Environment and Planning A 1995, volume 27, pages 353-378

Visualizing changing social structure from a census†

D Dorling
Department of Geography, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, England
Received 22 December 1993; in revised form 17 March 1994

Abstract. In this paper it is shown how new visualization techniques are being used to analyze the first results of the British 1991 Census and other large data sets. Of interest here are questions about how localities have developed over time; which neighbourhoods have experienced gentrification and in which places have the recessions of the previous decades had their worst effects? Overall, do we see a picture of polarization or a levelling out of social disparities either locally or nationally? It is argued that these questions cannot be answered by conventional quantitative techniques because the answers are unlikely to be simple enough to be presentable in tables or by equations. Pictures are needed to show how different processes occur in different places, and holistic patterns need also to be seen without generalizing out the detail. Neither traditional thematic mapping nor commercial geographic information systems can do this well. Spatial visualization is an alternative approach in which the researchers choose what they wish to see and how they wish to view it. Many problems require new methods of visualization for their exploration. A new census presents us not only with new statistics, but also with the opportunity and impetus to develop radically different ways of envisioning information to reveal more fully the human facts contained within a mass of social statistics.

Introduction

> "... space is not viewed as an existent, physically separable from process, but as an aspect of process which is apprehended in a different and mysterious way, and by a different kind of empirical science."
>
> (Blaut, 1961, page 4)

Why draw maps if we wish to understand how a country is changing socially? It can be argued that the spatial patterns of a population reflect the social structure. These patterns reflect a reality that impinges upon everyday lives; lives whose courses are governed by that social structure, a social structure which is changing spatially. Pred (1986, page 198) has argued that social and spatial structures "feed upon one another" to, in a literal sense, eventually become each other. Knowing where the rich and poor used to live, and now live, tells us as much about the changing social relationships between rich and poor as it does about their changing individual geographies.

What makes a census valuable for social research is not the few questions that are asked of a population, but the fact that almost all the people are asked those questions simultaneously. Only a few dozen answers are reported for these questions, but these answers are given for hundreds of thousands of places. If it proves possible to reconstuct some social relationships from knowing who lives near whom, and how that is changing, that is a basis for a better understanding of why society is organised as it is, and how it is likely to fashion itself and be fashioned in the future.

† The paper is based on presentations to the Research on the 1991 Census Conference, Newcastle University, 13-15 September 1993 and to the 8th European Theoretical and Quantitative Geographical Colloquium, Budapest, 12-16 September 1993.

The aspects of social structure focused on here are status and reward through work. The British census did not ask people what they were paid, but it did ask them what work they did. By looking at where different occupational groups choose or are constrained to live in different places it may be possible to improve our understanding of society. What I try to begin to show in this paper is how the rich and poor of Britain live in loose ghettos of each other's making; and that, in places, the divides between these areas are widening.

These findings were not made through the use of statistical techniques. Most of those techniques require the researcher to know beforehand precisely what he or she is looking for: "Significance tests cannot tell us what to think" (Tufte, 1970, page 438). Instead new methods of visualization are used here which allow us to see simultaneously the detail and the whole—a technique Tufte (1990, page 37) now promotes. The technique is called visualization—making visible what cannot be imagined or seen. Through visualization reams of figures are transformed into a single picture which we can then study in detail (Tukey, 1965).

When applied to geographical data, visualization is not the same as mapping. This is because conventional maps do not make the characteristics of people visible. They do not allow us to see the shape of society. Maps are designed to show land and to be used for military purposes, seaward navigation, and the demarcation of territory (Harley, 1989)—not to allow us to visualize spatial social structure. The argument being supported here is that we need new kinds of maps, maps which stretch areas so that those "containing many people would appear large, and areas containing few people would appear small" (Tobler, 1973, page 215).

To visualize spatial social structure a 'distorted' geographical map of Britain is required. It is possible to show the characteristics of thousands of neighbourhoods such that each neighbourhood is visible as a distinct entity. The new map is an area cartogram, the idea of which is not in fact at all new (Wallace, 1926; Raisz, 1934). Detail within cities can be identified only with difficulty on traditional (choropleth) maps. Different groups of people are given enormously varying degrees of prominence on such maps in accordance only with the amount of land area they are deemed to occupy (Williams, 1976).

On an area cartogram, however, every suburb and village becomes visible in a single image because they contain people. Each area is drawn in proportion to the number of people who live there so that each individual's circumstances receive equal representation. Areas containing few people are deliberately drawn small. Here area cartograms are used to illustrate the visualization of the spatially detailed results from the 1991 Census (including change from 1981). This is done not only to aid understanding of the census but also as an argument, by example, for why cartograms should be used in the cartography of social geography.

The argument for visualizing social geography in this way concentrates on issues of social justice (Dorling, 1994a). It should be possible, and is desirable, to give the characteristics of each individual shown in a map equal prominence. If a map of one spatial facet of a society claims, in its title, to show an aspect of that society, then the spatial patterns across everybody in that society are what should be shown—without extreme overemphasis and underemphasis of particular groups of people for no better reason than that the streets, houses, towns, or villages that they live in occupy much more or less land. A more practical argument is simpler—on a detailed conventional choropleth map, even with a magnifying glass, you often cannot see what the map purports to show. A well-known example is of a map of the 1974 Canadian Federal Election which was criticized because more than half the constituencies were 'indecipherable' [because they were drawn so small on the traditional projection used (Coulson, 1977, page 102)]. This map was produced, despite there being a detailed officially sponsored population cartogram of Canada available (Skoda and Robertson, 1972).

Examples of where cartograms have been used successfully range from epidemiology (for example, Howe, 1970) to political science (for example, Johnston et al, 1988). Conventional maps of neighbourhoods or even of towns and counties, are next to useless at the level of the nation because they hide the residents of dense urban areas while massively overemphasizing the characteristics of those living in the countryside. In social geography the interest is often in processes that are occurring across a nation, but most people live and move around in quite small areas. A default principal in mapping human geography should be that the population of each area is represented equitably.

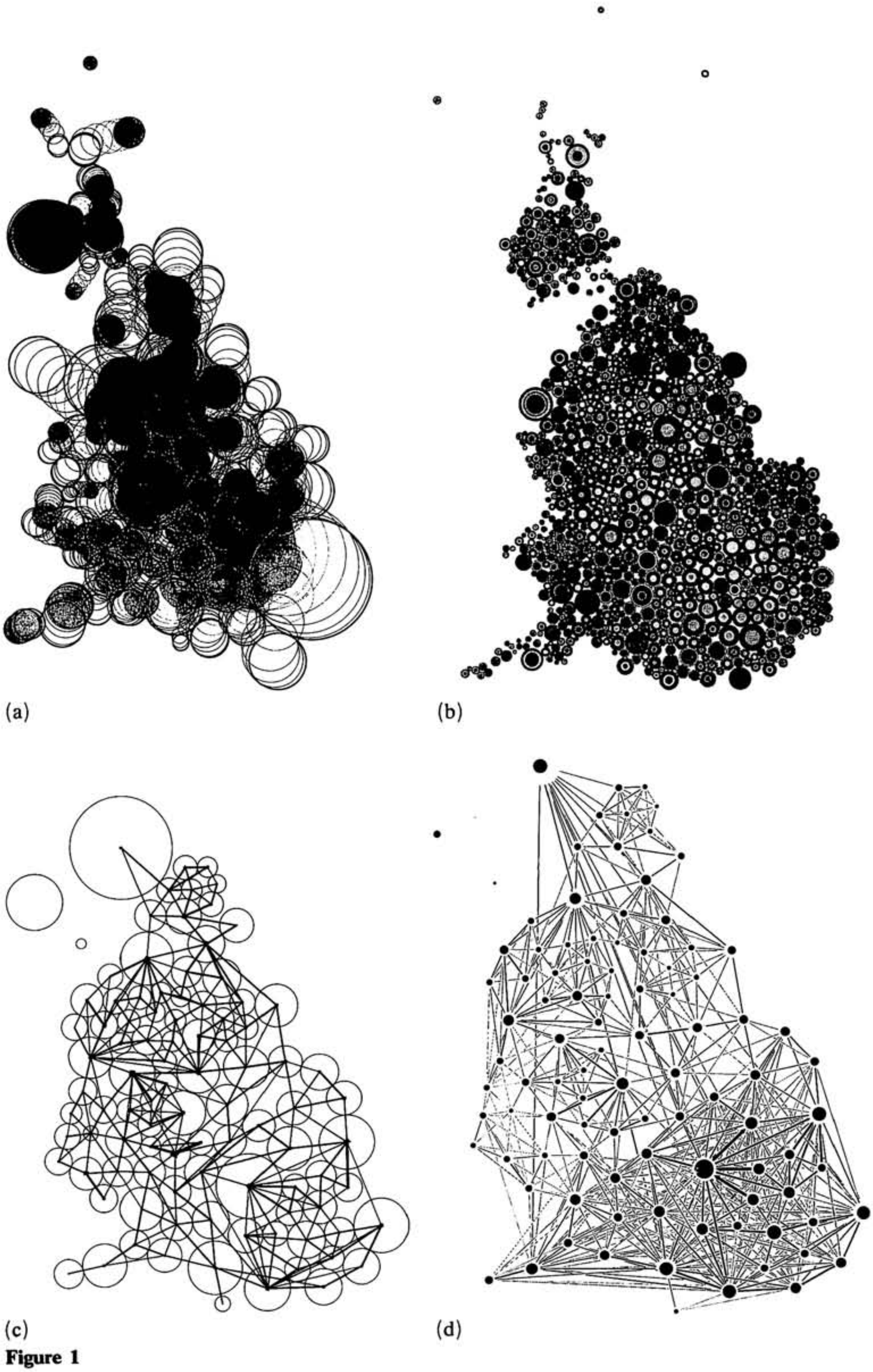
As visualization means making visible what can not easily be imagined or seen, the spatial structure of Britain's social geography is an ideal subject for this methodology. The aim is to grasp simultaneously the detail and the whole picture in full. To achieve that aim a population cartogram created from the human geography of the population of Britain is used here. A population cartogram is the appropriate base for seeing how social characteristics are distributed spatially across people rather than land.

