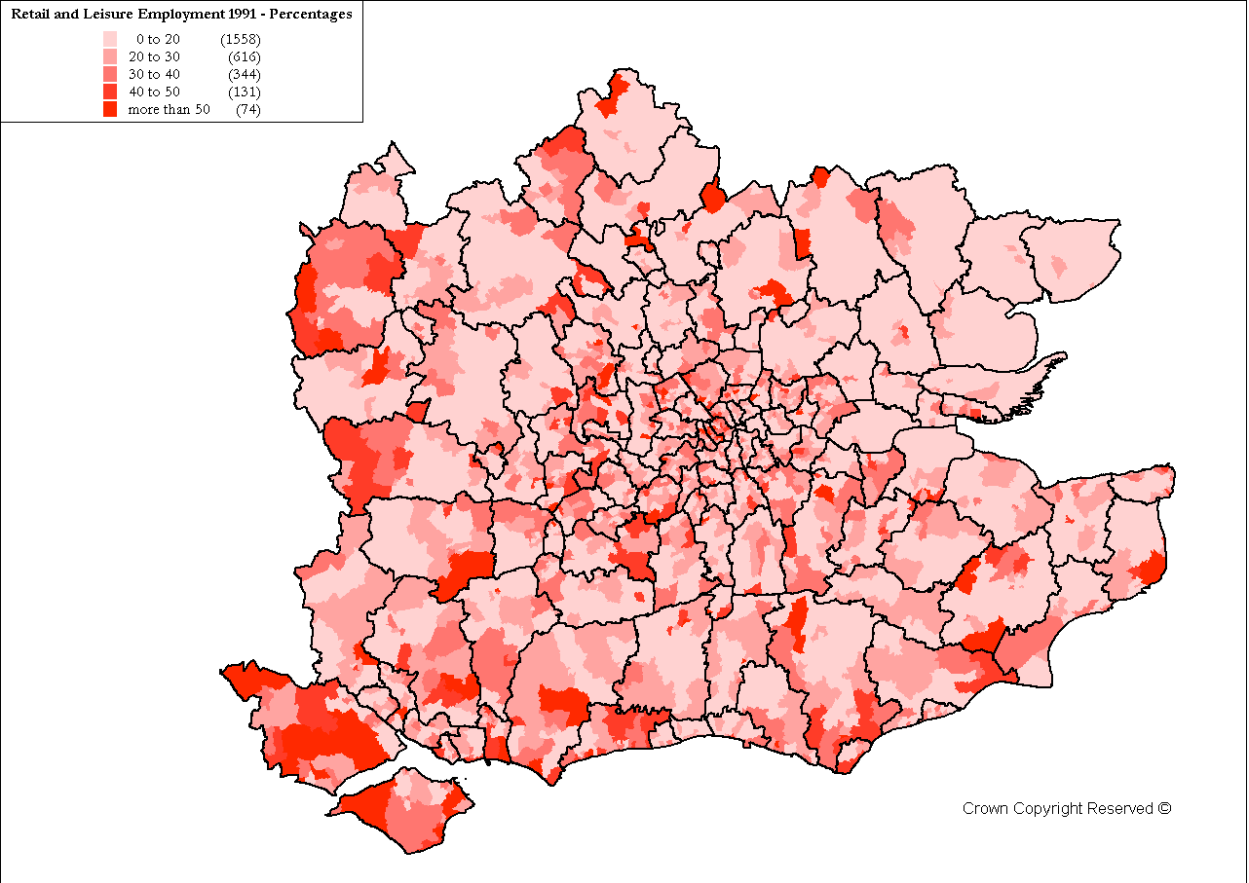


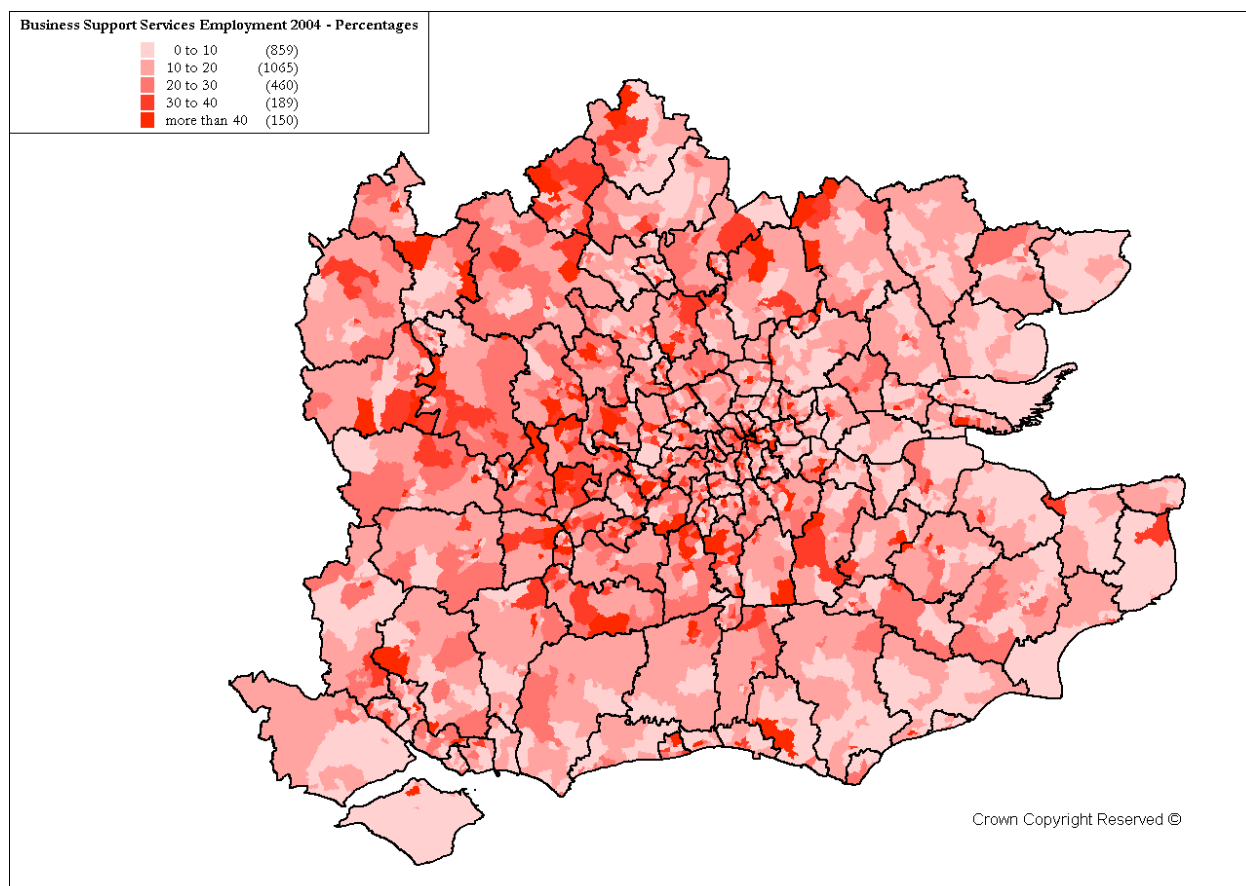
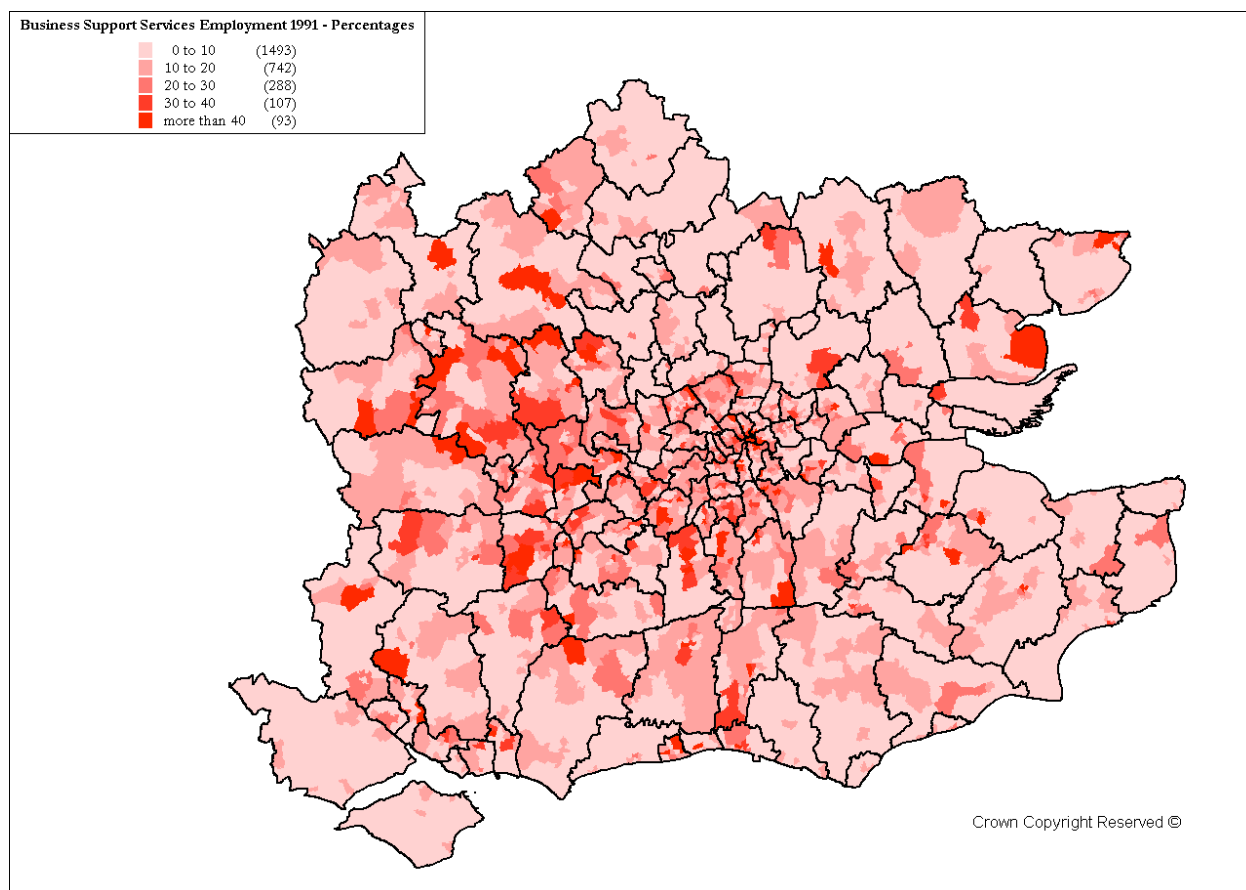


