Dorling, D. (1995) Visualizing the 1991 census, in S. Openshaw (ed), A Census User's Handbook, London: Longmans, 167-213.

Census User's Handbook Chapter X

Visualizing the 1991 census

Daniel Dorling
Department of Geography
Newcastle University

This is not a standard handbook chapter. It does not describe "best practice", take the reader through the options they have, or provide a succinct summary of the field. This is because there is, as yet, no accepted wisdom to the methodology of census visualization, no set of all-singing-all-dancing computer packages, no weighty academic legacy of papers on the subject. Visualization — the study of information through graphical means freed from manual restriction by the advances in scientific computing — is still very new. What this chapter aims to give is a flavour of visual techniques which can be used to study census data and which will become generally accessible in the decade following the 1991 census. First the arguments for visualization are made, examples are then given using the hierarchy of census geography to structure the chapter. Speculation on some possibilities for entirely new visual representations are given towards the end. The aim of this chapter is to show how visualization of the census can engage our imaginations in new ways to address age-old questions about society.

1 Visualization

> "We must create a new language, consider a transitory state of new illusions and layers of validity and accept the possibility that there may be no language to describe ultimate reality, beyond the language of visions."
>
> Denes 1979 p.3

Visualization is the dream that we can turn statistics into pictures which portray the complexity of the reality which we, from our experiences, know to exist. The British Census provides millions of numbers concerning the people of this country. These numbers can be reduced to a few totals, averages and deviations, but such simple statistical manipulation alone cannot convey to us the grand structure and near-infinite subtlety of society which we know them to have captured (McCormick et al 1987). What if it were possible to reconstruct part of what it actually means to live in Britain today — from the way in which twenty million census forms were completed — through pictures? This dream is not new:

From some of the first computer drawn maps of the Swedish 1980 census came the idea that there could be new ways of bringing masses of statistics to life through pictures:

> "From the diagrams which display households suddenly appeared a throng of people who with muted voices told of their lives, of their loneliness, of their joy in their children and of their hopes on their behalf."
>
> Szegö 1984 p.20

It needs to be made clear that the subject of this chapter is not illustration (to illustrate meaning, in dictionary terms, to clarify, make clear, pure or transparent). Graphic illustration of the census merely involves taking a few numbers and turning them into a pretty picture which is more interesting for publication than a table (for an example, see **Figure 1**). Simple graphs and charts, maps and diagrams are included in this category as they can contain only a few variables, even if they are of the most elaborate design. This is not visualization, which is defined in practical terms as the process of making visible what was hidden in huge tables of figures, often too large to be printed on paper (Tufte 1983). Visualization involves the creation of new solutions for a dilemma of our times — having too much information to understand when using conventional analytical techniques.

Social scientists in particular have a more serious and ironic problem in trying to understand census data using conventional statistics. They already *know* too much about the subjects they are studying for the simplistic findings from these studies to be enlightening. Statistics has been defined as "the study of the unknown" (Tanur et al 1985), being able to find structure in complex variable patterns — such as the influences of a variety of pesticides on crops or the discovery of clusters between alloys of many metals. But when it comes to the study of that with which we are most familiar — our social world — these techniques can often produce embarrassingly naive results, 'discovering' what was either already well-known, or results which are artefacts of the technique. The application of visualization is thus appropriate in this area because it can simultaneously show us both what we expect to see and allow us slowly to re-shape our perceptions of reality. Superficially the working of images is explicit (Marr 1982) — the literal opposite of *black box* techniques where the computer produces, say, a cluster analysis purporting to represent social inequality and you are implored by the weight of convention to accept its validity.