Area cartograms

Figure 1(a) shows the evolution of a noncontinuous area cartogram of the sixty-four standard counties (and Scottish regions) of Britain. The areas, as circles, appear to spring into place. Figures 1(b) to (f) illustrate various graphical uses to which the cartogram can be put, ranging from change and flow mapping, to depicting voting swings by arrows, or the social characteristics of places with a crowd of Chernoff faces [some of these techniques are developed further in Dorling (1992); they are included here to illustrate the range of graphical uses to which cartograms can be put]. Area cartograms can either be continuous, meaning that they preserve the original geographical topology, or noncontinuous (Härö, 1968; Olson, 1976). Noncontinuous cartograms can produce simpler images more suitable for visualization, the continuous ones are mathematically more elegant but usually rather more confusing to look at.

The true value of the noncontinuous area cartogram is not in producing maps of a few hundred places which manual solutions could achieve. Instead, computer-generated cartograms can now clearly make visible the social structure of thousands of neighbourhoods on a sheet of paper. Figures 2(a) and (b) (see over) use an equal land area map to show administrative boundaries and figures 3(a) and (b) (see over) show the same boundaries on a population cartogram. Each of over ten thousand wards are visible on the cartogram and there is enough space in which to name major cities, which can only appear as dots on a conventional map of Britain where their names have to be put in the countryside to be seen!

The cartogram is based on 1981 ward boundaries and topology because the 1991 digital boundaries have still to be released to academics and also because this is a suitable base for studying changes over time. An iterative procedure was developed to transform an equal land area projection into an equal population area projection which maintained the original geographical contiguity as far as possible (Dorling, 1991). Data from the 1991 Census are mapped on these units by means of a lookup table from 1991 enumeration districts and output areas to 1981 Census



(a) (b) (c) (d)

Figure 1

(e)

(f)

Figure 1. (a) The evolution of a cartogram of the British counties and the regions of Scotland. The grey circles lie on the original geographic centroids, successive rings depict the movement of the circles following the progress of the cartogram algorithm. The empty circles shown with thick outlines delimit the final resting places of these 64 areas to form a noncontinuous equal population area county-level cartogram of Britain. **(b) Unemployment rates over twelve years shown by concentric rings in 'Office Areas'.** Outer rings depict levels in more recent years. Dark shading of a ring implies high relative unemployment rates for that year and area. Shown like this the pattern is very complex. An enormous amount of information can be presented on a cartogram, but showing patterns across time as well as space is not at all simple in a single flat image. **(c) Population cartogram of United Kingdom Health Service Areas showing connectivity.** The original contiguity could not be maintained in all places; difficult cases, however, occur infrequently. The thickness of the lines between areas is proportional to the length of their common border as a proportion of their total perimeters. This cartogram is the base map for figure 1(d). The two unconnected circles represent Northern Ireland and the Isle of Man. **(d) Significant migration flows drawn between Health Service Areas.** Shown by lines the width of which varies with the size of flow. These lines are not drawn for very low flow rates. The flow rates can be seen to be higher between areas in the South and to be very low, in general, across the Pennines. Scotland is shown as a single area at the top of the image. It is most closely connected, in migration terms, with parts of London. **(e) Swings of votes between three parties in two elections and 633 constituencies.** Shown by arrows giving the direction and magnitude of swings, left to the Labour Party, right to the Conservative Party and up showing swings to the Liberal Democrat Party. The swings shown here are for the votes between the 1987 and 1992 elections. In general the country swung towards Labour (southwesterly arrows dominate), but in Scotland the gross trend was generally equal for both Labour and the Conservatives, shown as arrows moving away from the Liberals (downwards). **(f) Social characteristics of constituencies shown by Chernoff faces.** The shape of the mouth is determined by employment levels, the cheek size by wealth, the position of the eyes by age of local industries, the size of the nose by electoral turnout, and the size of the face by population. Five variables are thus depicted for each constituency. Positioning Chernoff faces on a conventional map requires a great many insets. This kind of cartogram is well suited to show areas as faces, although the usefulness of these symbols is debatable.

wards (Atkins et al, 1993). Use of the 1981 wards also simplifies the analysis of information on change over time[1].

The algorithm to construct this cartogram is available for academic use (and is listed in the appendix). It differs from other published algorithms (for example, Tobler, 1973; Dougenik et al, 1985; Selvin et al, 1988) in that a noncontinuous cartogram is produced. This is done to create a new map of many areal units which satisfies the mathematical requirements of area being proportional to population, while also being cartographically useful. The algorithm begins by placing each area at its original geographical location, but as a circle of the correct size. Each area

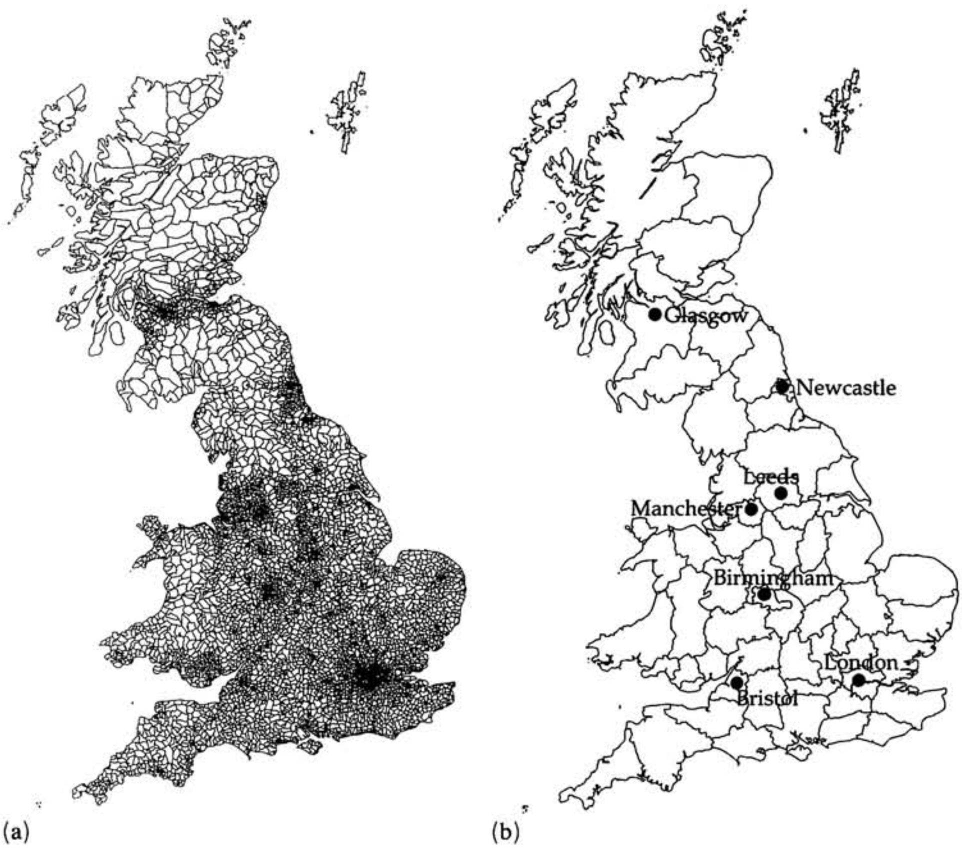


Figure 2. (a) **Ward boundaries on an equal land area map of Britain.** The vast majority of wards are simply not visible on a conventional map no matter how many insets are drawn. Most of the areas which can be seen contain the fewest people. These are the boundaries used in taking the 1981 Census and so in Scotland the areas are 'part-postcode-sectors'. This is also the geographical base from which figure 3(a) was created and which is used for depicting most of the geographical distributions shown here. **(b) County and Scottish region boundaries on an equal land area map of Britain.** Even the largest cities appear small on a conventional map of the country. This figure serves as a key to figure 3(b) and also shows the areas which were used as the base map from which figure 1(a) was created. To construct the topology matrix for Britain, which is needed in the creations of the area cartograms shown here, the locations of major bridges and tunnels are also required as well as these geographical boundaries.

[1] The 1991 Census data which were used to produce the illustration shown here have been deposited with the ESRC Data Archive at the University of Essex, which now holds the complete set of 1991 Census Small Area Statistics aggregated from enumeration districts to the best fitting boundaries of 1981 Census wards for all of Britain, in the SASPAC 91 file format.

then simultaneously repels those with which it overlaps and is attracted by those with which it should share a common geographical boundary. This process is repeated until all the circles have reached a stable position in which none of them overlap, and in which as many as possible are still contiguous with their original geographical neighbours.

Figures 4(a) and (b) (see over) show the ward cartogram being used to illustrate the spatial distribution of ethnic minorities in Britain. On the ward map it appears that almost everyone is white, with the most significant feature being two ghettos in the mountains of Scotland. This map is completely misleading, as are all maps of detailed social geography based on an equal land area projection. Most people in Britain live in neighbourhoods where more than 1% of the residents classified themselves as belonging to ethnic minorities. The most significant concentrations