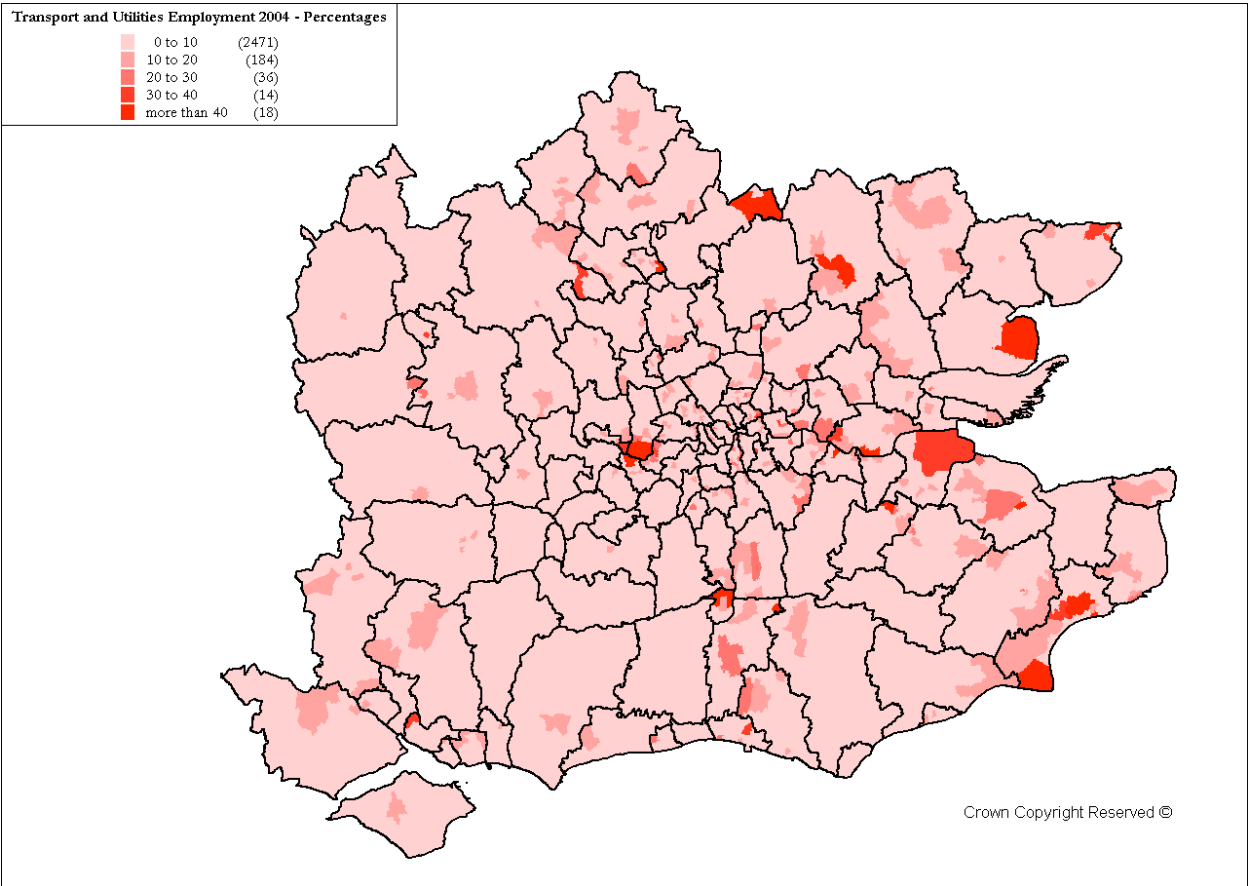
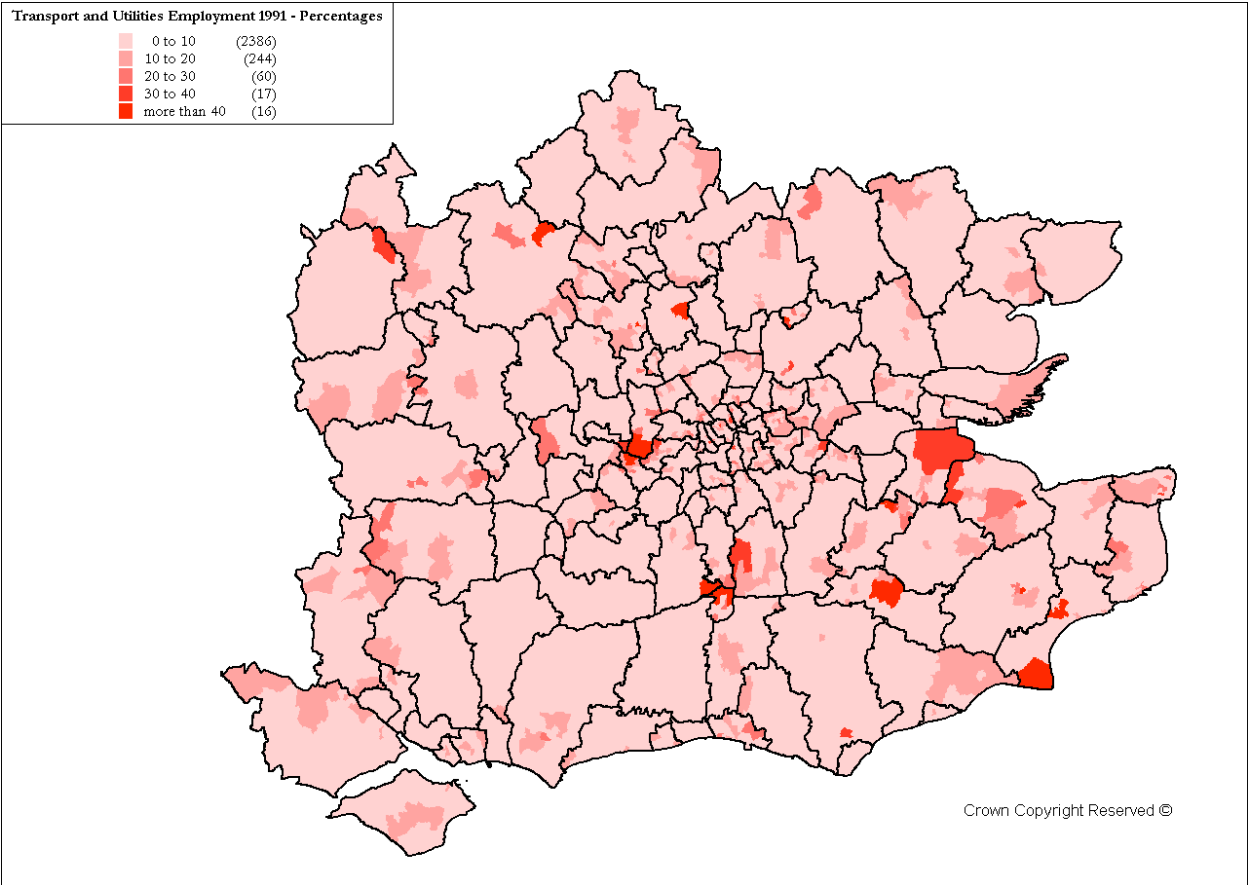




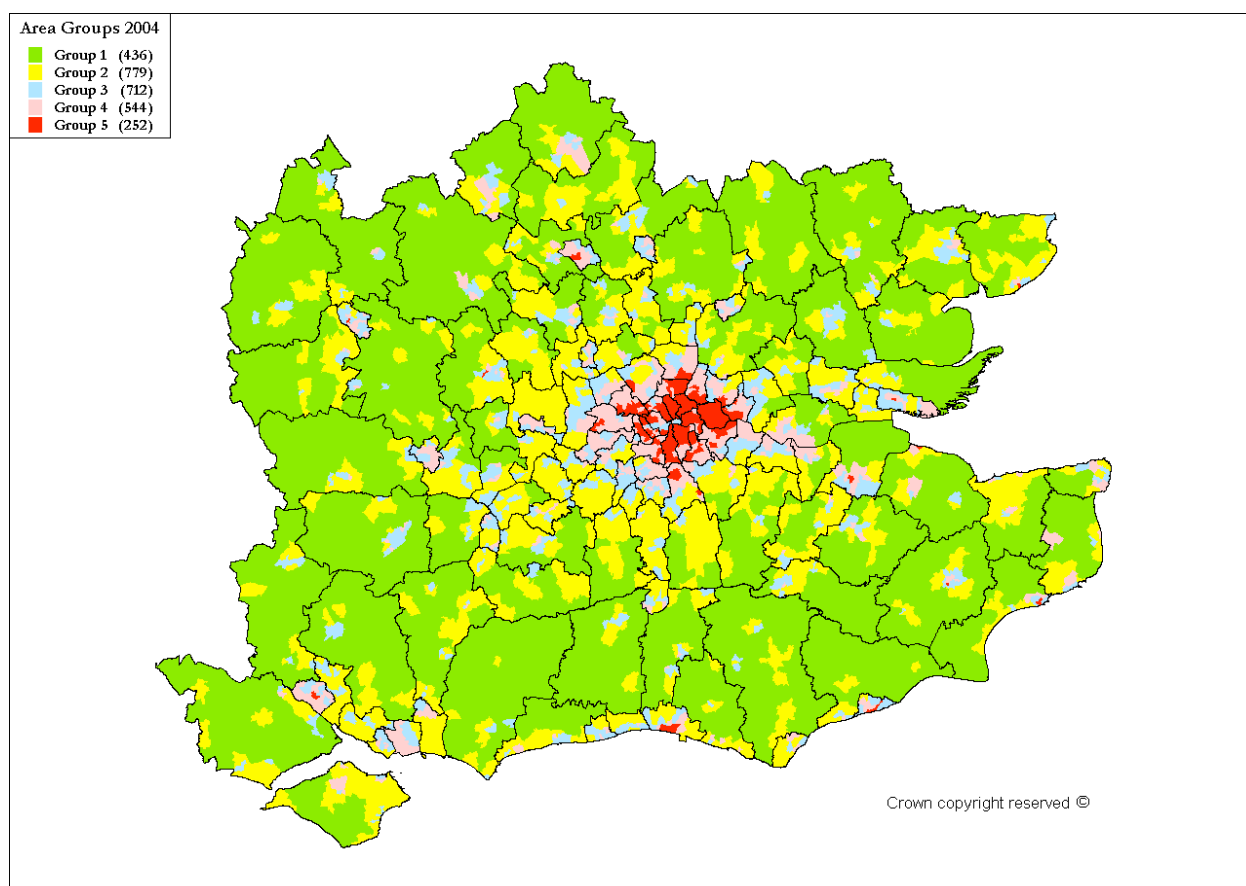
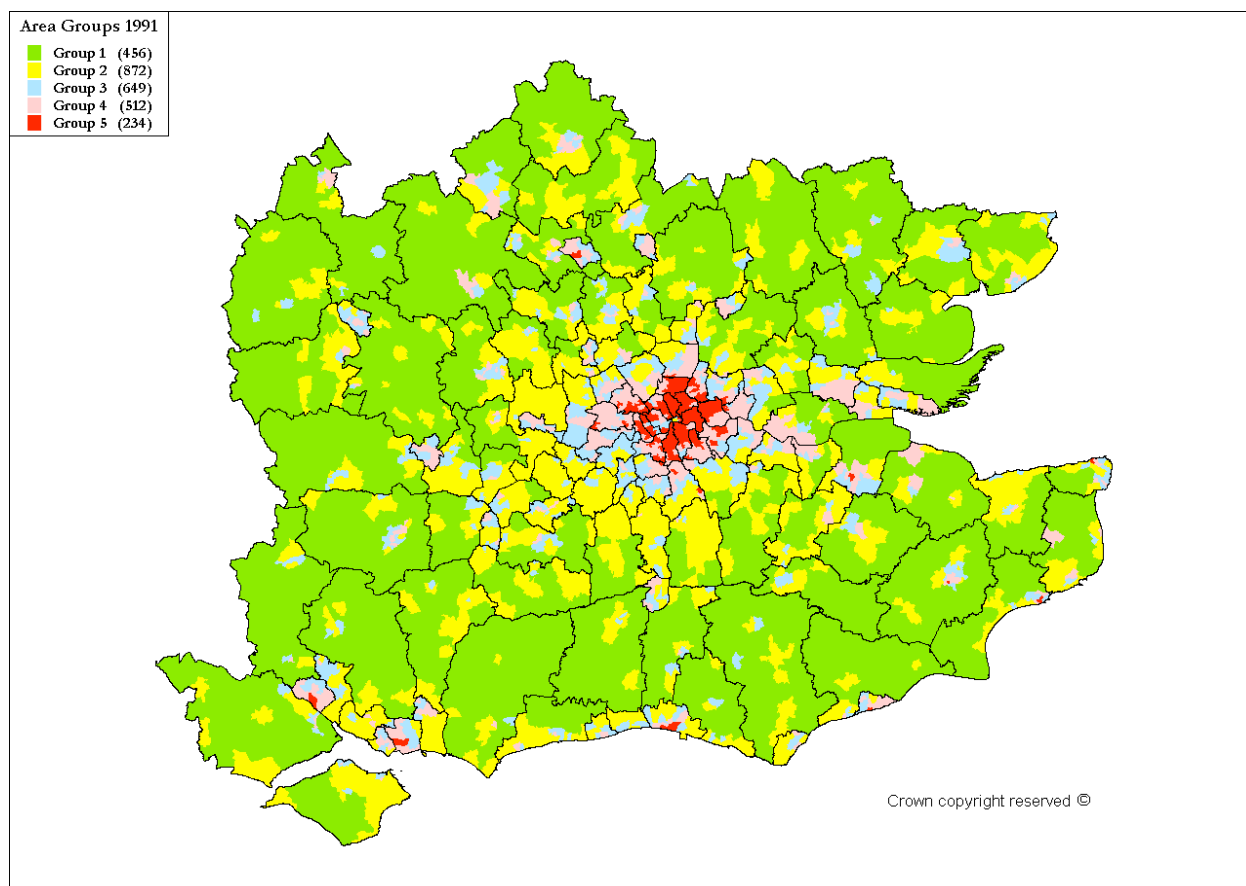


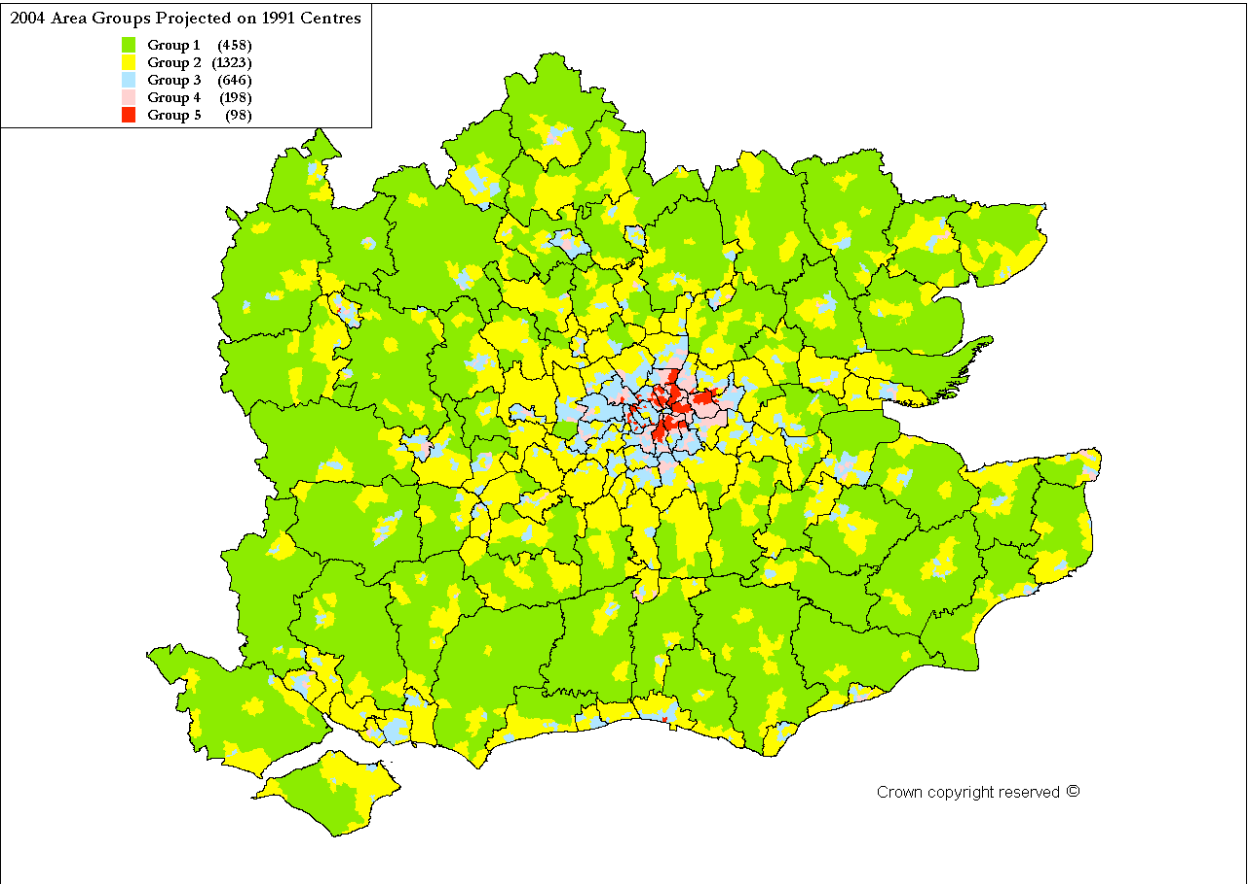