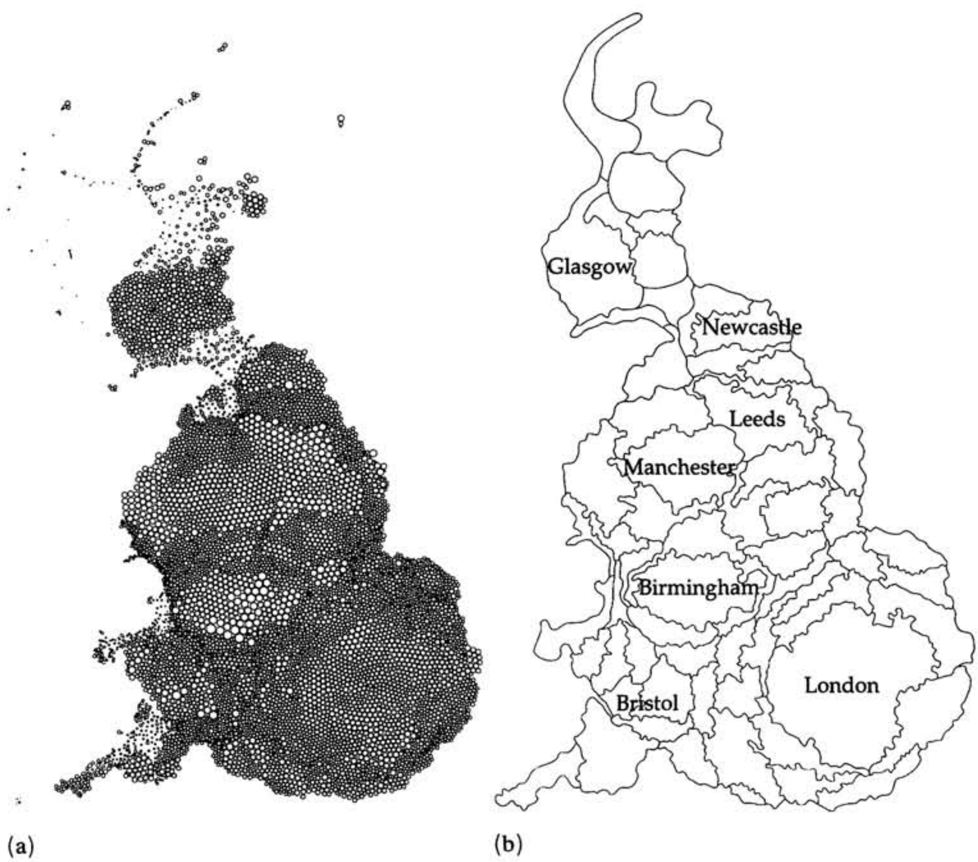


Figure 3. **(a) Ward boundaries on an equal population cartogram of Britain.** Each ward is shown by a hexagon the area of which is in proportion to its population. The hexagons do not touch each other in a few areas of very low population density. Hexagons are used here to illustrate how shapes other than circles could be employed and also to show how the algorithm leads to wards being positioned in arrangements which tesselate hexagonally. This is fortunate as the mean number of geographical neighbours for all British wards is 5.69. **(b) County and Scottish region boundaries on an equal population cartogram of Britain.** The county boundaries are drawn around the wards as they are positioned on the cartogram. Their topology can be seen to be intact except that the Isle of Wight becomes attached to Hampshire as the ferry link is defined as a contiguity constraint [see figure 2(b)]. Each city's name is placed within its respective county, for example the boundary around Leeds is of the wards of West Yorkshire.

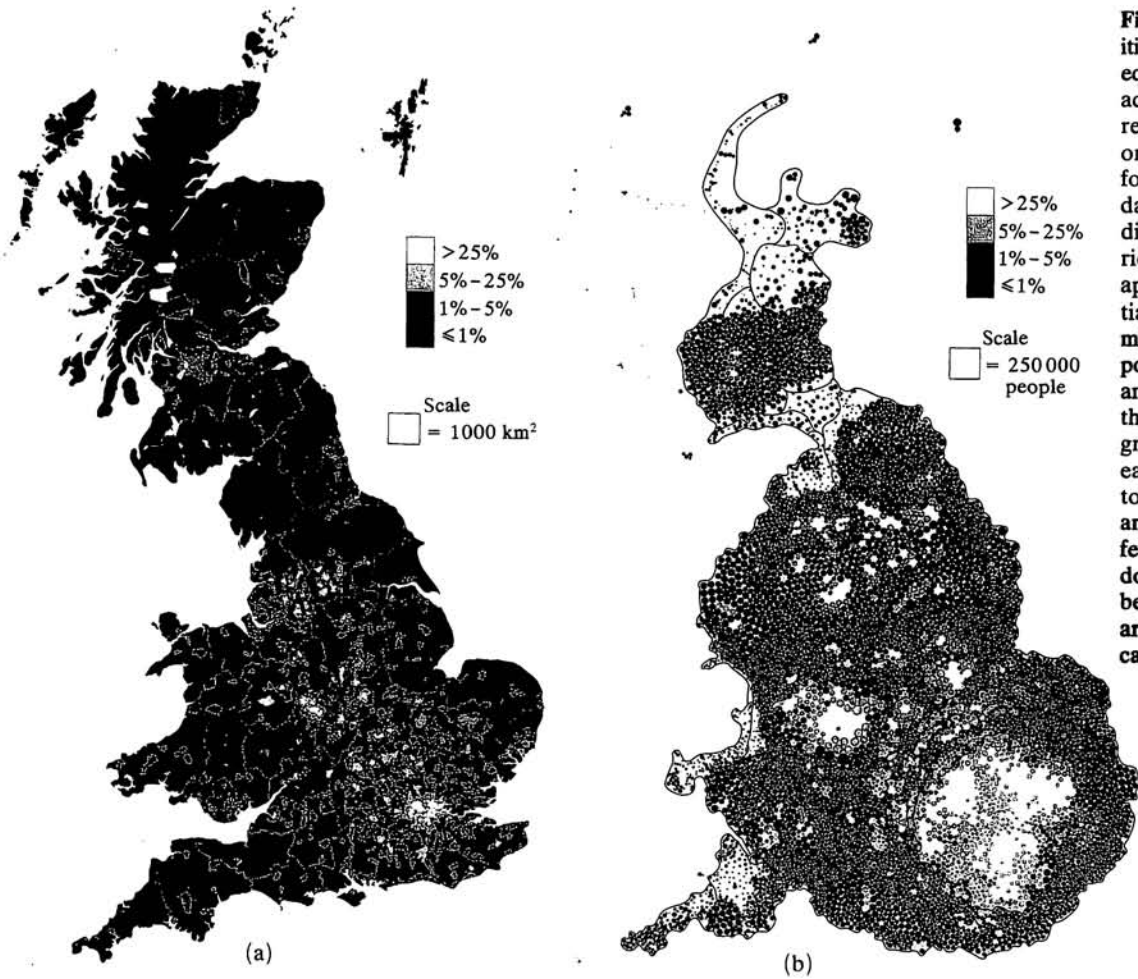
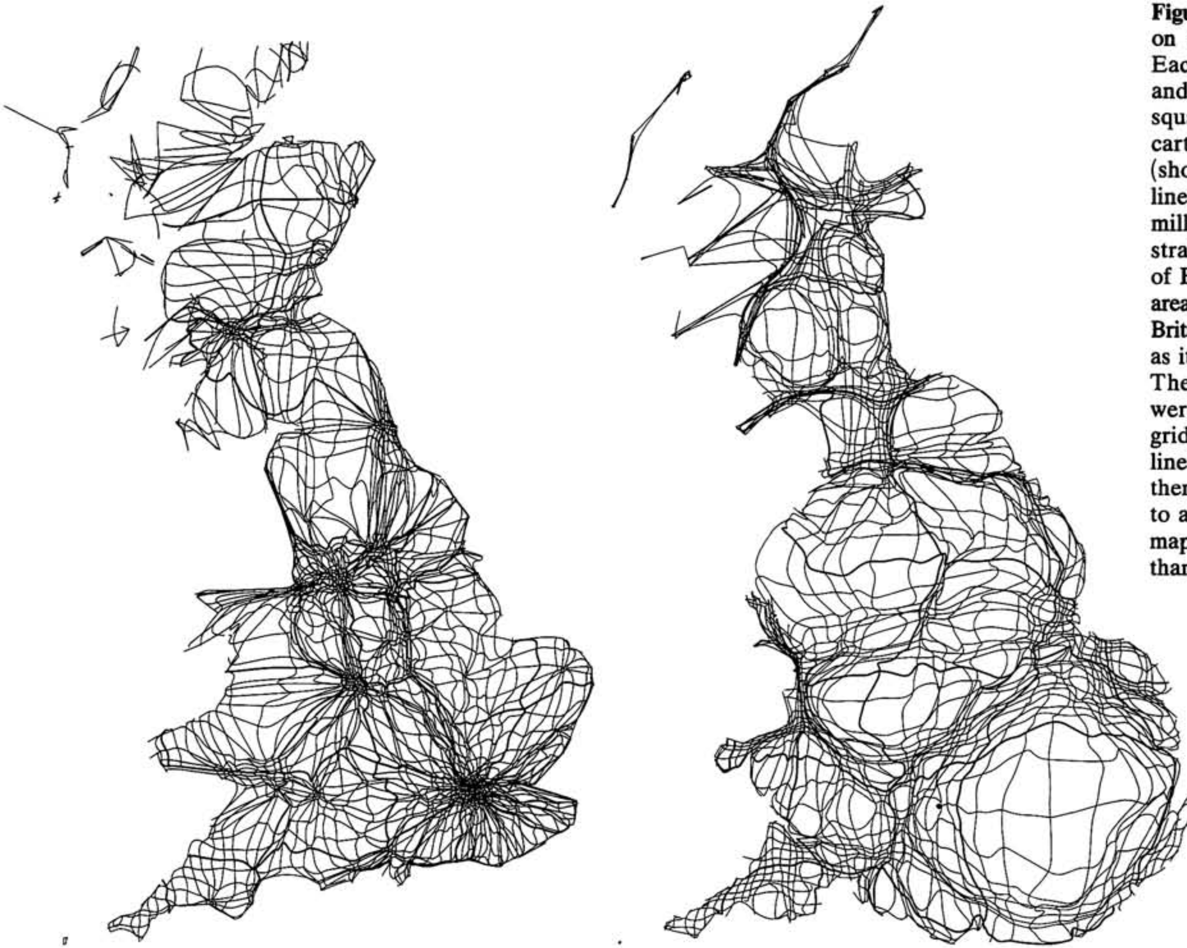


Figure 4. (a) Residents in ethnic minorities by ward in Britain, 1991, on an equal land area map. Wards are shaded according to the proportion of their residents who claimed to have an ethnic origin other than 'White' on the census form. County and Scottish region boundaries are shown by white lines. This diagram uses the same shading categories and areal units as figure 4(b) but apparently shows a very different spatial distribution. **(b) Residents in ethnic minorities by ward in Britain, 1991, on a population cartogram.** The shadings and variable shown in this diagram are the same as for figure 4(a). The cartogram allows the population sizes of each of Britain's ethnic minority areas to be compared. Sharp spatial divides are made evident and the areas in which fewer people live are not allowed to dominate the image. It should be remembered that in remote rural areas there are gaps between the wards on this cartogram.