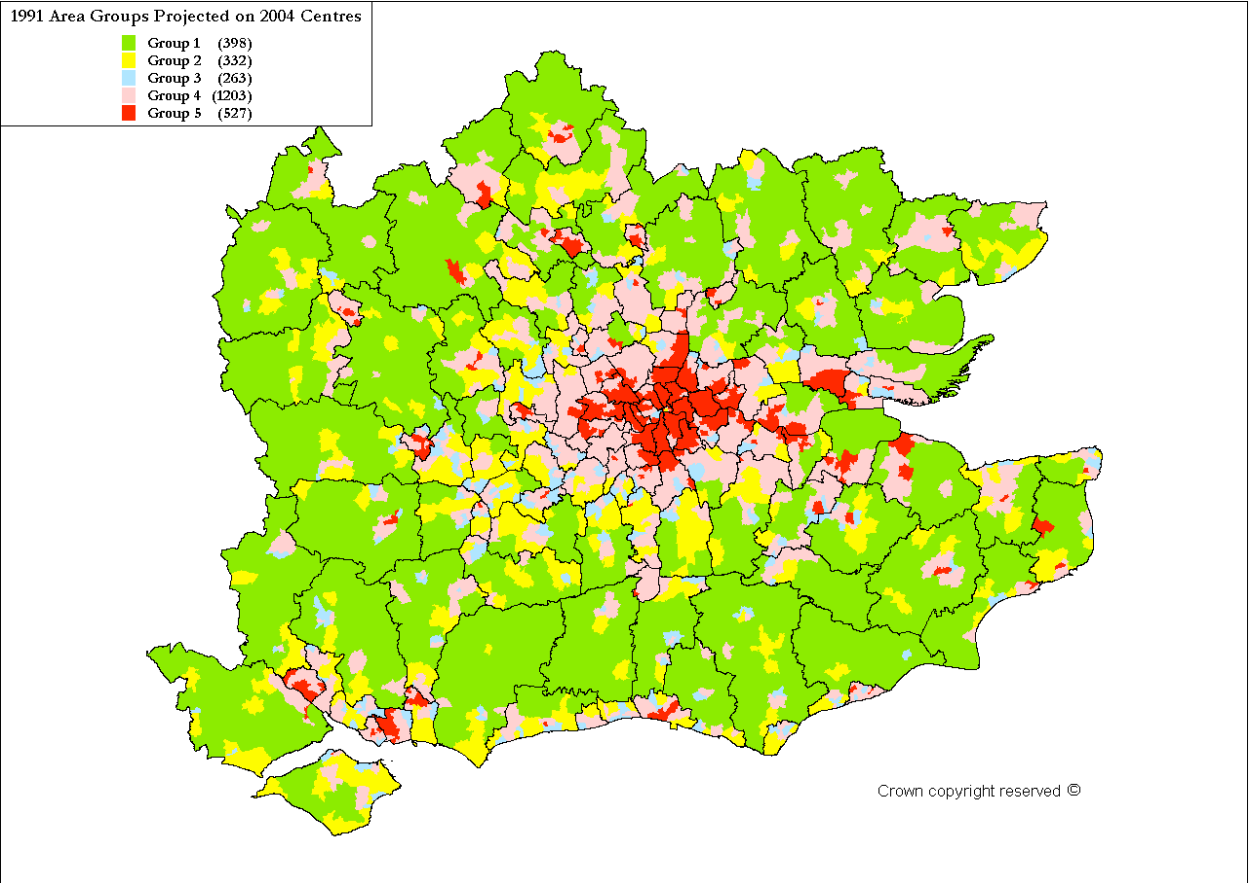


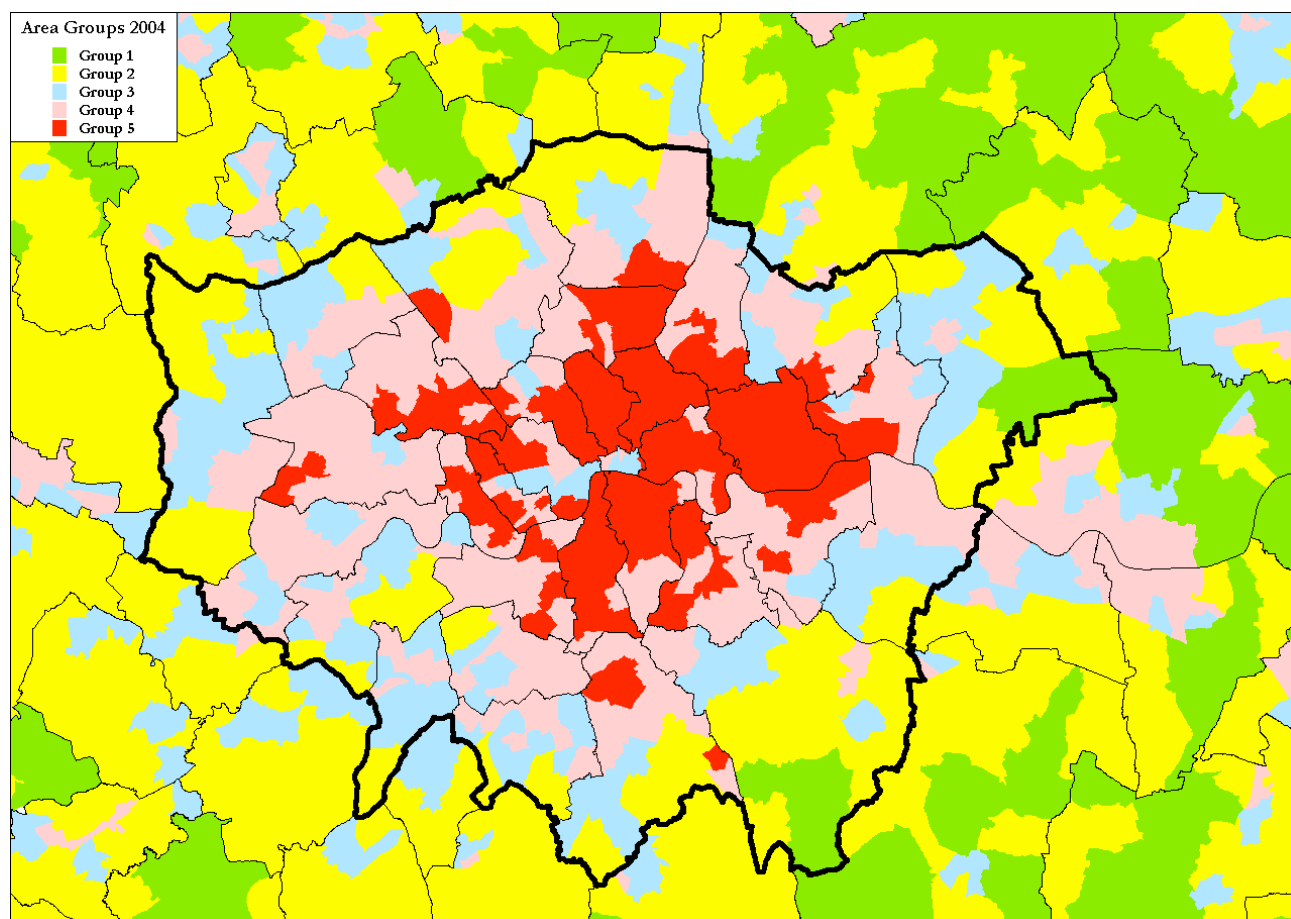
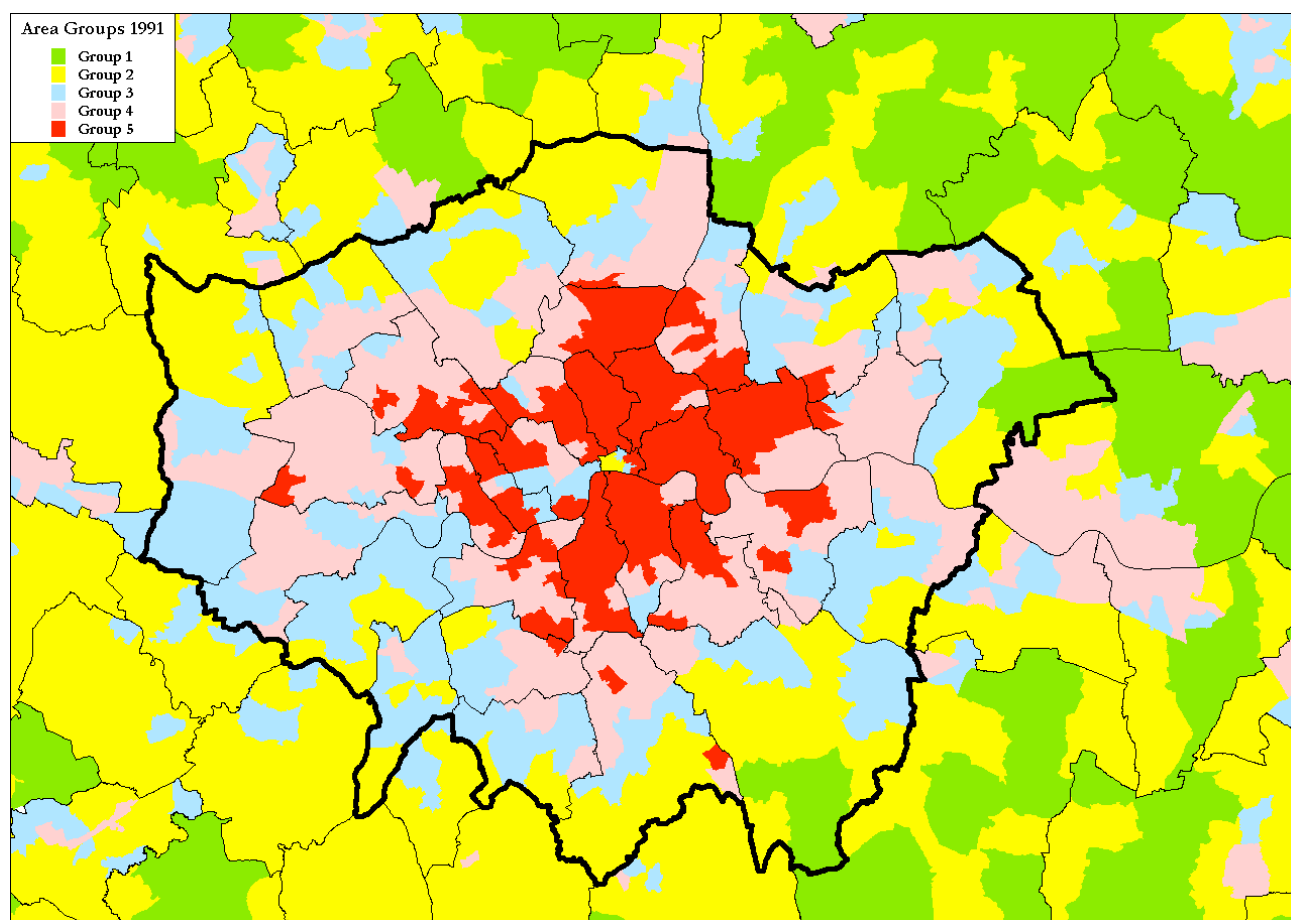


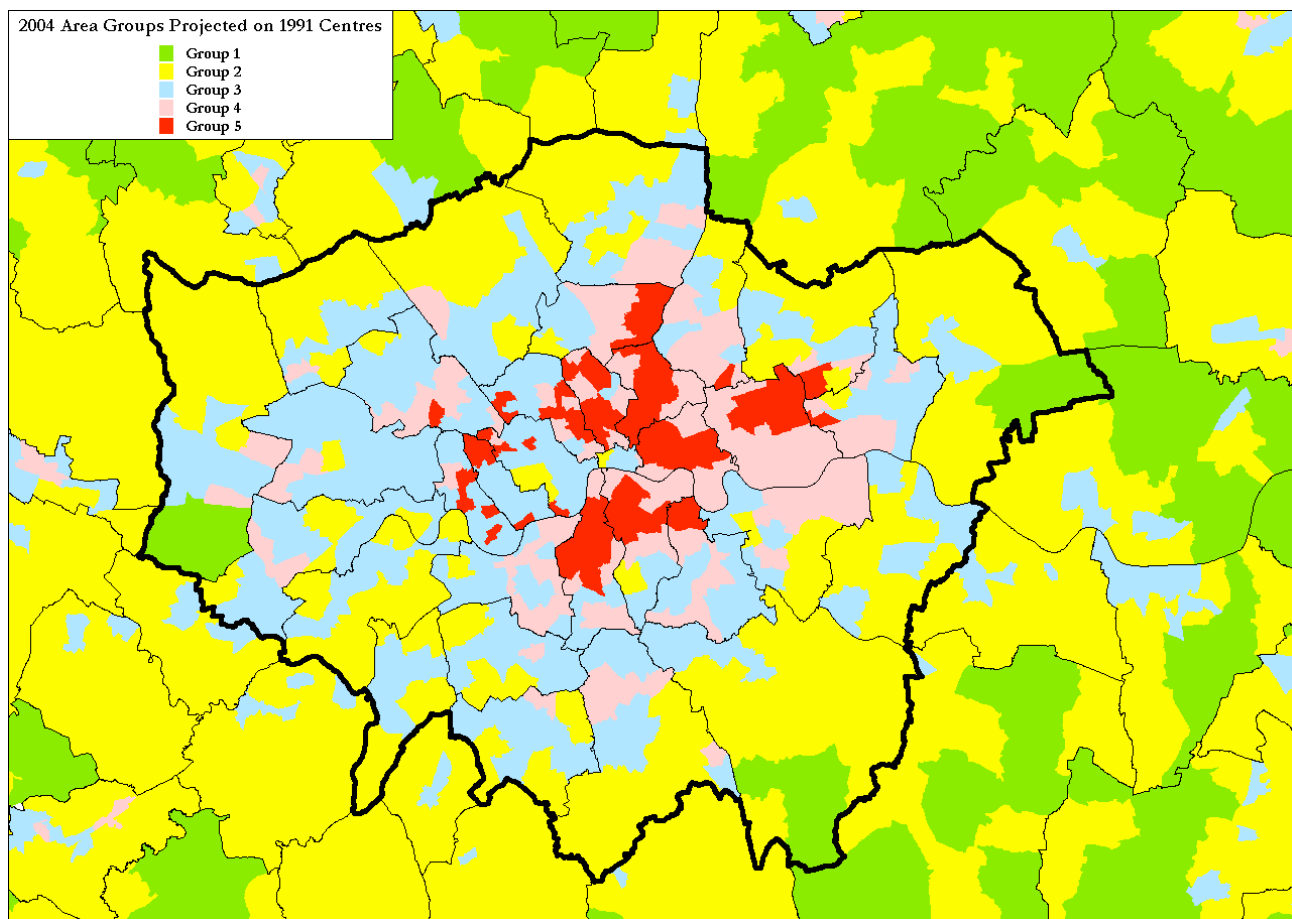
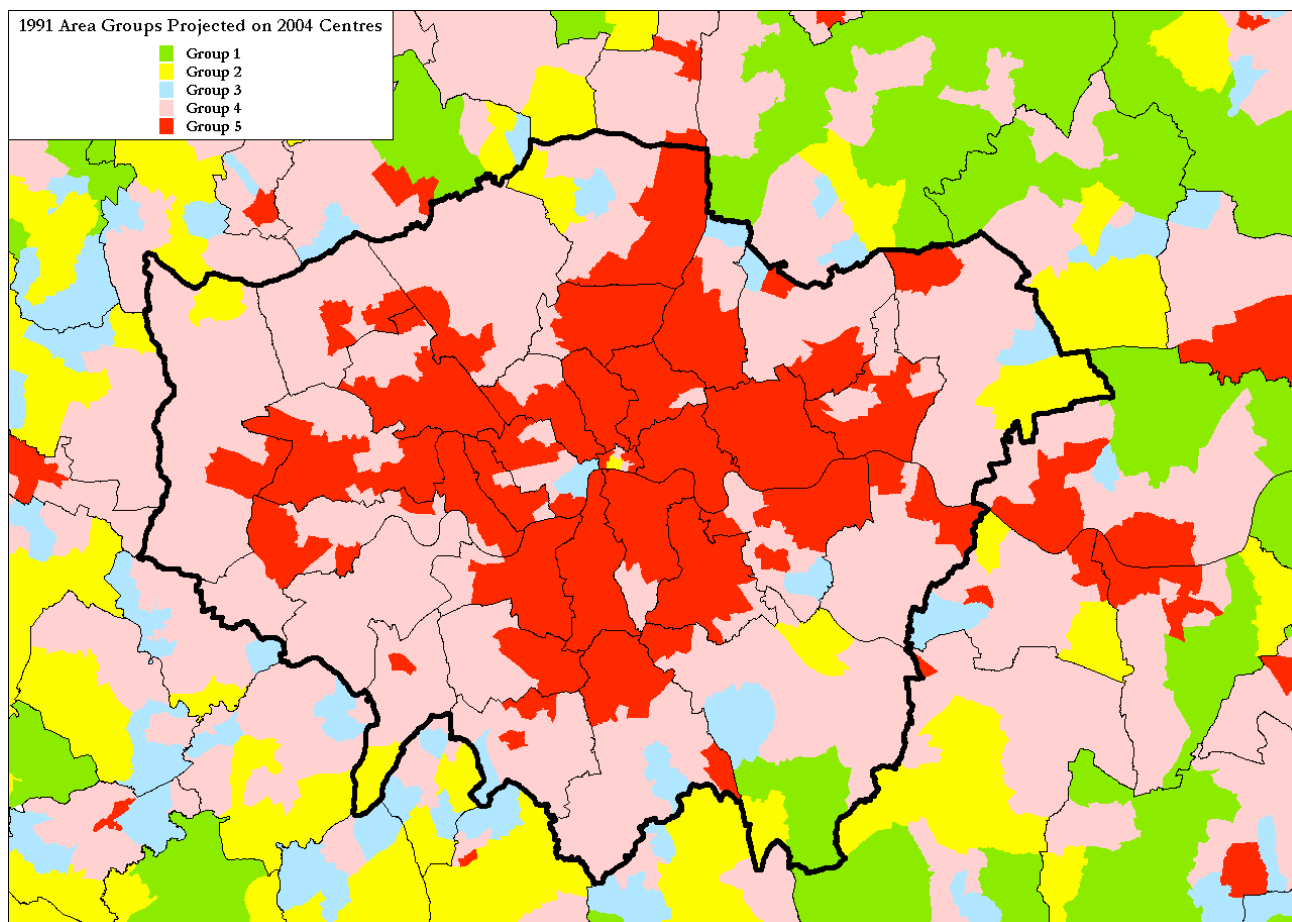


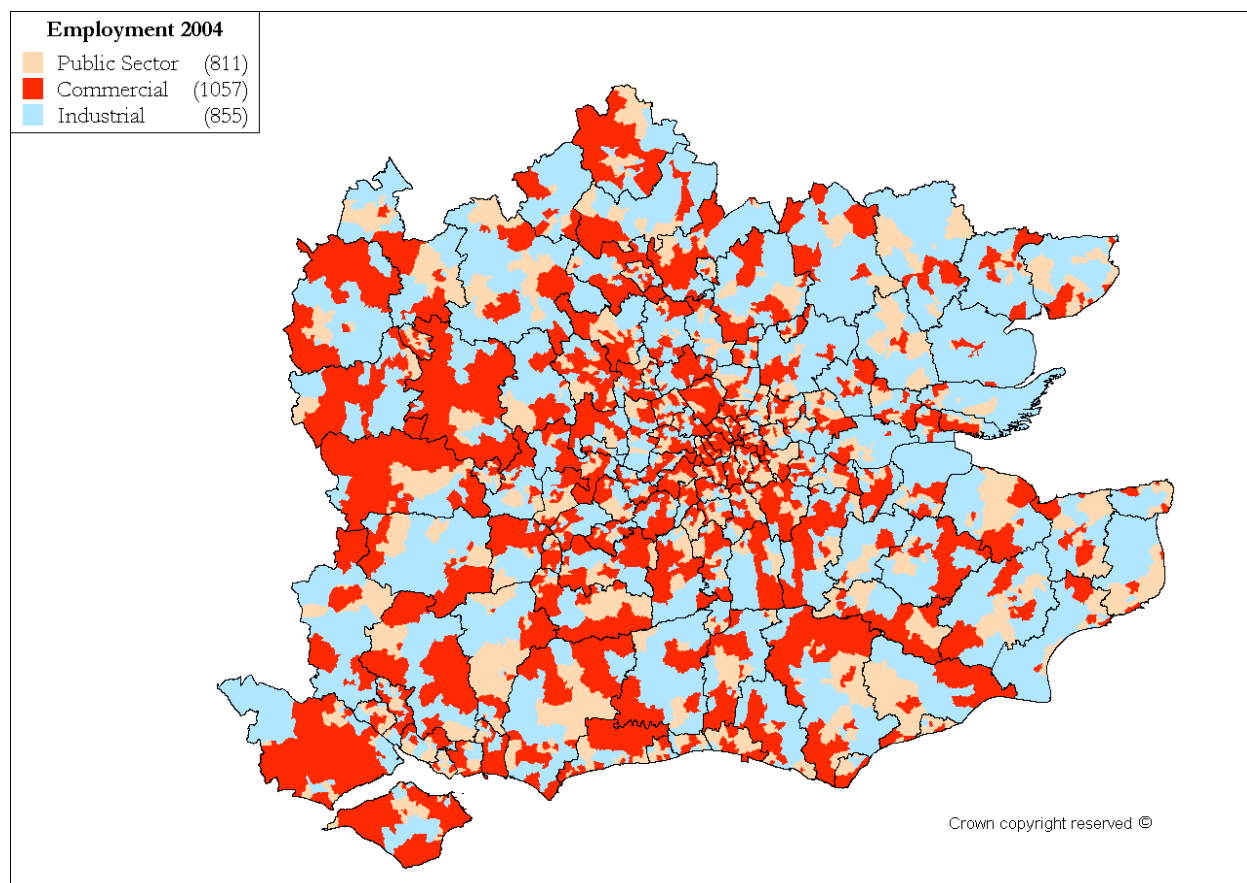