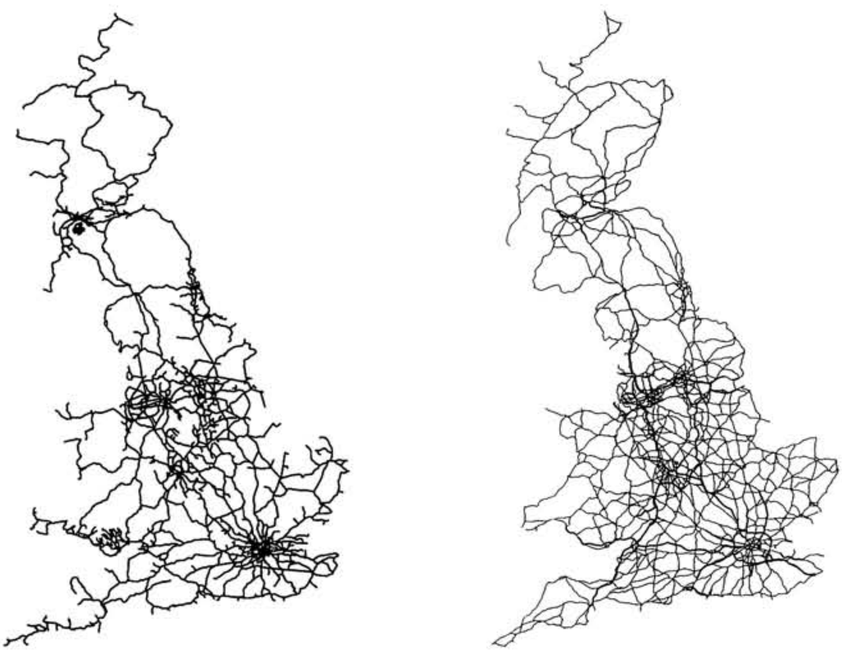
(a) (b)

Figure 5. **(a) An equal population grid on an equal land area map of Britain.** Each 'square' contains 30000 people and is constructed by reprojecting a square grid drawn on a population cartogram of 130000 census districts (shown in Dorling, 1991). The bold lines delimit squares containing up to 3 million people. If this grid is 'pulled-straight' an equal population cartogram of Britain is created. **(b) An equal land area grid on a population cartogram of Britain.** Each 'square' contains 100 km^2 as it appears on a population cartogram. The 10 km lines of the national grid were used. The 100 km-wide national grid squares are drawn with thicker lines. If these lines are 'pulled-straight' then the cartogram would revert back to a traditional (almost) equal land area map. Only grid-lines across land (rather than across sea) are drawn.

are in Birmingham, Leicester, Manchester, Leeds, and three parts of London, where 'minorities' comprise more than a quarter of the population of some areas. Conventional maps are biased in terms of which social group's neighbourhoods they conceal. The two Scottish ghettos are no different from numerous other wards, apart from being sparsely populated and therefore occupying a large tract of land and so dominating part of the image. Several other examples of this problem, illustrated with information from the 1991 Census, are given in Dorling (1994b).

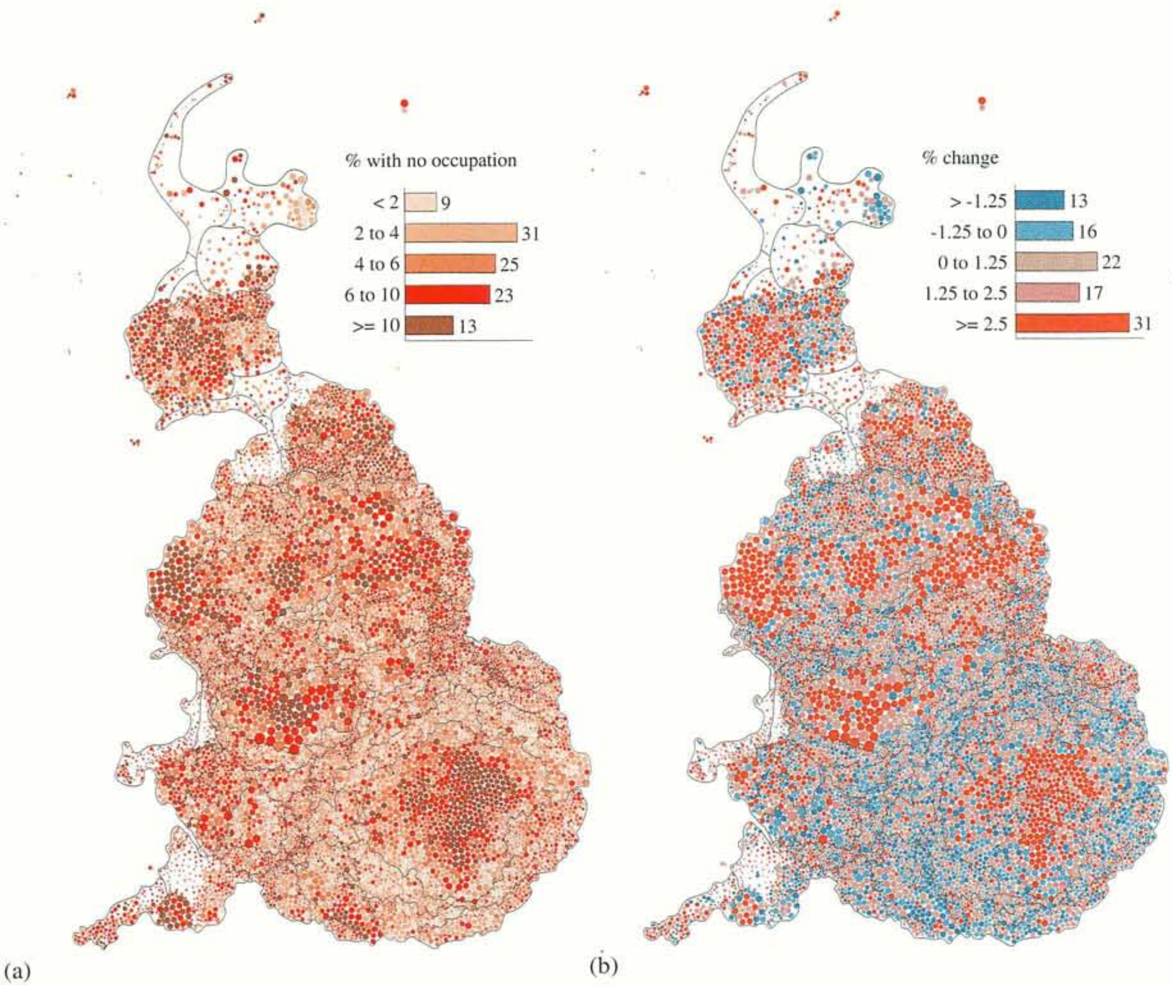
The algorithm which has been used to construct these ward cartograms can also be used to create cartograms of over 100 000 areal units. To show social characteristics effectively upon these requires more space than is available here and also benefits from the use of colour (Dorling, 1993). Figures 5(a) and (b) have used such a cartogram as a base to illustrate the spatial distribution of people in Britain following the methodology of Tobler (1973) for the United States. Once a resolution such as this has been achieved, the cartogram can be viewed as a continuous transformation and used for the mapping of incidences of disease or, for instance, the smooth reprojection of conventional maps.

Figures 6(a) and (b) show the main rail and road routes in Britain on a conventional map projection. These have been reprojected onto the equal population cartogram in figures 7(a) and (b). This was done by transforming each vertex of every route to follow the movements of their nearest enumeration districts to a new



(a) (b)

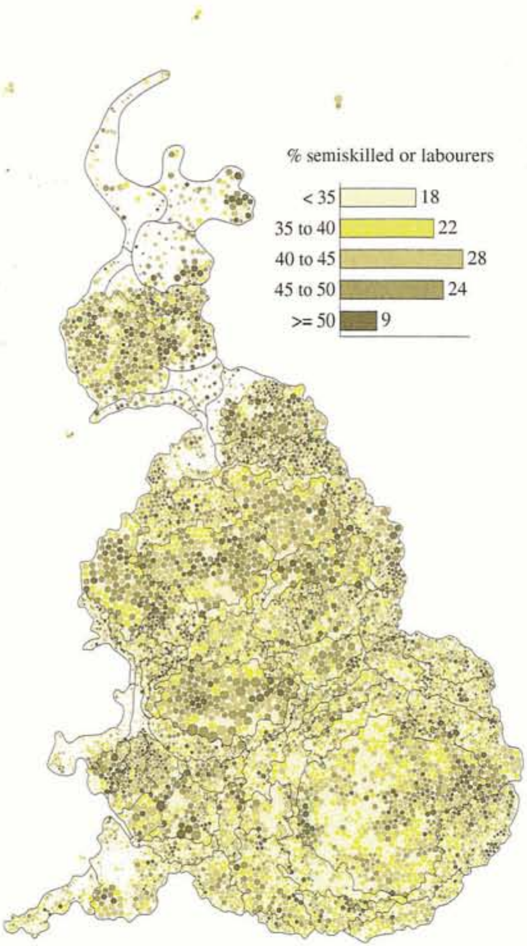
Figure 6. (a) The British rail mainline network on an equal land area map projection. The major routes of the railway system are shown, drawn as bold lines as they would be seen on a conventional map of Britain on which the coastline is not shown. On this projection they appear clustered in London and other major cities with a lack of provision in the more rural parts of Britain. (b) The British primary road network on an equal land area map projection. The main roads are shown by thin lines, the motorways are drawn with bold lines. This pattern appears to be more evenly spread across the land than that of the provision of the railways [figure 6(a)]. The motorways circle and connect the major cities of Britain and the main roads extend to reach almost every corner of the country.



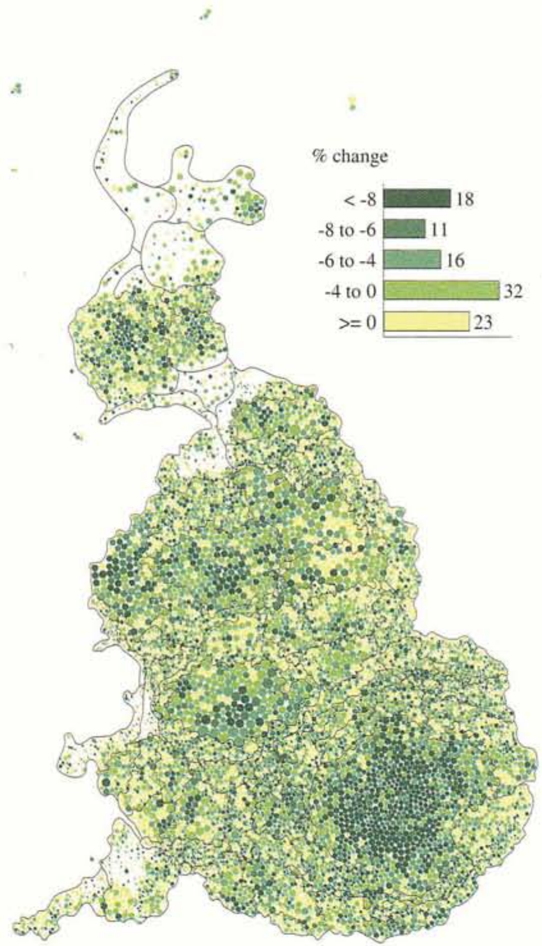
The numbers to the right of the histogram are the percentage of the work force in Britain living in each category of ward.