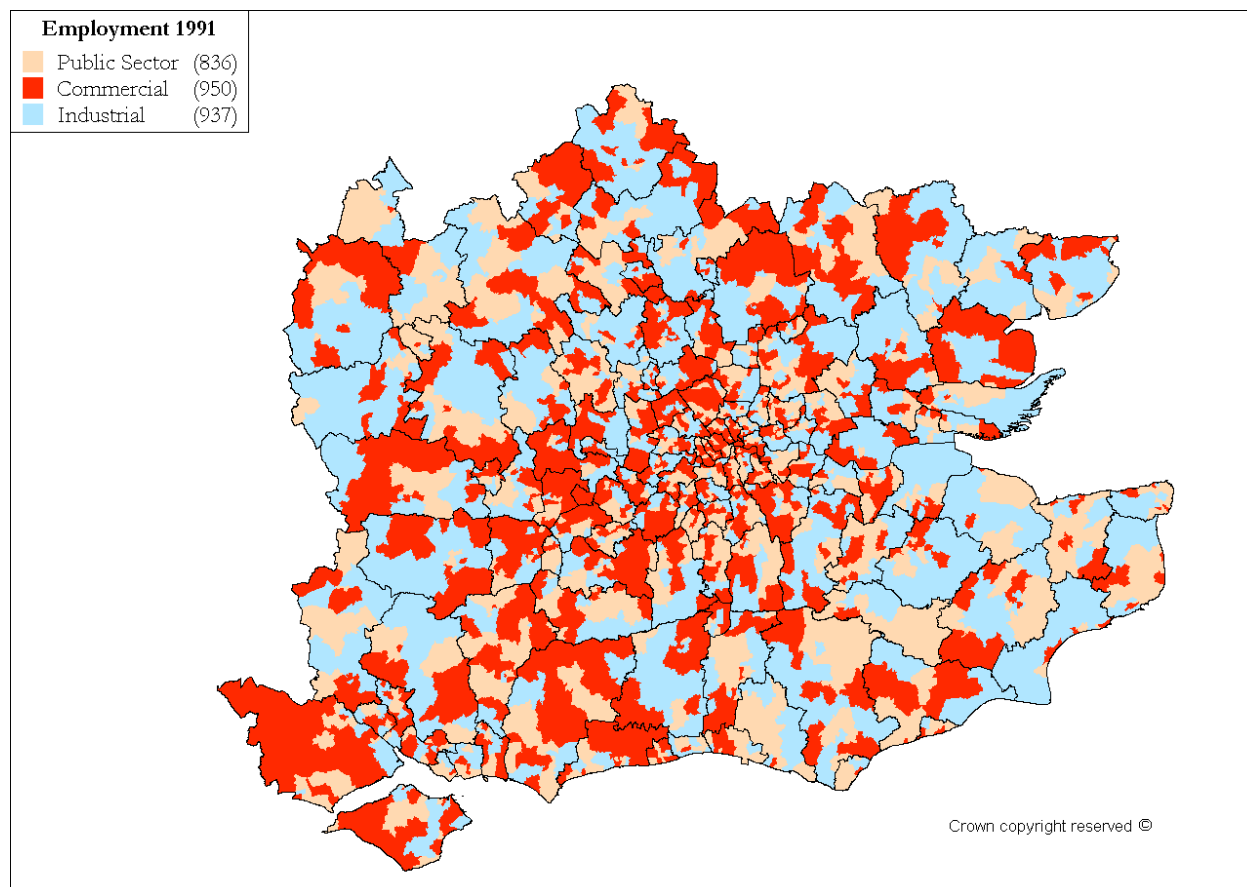
Appendix IV: Group Maps

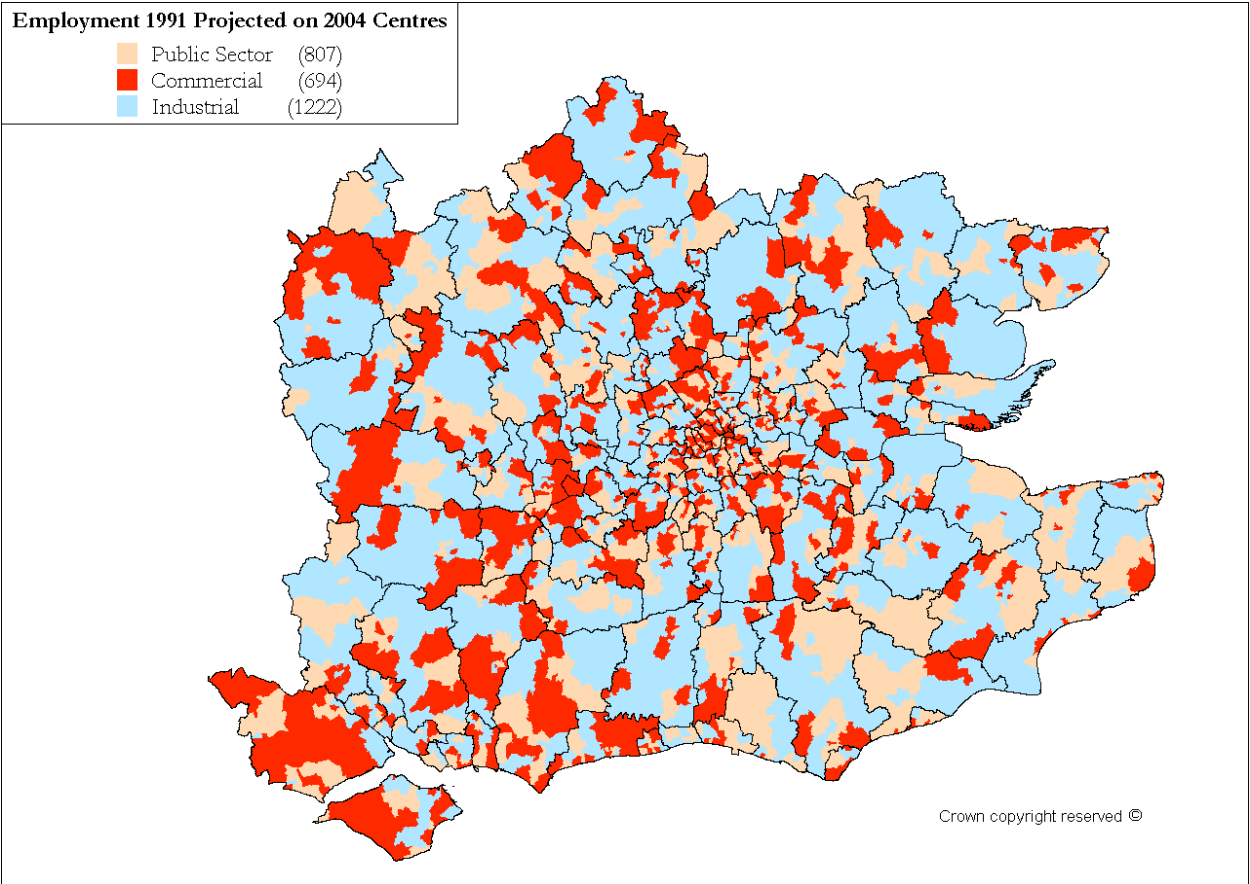