Scale □ = 250 000 people

Plate 1. *(a) Proportion of ward work forces with no occupation in 1991.* The histogram in the key to this figure shows facts such as that 13% of the British work force live in wards where more than one in ten 'economically active' people cannot state an occupation (having had no job in the last decade). The great majority of the group shown here are young people who have recently left school and who have yet to find their first job. The concentrations of areas where more than 10% of the work force have no occupation are stark. *(b) Change in ward work forces with no occupation, 1981-91.* 70% of the work force live in areas (grouped in the three lower classifications on the key) where the social group shown in plate 1(a) is growing in size. The spatial concentration of this change is very clear. It should be borne in mind that this distribution could be easily distorted by census underenumeration. However, any distortion is likely to lead to this illustration being an underestimation of the growth of this group (Dorling and Simpson, 1993).



(a)

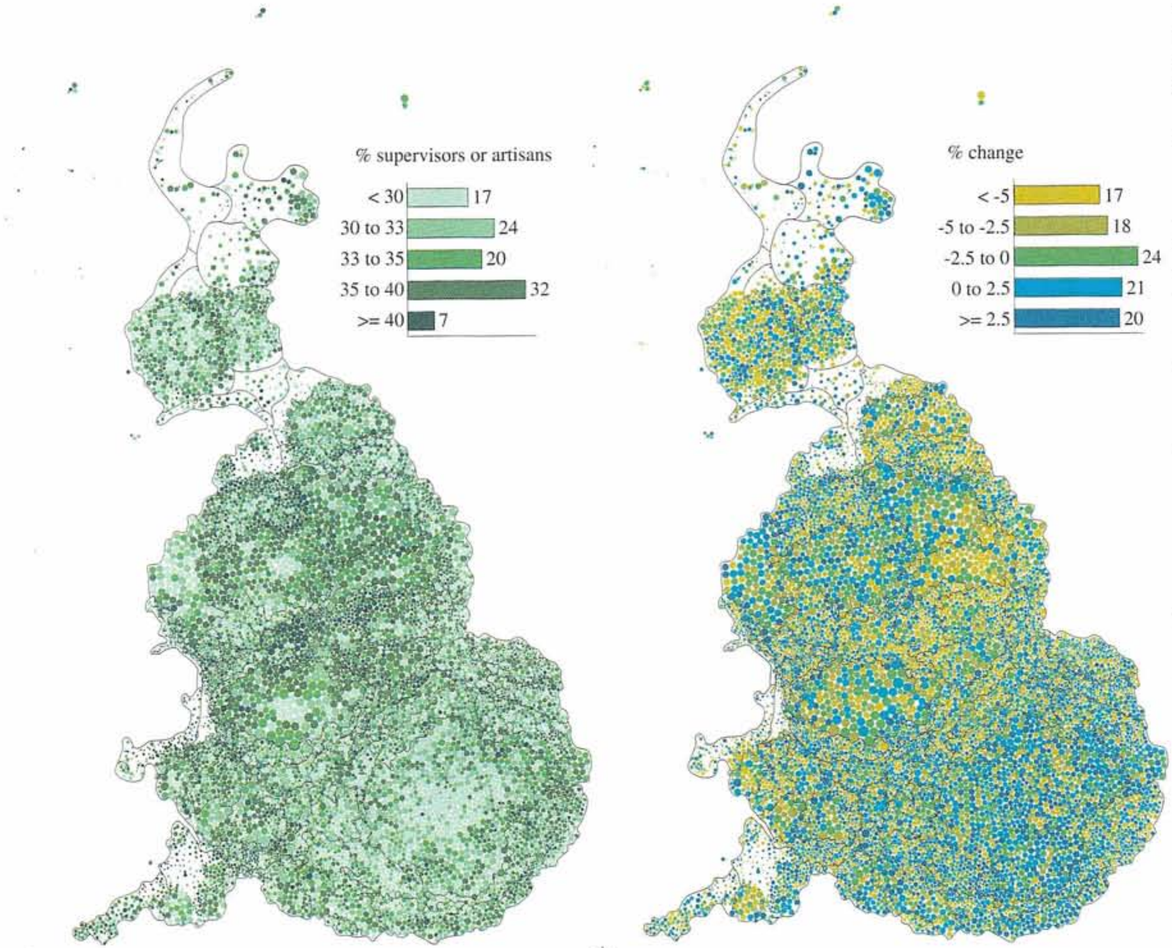


(b)

The numbers to the right of the histogram are the percentage of the work force in Britain living in each category of ward.

Scale ☐ = 250 000 people

Plate 2. *(a) Proportion of ward work forces who were semiskilled or labourers in 1991.* This group is lowest proportionally in the South East, excluding some concentrations in the East End of London and elsewhere. A strong degree of spatial autocorrelation is evident. This groups is made up of people who are working (or have worked) as junior nonmanual workers, personal service workers, semiskilled and unskilled manual workers, or as agricultural workers. *(b) Change in ward work forces made up of semiskilled or labourers, 1981-91.* The whole of London, central Birmingham, Liverpool, Manchester, and Glasgow have all seen falls of more than 8% in the size of this section of the work force. These falls coincide with the areas of increase of people with no occupation [shown in plate 1(b)]. It might well be that many of the children of people in this group in the 1980s have been unable to find any work in the 1990s in these areas, and that this has caused the group to appear to be diminishing in size.

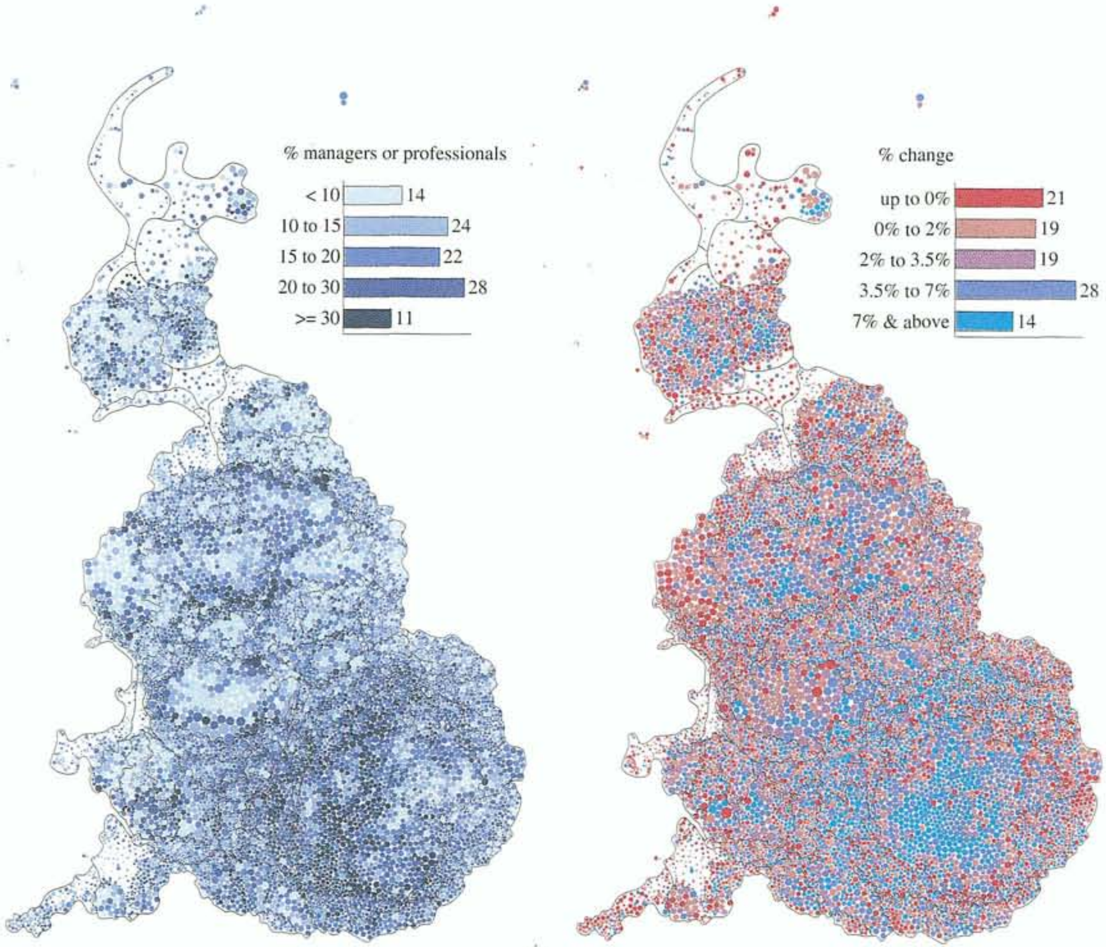


(a) (b)

The numbers to the right of the histogram are the percentage of the work force in Britain living in each category of ward.

Scale ☐ = 250 000 people

Plate 3. *(a) Proportion of ward work forces who were supervisors or artisans in 1991.* Northern towns tend to contain wards with the higher proportions of this group. The concentration of dark green wards in the centre is Stoke on Trent. This group includes people who stated their occupation as ancillary workers, artists, foremen, supervisors, skilled manual workers, own-account workers (other than professional), and farmers who are working on their own account. *(b) Change in ward work forces made up of supervisors or artisans, 1981-91.* There is very little pattern to this distribution, indicating that there is not much spatial structure to the changes. A southward movement can just be detected and, again, a movement away from the centres of cities. The key indicates facts such as that 20% of the work force of Britain live in areas where the proportion of people working in this group has increased by more than 2.5 workers in every hundred.



(a) (b)

The numbers to the right of the histogram are the percentage of the work force in Britain living in each category of ward.