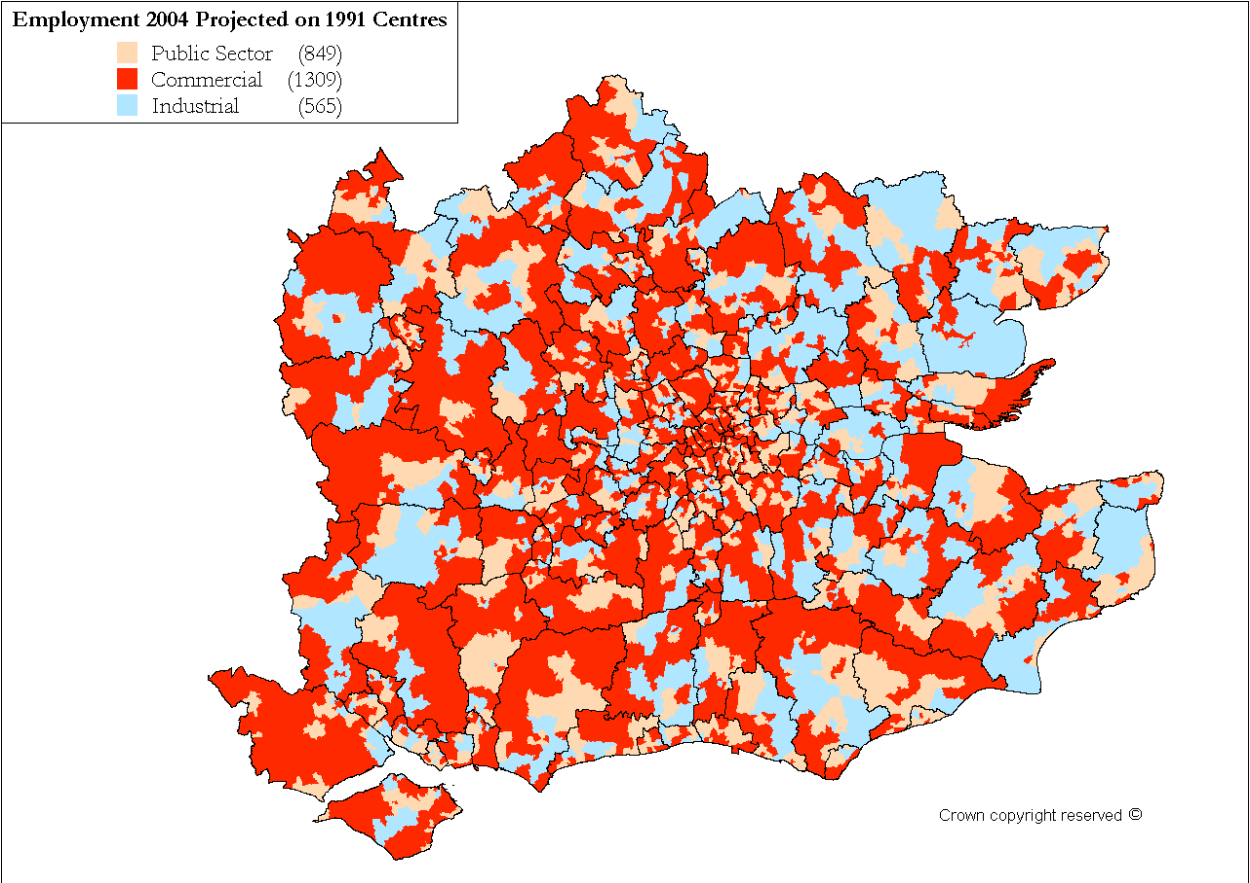


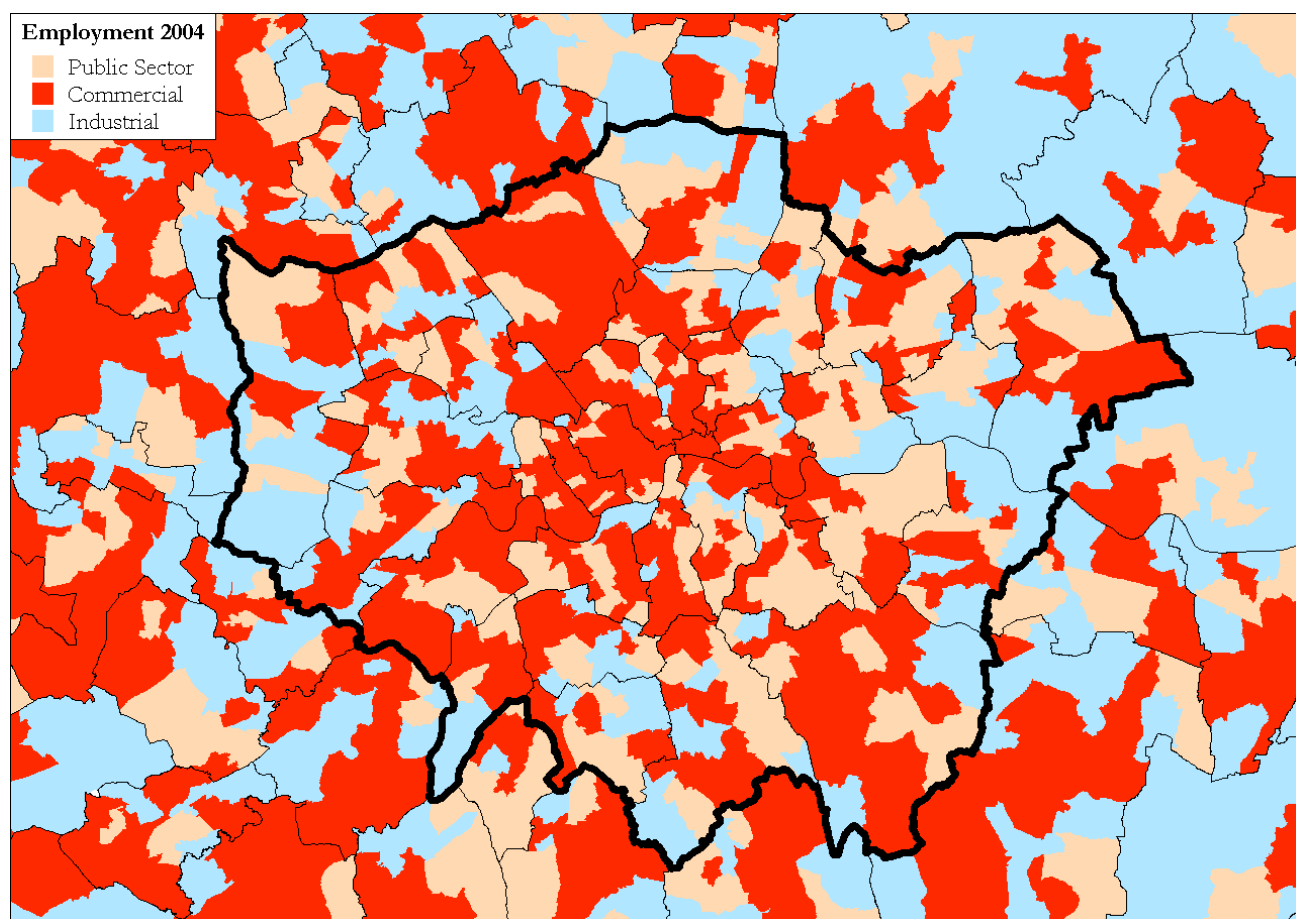