Scale ☐ = 250 000 people

Plate 4. *(a) Proportion of ward work forces who were managers or professionals in 1991.* This group exhibits the most marked spatial structure. Areas of very low and very high concentrations dominate the image reflecting the 'geography of affluence' in Britain. Managers in industry and government, professionals who are either employed or self-employed, and farmers who are employers are included here. Apart from in Edinburgh and London, this group is generally found furthest away from the centres of cities. *(b) Change in ward work forces made up of managers or professionals, 1981-91.* Again the change is less clear than the static structure although a marked cluster of growth can be seen in the western half of London with smaller clusters elsewhere. Where there have been increases, those increases are often quite stark, with 14% of the population of Britain living in wards where more than seven extra people in every hundred are now in this group as compared with a decade ago. These changes could be partly a result of peoples' job labels changing rather than the nature of their work being different from before.

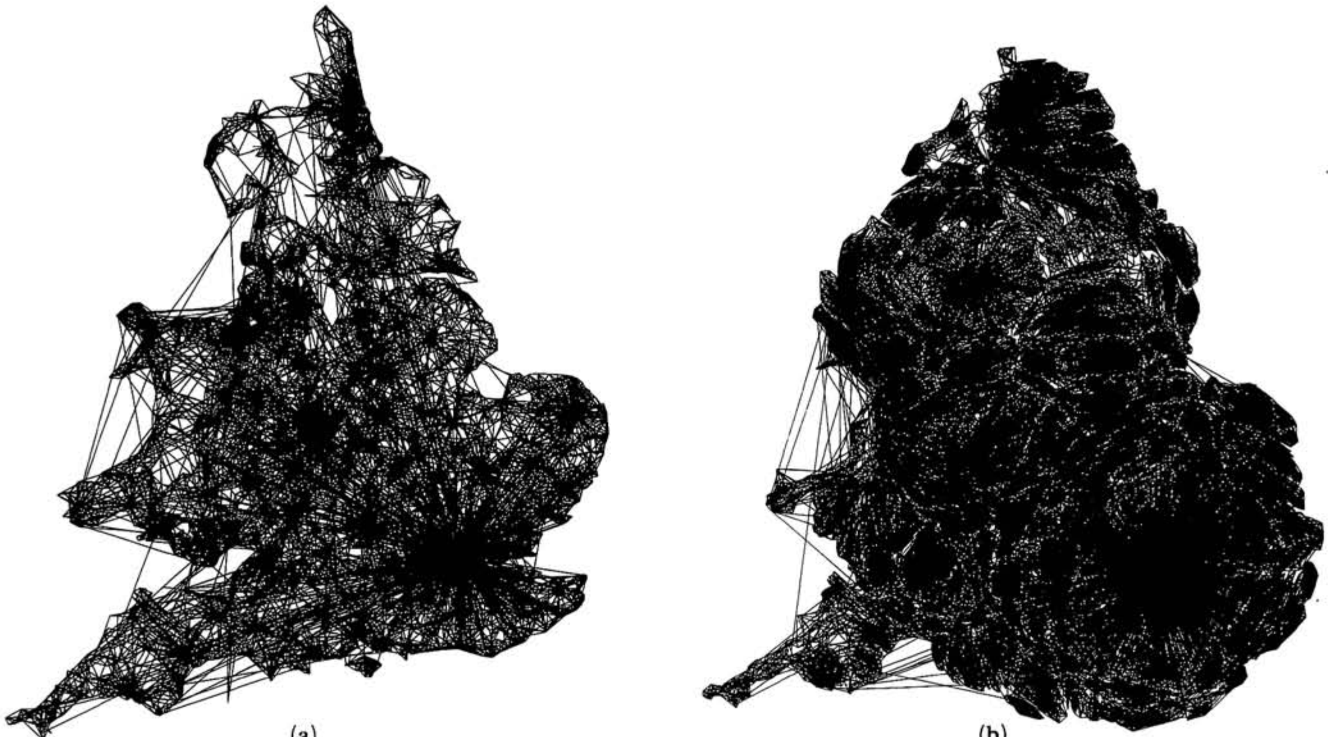
(a) (b)

Figure 7. (a) The British mainline rail network on an equal population cartogram. The same lines are shown as in figure 6(a) but here reprojected onto the equal population cartogram. The railway lines can be seen to be distributed extremely evenly across the population. This may be surprising in that most of these railways were laid down over one hundred years ago. Many rural routes have been removed since then, however, to produce the roughly equitable distribution made visible here. (b) The same lines are shown as in figure 6(b) reprojected onto the cartogram. Roads are more concentrated (per person) in rural areas, particularly just outside the major conurbations. The motorways appear as more jagged lines because of the constraint that they must connect correctly to all the major routes on this illustration. The provision of roads within London is obviously very sparse purely in terms of the numbers of people who actually live in the capital, let alone work there or travel in by car.

position on a population cartogram [based on the 1981 Census and detailed further in Dorling (1991)]. In 'population space' the railways of Britain can be seen to be distributed evenly across the people. As is evident from the figure, however, there is a distinct lack of roads per person in London. It is not surprising that that city sees the worst congestion in the country. The cartogram is based, like the census, on the population's places of residence. On a cartogram of daytime population the problem of too few roads for too many people would be seen to be even more stark. Figures 8(a) and (b) (see over) show those 'travel-to-work' flows on both bases. Cartograms, then, can give us a new view of a country. How therefore can they be used to understand better that country's changing society?

Polarization

For social scientists, one of the most interesting measures provided uniquely by the census is a count of people in different socioeconomic groups. Changes in the proportions of people in these various groups over time can indicate changes in the underlying social structure of the country: which areas are 'gentrifying'; where is 'residualization' occurring? There are many ways of measuring polarization and a relatively simple one is used here [for a summary and critique of the field see White (1983) and Smith (1993)].



(a) (b)

Figure 8. (a) Daily commuting flows between English and Welsh wards on a land map. Each flow of more than 2% of the employed population of a ward is drawn as a thin line to show, in total, the movement of 10 319 000 people—50% of all the commuters. This is the distribution taken from the census of 1981. The 1991 travel-to-work matrix has still to be released to the academic community at the time of writing. (b) Daily commuting flows between English and Welsh wards on a cartogram. Exactly the same lines are drawn as in figure 8(a). Whereas the conventional map shows the city structure, the cartogram shows the structure of flows within those cities containing the majority of commuter traffic. London and Manchester can be seen to represent the most centripetal forces in this system. This figure illustrates that, in terms of population, few people live in areas where they cannot meet many people travelling to work in the morning.

People's occupation was only tabulated for 10% of the population and, as some social groups are quite small and changes over time can be quite subtle, aggregations of enumeration districts have to be used in looking for spatial patterns of polarization (polarization meaning, here, 'the concentration over time of particular social groups in space'). For a discussion of the practical problems which can arise when studying these data, see Cole (1993) and Barke et al (1993).

Four categorizations of 'economically active' residents were defined for the purposes of this illustrative study. Managers and professionals [socioeconomic groups (SEGS) 1, 2, 3, 4, and 13], supervisors and artisans (SEGS 5, 8, 9, 12, 14), semiskilled and labourers (SEGS 6, 7, 10, 11, 15), and people with no occupation (all other 'economically active' adults including those in the armed forces and those on government schemes, but excluding students or 'housewives'). The definition and the derivation of the groups from which these categories are amalgamated is given in the official publications of the Office of Population Censuses and Surveys (OPCS, 1991; 1992), the argument for choosing these groups according to income is given in Dorling (1991).

Coloured plates 1 to 4 show, nationally, for each of the four groups, their contributions to the total economically active populations in 1991 and the 1981–91 percentage point shift of each of these by ward. Only five shading classes are used on each map for clarity and also because it is not possible to be certain of the exact proportions when using 10% sampled data drawn from a census which mislaid 1.2 million people (Dorling and Simpson, 1993; OPCS, 1994). In places even wards might be too small for taking such measurements with such data; although where similar characteristics can be seen in a spatial cluster of wards it can be assumed that a sizable concentration is present or that a significant change has occurred.

Despite there being just a few colours used, the patterns shown are quite revealing. The bulk of people with no occupation (economically active people who have never worked, have no stated occupation, or were or are in the armed forces) are to be found in the centres of London, Birmingham, Liverpool, Leeds, Manchester, Newcastle, and Glasgow. These are also the places where that group has increased in numbers most significantly (nationally they have increased on average by 30%). In contrast, there are very few semiskilled workers or labourers in London (apart from the East End) and, in general, that category of workers has been diminishing in numbers in large cities over the 1980s. The groups classified as supervisors and artisans are concentrated in the Midlands and the North West but have been gradually increasing in numbers towards the South. Managers and professionals are most densely concentrated in the home counties; however, their numbers have been growing strongly in London—the only large city in Britain to experience significant gentrification during the 1980s.

The last paragraph has just given a superficial impression of what can be gleaned from studying the eight colour cartograms in more detail. They do not show clear-cut answers because the reality that they seek to portray is not clear-cut, but they do attempt to show that reality without gross spatial bias. For quantitative analysis there is no substitute for being able to manipulate these images on the computer screen, experimenting with different categorizations and measures, shading techniques, and projections. Figure 9 shows the bivariate relationships between each of the four classifications using scatterplots created from measuring the density of dots which would have had to be drawn if the figures for wards had been plotted (as a simple scatter plot of 10000 points obscures more than it portrays).

Finally an attempt is made to depict visually one aggregate measure: within-area social polarization at the local level. In each ward, each social group can either be overrepresented or underrepresented in relation to the national average proportions; it can also either be increasing or decreasing in size more or less rapidly than the national average shift over the decade. If a group is overrepresented and increasing more rapidly, or underrepresented and increasing less rapidly (or decreasing), then that group can be said to be polarizing—moving towards an extreme. If a majority of the four social groups in a ward are polarizing, then that whole ward's population is polarizing, in terms of the occupational structure of the residents who live there.

Figure 10 shows those wards where at least three of the four social groups in a ward were polarizing. This figure, in effect, summarizes the colour plates (although the distribution in 1981 was used to measure the static structures for figure 10). The threshold chosen here, that at least a majority of the four groups are polarizing, is an important influence on the pattern shown. The figure shows that under this definition the majority of the population do not live in wards which are obviously polarizing, but there are clusters of wards—in the West End of London, Inner Birmingham, Manchester, and Aberdeen, for instance—which are becoming less heterogeneous and so can be more easily stigmatized as 'social ghettos' of one form or another.