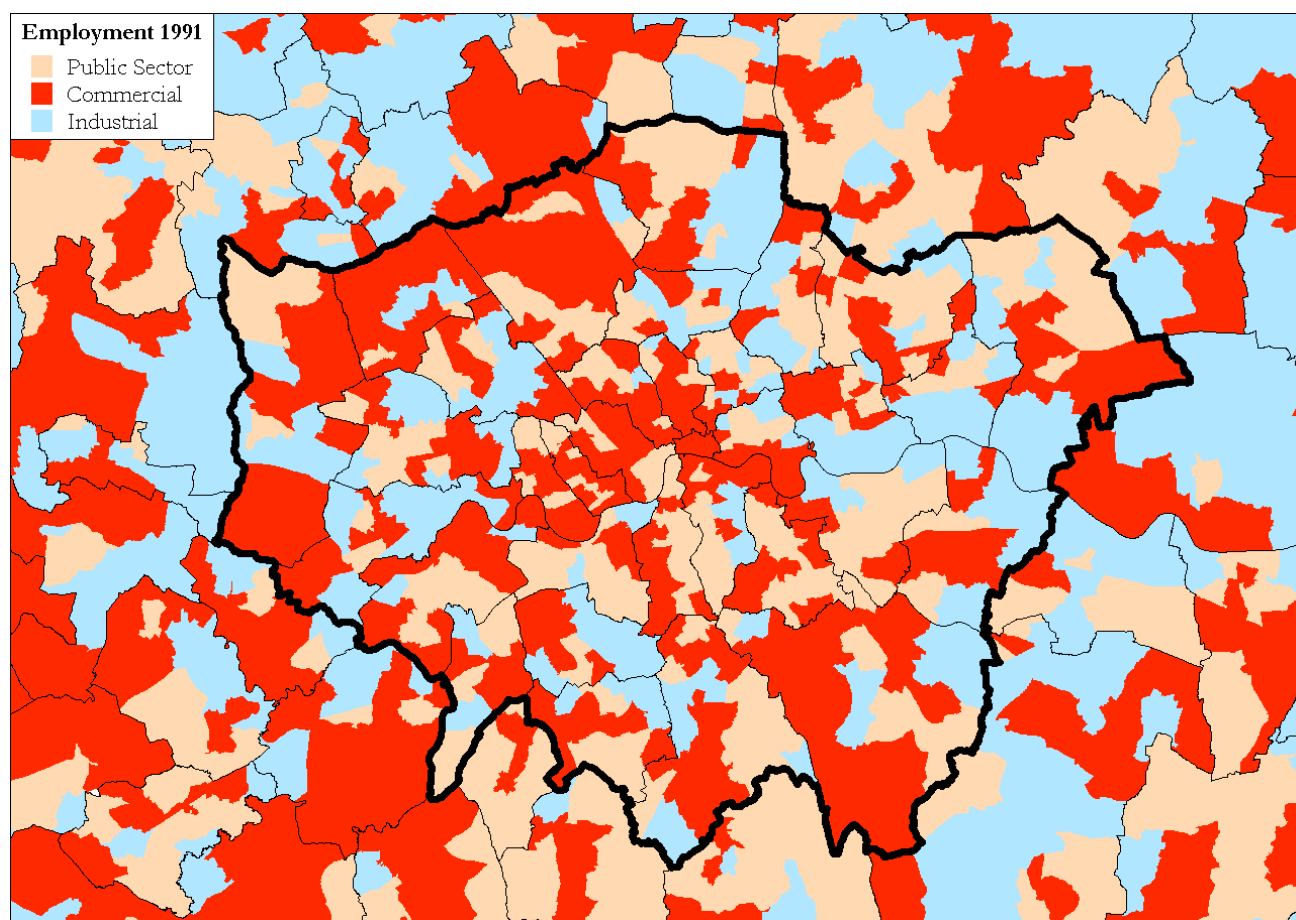


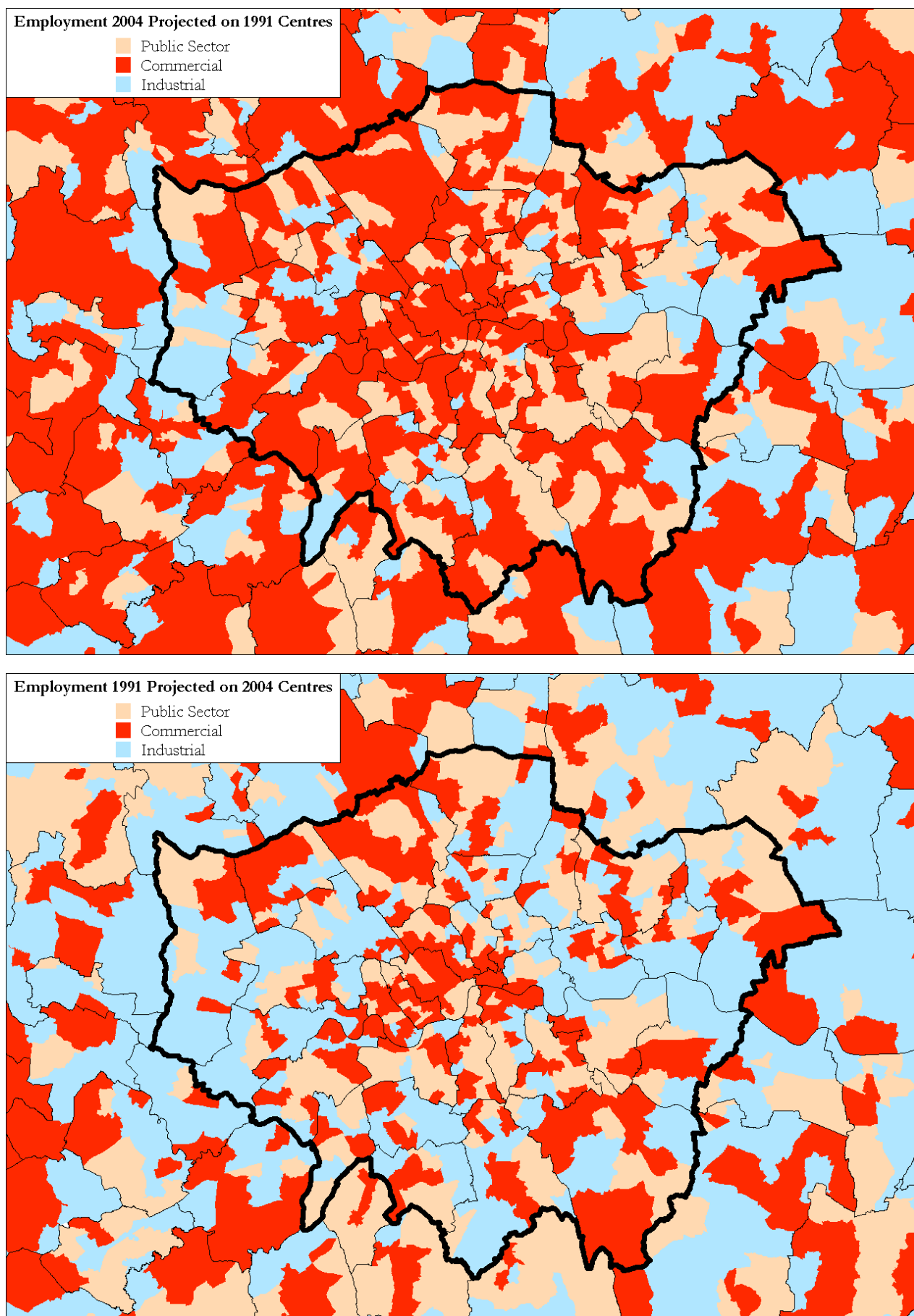












Appendix V: Degrees of Membership