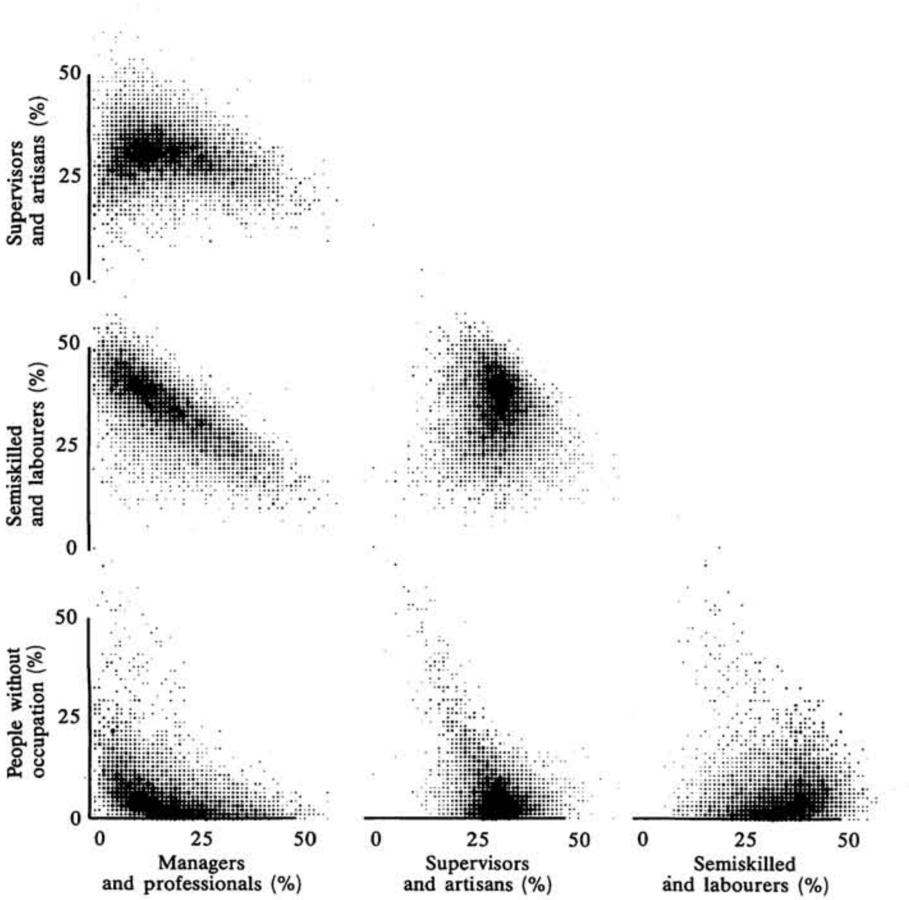


Figure 9. Social stratification of the British work force by occupation in wards in 1991. Circles are drawn on the scatter plots to show how many people live in areas with the various mixtures of the work-force categories shown. The area of each point is in proportion to the population of the wards it characterizes. It is evident from the degree of dispersal in these illustrations that the spatial relationship between the four social groups are not clear-cut.

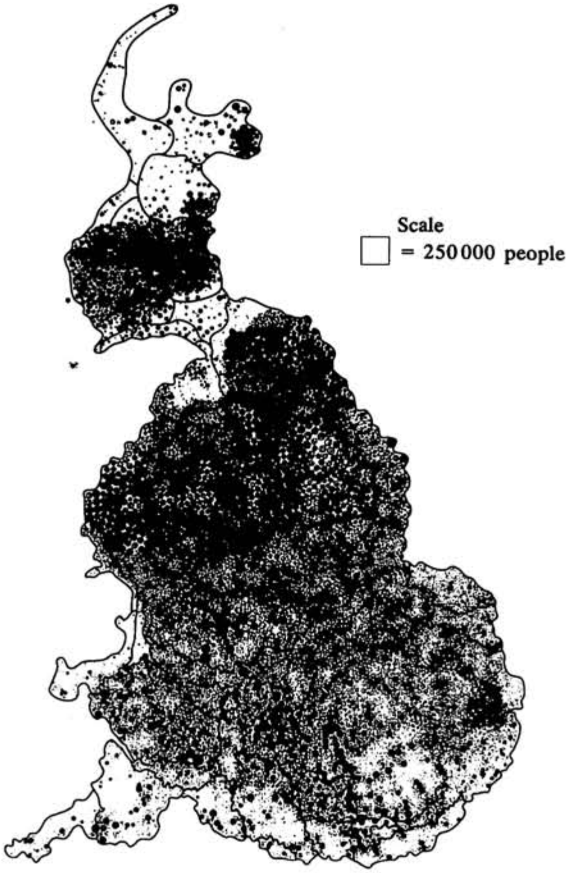


Figure 10. Occupational polarization of residents in Britain by ward, 1981–91. Wards with increasingly distinctive work-forces appear here highlighted by shading black those places where the proportion of at least three of the four occupational categories have moved significantly *away* from the national average over the decade. These wards are thus becoming more distinctive places socially. This map, however, does not show in what way that is happening—which social groups are becoming more or less numerous in which places.

Conclusion

The creation and use of high-resolution population cartograms make spatially detailed visualization of social information from the censuses possible. The images created can encompass the detail of the whole country in a single picture. The age-old constraints that come from conventional projections are broken as we move beyond the traditional map to choose how we wish to view the spatial structure of society (Goodchild, 1988). Conventional projections are not only uninformative, they are unjust—exaggerating the prevalence of a few people's life-styles at the expense of the representation of those who live inside cities, and hence presenting a biased view of society as a whole. Just as we can, to an extent, choose what we wish to draw, we should choose more carefully how we wish to draw it. In census mapping we have a choice over whether to see the people or the land.

Acknowledgements. Thanks are due to Tony Champion, Mike Coombes, James Cornford, Peter Taylor, and two anonymous referees who commented on earlier versions of this paper. This work was funded under a British Academy Postdoctoral Fellowship and also through research grants from the Joseph Rowntree Foundation, Newcastle University, and the Economic and Social Research Council (H507255135). The census data used in the analyses presented here are, as always, Crown Copyright.

References

Atkins D, Charlton M, Dorling D, Wymer C, 1993, "Conecting the 1981 and 1991 censuses", WP 93/9, North East Regional Research Laboratory, Department of Geography, University of Newcastle upon Tyne, Newcastle upon Tyne

Barke M, Henderson P A, Morphet C S, 1993, "Mapping the census" *Geographical Magazine* **65**(7) 36-39

Blaut J, 1961, "Space and process" *The Professional Geographer* **13**(4) 1-7

Cole K, 1993, "The 1991 local base statistics", in *The 1991 Census User's Guide* Eds A Dale, C Marsh (HMSO, London) chapter 8

Coulson M R C, 1977, "Political truth and the graphic image" *The Canadian Cartographer* **14** 101-111

Dorling D F L, 1991 *The Visualization of Spatial Social Structure* unpublished PhD thesis, Department of Geography, University of Newcastle upon Tyne (available from the British Library)

Dorling D F L, 1992, "Stretching space and splicing time: from cartographic animation to interactive visualization" *Cartography and Geographic Information Systems* **19** 215-227; 267-270

Dorling D F L, 1993, "Map design for census mapping" *The Cartographic Journal* **30** 167-183

Dorling D F L, 1994a, "Cartograms for visualizing human geography", in *Visualization and GIS* Eds D Unwin, H Hearnshaw (John Wiley, Chichester, Sussex) pp 85-102

Dorling D F L, 1994b, "Visualizing the geography of the population with the 1991 census" *Population Trends* **76** 29-39

Dorling D, Simpson S, 1993, "Those missing millions: implications for social statistics of the undercount in the 1991 census" *Radical Statistics Newsletter* number 55, 14-35

Dougenik J A, Chrisman N R, Niemeyer D R, 1985, "An algorithm to construct continuous area cartograms" *Professional Geographer* **37**(1) 76-81

Goodchild M F, 1988, "Stepping over the line: technological constraints and the new cartography" *The American Cartographer* **15** 311-319

Harley J B, 1989, "Deconstructing the map" *Cartographica* **26**(2) 1-20

Härö A S, 1968, "Area cartograms of the SMSA population of the United States" *Annals of the Association of American Geographers* **58** 452-460

Howe G M, 1970 *National Atlas of Disease Mortality in the United Kingdom* (Thomas Nelson, London)

Johnston R J, Pattie C J, Allsopp J G, 1988 *A Nation Dividing?* (Longman, Harlow, Essex)

Olson J, 1976, "Noncontiguous area cartograms" *The Professional Geographer* **28** 371-380

OPCS, 1991 *Standard Occupational Classification, Volume 3: Social Classifications and Coding Methodology* Office of Population Censuses and Surveys (HMSO, London)

OPCS, 1992 *1991 Census Definitions* Office of Population Censuses and Surveys (HMSO, London)

OPCS, 1994 *1991 Census User Guide 58: Undercoverage in Great Britain* Office of Population Censuses and Surveys (HMSO, London)

Pred A, 1986 *Place, Practise and Structure: Social and Spatial Transformations in Southern Sweden: 1750-1850* (Barnes and Noble, Totowa, NJ)

Raisz E, 1934, "The rectangular statistical cartogram" *The Geographical Review* **24** 292-296

Selvin S, Merrill D W, Schulman J, Sacks S, Bedell L, Wong L, 1988, "Transformations of maps to investigate clusters of disease" *Social Sciences in Medicine* **26** 215-221

Skoda L, Robertson J C, 1972, "Isodemographic map of Canada", geographical paper 50, Department of the Environment, Ottawa, Canada

Smith S J, 1993, "Residential segregation and the politics of racialization", in *Racism, the City and the State* Eds M Cross, M Keiths (Routledge, London) pp 128-143

Tobler W R, 1973, "A continuous transformation useful for districting" *Annals of the New York Academy of Sciences* **219** 215-220

Tufte E R, 1970, "Improving data analysis in political science", in *The Quantitative Analysis of Social Problems* Ed. E Tufte (Addison-Wesley, Reading, MA) pp 437–449
Tufte E R, 1990 *Envisioning Information* (Graphics Press, Cheshire, CT)
Tukey J W, 1965, "The future process of data analysis", in *Proceedings of the Tenth Conference on the Design of Experiments in Army Research Development and Testing* report 65-3, US Army Research Office, Durham, NC, pp 691–725
Wallace J W, 1926, "Population map for health officers" *American Journal of Public Health* **16** 1023
White M J, 1983, "The measurement of spatial segregation" *American Journal of Sociology* **88** 1008–1018
Williams R L, 1976, "The misuse of area in mapping census-type numbers" *Historical Methods Newsletter* **9** 213–216

APPENDIX

Cartogram algorithm†

```
program cartogram (output);

[Pascal implementation of the cartogram algorithm]

[Expects, as input, a comma-separated-value text file giving each zone's number, name,
population, x and y centroid, the number of neighbouring zones, and the number and border
length of each neighbouring zone. Outputs a radius and new centroid for each zone. The two
recursive procedures and a tree structure are included to increase the efficiency of the
program.]

[Constants are currently set for the 10444 1981 Census wards of Great Britain and for
15000 iterations of the main procedure—exact convergence criteria are unknown. Wards do
actually converge quite quickly—there are no problems with the speed of the algorithm—it
appears to move from O(n^2) to O(n ln n) until other factors come into play when n exceeds
about 100000 zones.]

const
  iters = 15000;
  zones = 10444;
  ratio = 0.4;
  friction = 0.25;
  pi = 3.141592654;

type
  vector = array [1 ... zones] of real;
  index = array [1 ... zones] of integer;
  vectors = array [1 ... zones, 1 ... 21] of real;
  indexes = array [1 ... zones, 1 ... 21] of integer;
  leaves = record
              id     : integer;
              xpos   : real;
              ypos   : real;
              left   : integer;
              right  : integer;
  end;
  trees = array [1 ... zones] of leaves;
```

† Reproduced for free academic use only, Daniel Dorling, 1994.

```
var
  infile, outfile               : text;
  list                          : index;
  tree                          : trees;
  widest, dist                  : real;
  closest, overlap              : real;
  xrepel, yrepel, xd, yd        : real;
  xattract, yattract            : real;
  displacement                  : real;
  atrdst, repdst                : real;
  total_dist                    : real;
  total_radius, scale           : real;
  xtotal, ytotal                : real;
  zone, nb                      : integer;
  other, itter                  : integer;
  end_pointer, number           : integer;
  x, y                          : index;
  xvector, yvector              : vector;
  perimeter, people, radius     : vector;
  border                        : vectors;
  nbours                        : index;
  nbour                         : indexes;
```

[Recursive procedure to add global variable 'zone' to the 'tree' which is used to find nearest neighbours]

```
procedure add_point(pointer,axis : integer);
  begin
    if tree[pointer].id = 0 then
      begin
        tree[pointer].id    := zone;
        tree[pointer].left  := 0;
        tree[pointer].right := 0;
        tree[pointer].xpos := x[zone];
        tree[pointer].ypos := y[zone];
      end
    else
      if axis = 1 then
        if x[zone] >= tree[pointer].xpos then
          begin
            if tree[pointer].left = 0 then
              begin
                end_pointer := end_pointer +1;
                tree[pointer].left := end_pointer;
              end;
            add_point(tree[pointer].left,3-axis);
          end
        else
          begin
            if tree[pointer].right = 0 then
              begin
                end_pointer := end_pointer +1
                tree[pointer].right :=  end_pointer;
              end;
            add_point(tree[pointer].right,3-axis);
          end
```

```
    else
        if y[zone] > = tree[pointer].ypos then
          begin
            if tree[pointer].left = 0 then
              begin
                end_pointer := end_pointer +1;
                tree[pointer].left := end_pointer;
              end;
            add_point(tree[pointer].left,3-axis);
          end
        else
          begin
            if tree[pointer].right = 0 then
              begin
                end_pointer := end_pointer +1;
                tree[pointer].right := end_pointer;
              end;
            add_point(tree[pointer].right,3-axis);
          end
  end;
```

[Procedure recursively recovers the 'list' of zones within 'dist' horizontally or vertically of the 'zone', from the 'tree'. The list length is given by the integer 'number'. All global variables exist prior to invocation.]

```
procedure get_point(pointer, axis :integer);
  begin
    if pointer > 0 then
      if tree[pointer].id > 0 then
        begin
          if axis = 1 then
            begin
              if x[zone] - dist < tree[pointer].xpos then
                get_point(tree[pointer].right,3-axis);
              if x[zone]+dist > = tree[pointer].xpos then
                get_point(tree[pointer].left,3-axis);
            end;
          if axis = 2 then
            begin
              if y[zone] - dist < tree[pointer].ypos then
                get_point(tree[pointer].right,3-axis);
              if y[zone]+dist > = tree[pointer].ypos then
                get_point(tree[pointer].left,3-axis);
            end;
          if (x[zone] - dist < tree[pointer].xpos)
            and (x[zone]+dist > = tree[pointer].xpos) then
            if (y[zone] - dist < tree[pointer].ypos)
              and(y[zone]+dist > = tree[pointer].ypos) then
              begin
                number := number +1;
                list[number] := tree[pointer].id;
              end;
        end;
  end;
```

[The main program]
```
begin
  reset(infile,'FILE = ward.in');
  rewrite(outfile,'FILE = ward.out');
  total_dist := 0;
  total_radius := 0;

  for zone := 1 to zones do
    begin
      read(infile,people[zone],x[zone],y[zone],nbours[zone]);
      perimeter[zone] := 0;
      for nb := 1 to nbours[zone] do
        begin
          read(infile,nbour[zone,nb],border[zone,nb]);
          perimeter[zone] := perimeter[zone]+border(zone, nb];
          if nbour[zone,nb] > 0 then
            if nbour[zone,nb] < zone then
              begin
                xd := x[zone]-x[nbour[zone,nb]];
                yd := y[zone]-y[nbour[zone,nb]];
                total_dist := total_dist+sqrt(xd*xd+yd*yd);
                total_radius := total_radius+sqrt(people[zone]/pi)
                                +sqrt(people[nbour[zone,nb]]/pi);
              end;
        end;
      readln(infile);
    end;
writeln ('Finished reading in topology');

scale := total_dist / total_radius;
widest := 0;
for zone := 1 to zones do
  begin
    radius[zone] := scale * sqrt(people[zone]/pi);
    if radius[zone] > widest then
      widest := radius[zone];
    xvector[zone] := 0;
    yvector[zone] := 0;
  end;
writeln ('Scaling by ',scale,' widest is ',widest);
```

[Main iteration loop of cartogram algorithm.]
```
for itter := 1 to iters do
  begin
    for zone := 1 to zones do
      tree[zone].id := 0;
    end_pointer := 1;
    for zone := 1 to zones do
      add_point(1,1);
    displacement := 0.0;
```

[Loop of independent displacements—could run in parallel.]

```
      for zone := 1 to zones do
        begin
          xrepel   := 0.0;
          yrepel   := 0.0;
          xattract := 0.0;
          yattract := 0.0;
          closest  := widest;
```

[Retrieve points within widest + radius(zone) of 'zone' to 'list' which will be of length 'number'.]

```
          number := 0;
          dist   := widest+radius[zone];
          get_point(1,1);
```

[Calculate repelling force of overlapping neighbours.]

```
          if number > 0 then
            for nb := 1 to number do
              begin
                other := list[nb];
                if other <> zone then
                  begin
                    xd := x[zone]-x[other];
                    yd := y[zone]-y[other];
                    dist := sqrt(xd*xd+yd*yd);
                    if dist < closest then
                      closest := dist;
                    overlap := radius[zone]+radius[other]-dist;
                    if overlap > 0.0 then
                      if dist > 1.0 then
                        begin
                          xrepel := xrepel-overlap*(x[other]-x[zone])/dist;
                          yrepel := yrepel-overlap*(y[other]-y[zone])/dist;
                        end;
                  end;
              end;
```

[Calculate forces of attraction between neighbours.]

```
            for nb := 1 to nbours[zone] do
              begin
                other := nbour[zone,nb];
                if other <> 0 then
                  begin
                    xd := x[zone]-x[other];
                    yd := y[zone]-y[other];
                    dist := sqrt(xd*xd+yd*yd);
                    overlap  := dist-radius[zone]-radius[other];
                    if overlap > 0.0 then
                      begin
                        overlap := overlap*border[zone,nb]/perimeter[zone];
                        xattract := xattract +overlap*(x[other]-x[zone])/dist;
                        yattract := yattract+overlap*(y[other]-y[zone])/dist;
                      end;
                  end;
              end;
```

[Calculate the combined effect of attraction and repulsion.]

```
            atrdst := sqrt(xattract*xattract+yattract*yattract);
            repdst := sqrt(xrepel*xrepel+yrepel*yrepel);
            if repdst > closest then
              begin
                xrepel := closest*xrepel/(repdst+1);
                yrepel := closest*yrepel/(repdst+1);
                repdst := closest;
              end;
            if repdst > 0 then
              begin
                xtotal := (1-ratio)*xrepel+ratio*(repdst*xattract/(atrdst+1));
                ytotal := (1-ratio)*yrepel+ratio*(repdst*yattract/(atrdst+1));
              end;
            else
              begin
                if atrdst > closest then
                  begin
                    xattract := closest*xattract/(atrdst+1);
                    yattract := closest*yattract/(atrdst+1);
                  end;
                xtotal := xattract;
                ytotal := yattract;
              end;
```

[Record the vector]

```
            xvector[zone] := friction*(xvector[zone]+xtotal);
            yvector[zone] := friction*(yvector[zone]+ytotal);
            displacement := displacement+sqrt[xtotal*xtotal+ytotal*ytotal);
        end;
```

[Update the positions.]

```
      for zone := 1 to zones do
        begin
          x[zone] := x[zone]+round(xvector[zone]);
          y[zone] := y[zone]+round(yvector[zone]);
        end;
      displacement := displacement/zones;
      writeln('Iter: ', iter, ' disp: ', displacement);
    end;
```

[Having finished the iterations write out the new file.]

```
    for zone := 1 to zones do
      writeln(outfile,radius[zone]:9:0,',',x[zone]:9,',',y[zone]:9);
  end.
```