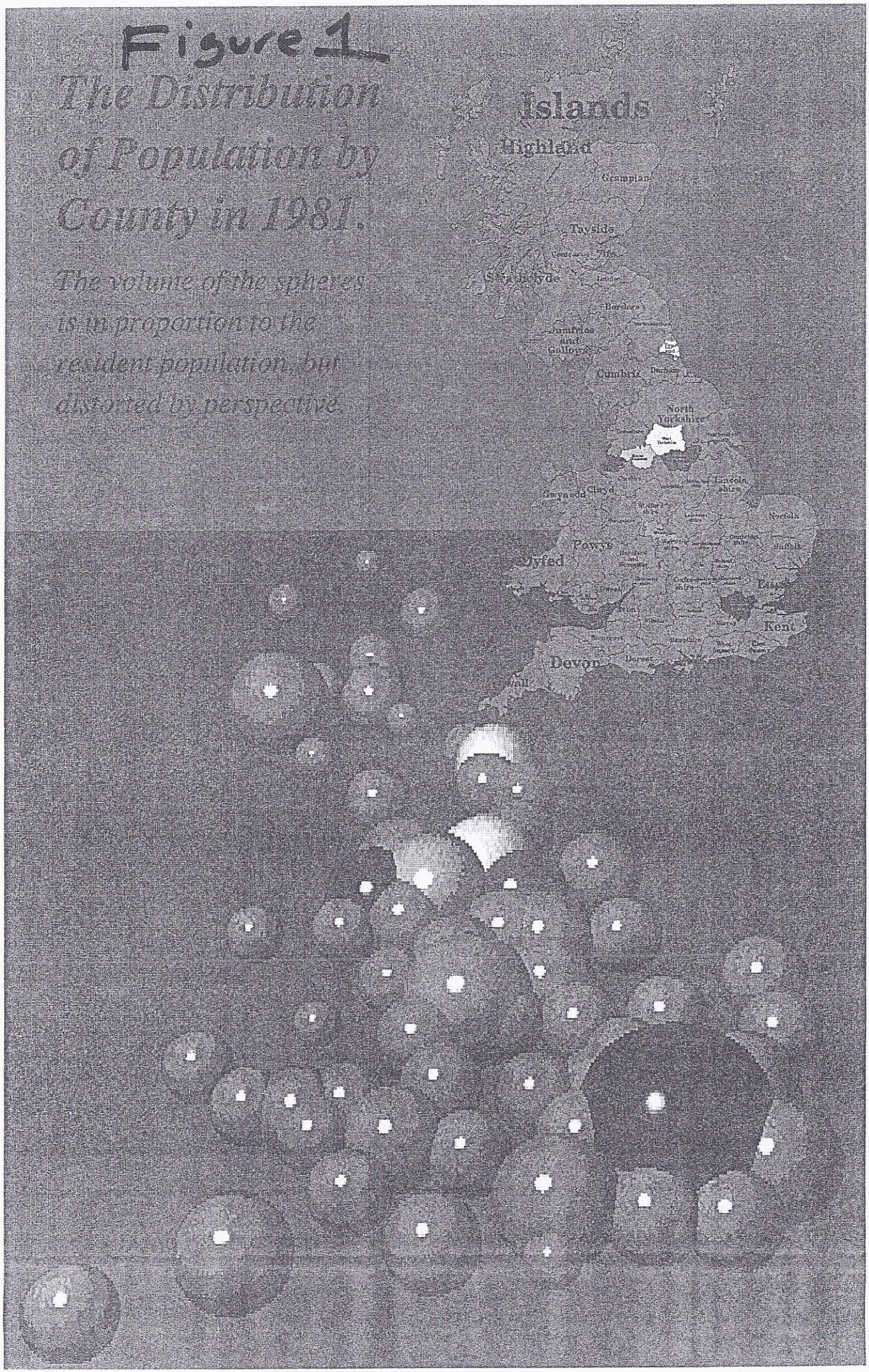
The pictures created through visualization may not show a pattern. This may be either because there is no pattern to find or because the information has not been looked at in the right way. Imagination in design is needed, but also a great deal of trial and error, in order to find an effective way of uncovering unknown patterns in census statistics (Tukey 1965). A scientific approach is needed to make sure that the pictures illustrate what was intended without unintended bias. An artistic approach is needed to ensure that the overall creation is acceptable to the mind and eye which are being asked to analyse them (Robinson 1989). Visualization requires the combination of disparate skills, and the existence of some structure to find, for it to be seen as successful.

The eye reads pictures, the mind analyses them: people respond and the public can react. To have presented, before our eyes, images detailing the nature of the society in which we live forces us to question more emotively the reasons for society's current evolution; by thinking about the quality of people's lives and of the lives of others around us. The census is our

Figure 1

The Distribution of Population by County in 1981.

The volume of the spheres is in proportion to the resident population, but distorted by perspective.



decennial snapshot of British society (Rees 1992). The challenge is to show, as best we now can, how the picture looks this time round.

The census is amenable to visualization because of the strengths of the inherent structures to be found in society and the quality of the geographical referencing which census tables contain. What makes the census particularly valuable for social science is not the wealth or depth of the questions that it asks (for they are few and shallow), but rather the great spatial detail that is provided — showing how each neighbourhood, each block of streets, each hamlet, differs socially from its neighbours (for every place in the country simultaneously). The social structure of our society is manifest most directly through the spatial distribution of its people (Pred 1986) and so through mapping the census we can obtain a glimpse of the complex social landscape which we imagine to exist (Census Research Unit 1980).

In the first Census User's Handbook there is a chapter by the editor on *mapping census data* (Rhind 1983). The advances reported in that chapter had altered census cartography significantly, and took many previously manual methods to their automated limits. Computer cartography had only recently become widespread, so for the first time maps of the census were not constrained by what it was possible to produce by hand (SEE GIS CHAPTER). Cartographers began to experiment with new forms of mapping, but forms which still largely thought in terms of traditional land maps — constrained by the tenets of a discipline rooted in mapping for military purposes, where the size and value of territory, rather than the lives of people, were all-important (Harley 1989). It is intriguing that census cartography combines two ancient tools of war: to put it simply, cartography maps land to be fought over (Sternberg, 1993), and census statistics count men to fight! It is important not to forget the original reasons why all this information was collected: that the state paid for it to happen.

Now that we can shed some of the constraints of traditional mapping (both in theory and practice) and adopt some of the new possibilities of computer graphics, many researchers are moving on to visualization (MacEachren 1992). It is worth reiterating that visualization, in a theoretical sense, means to make visible what is hidden; in this case, what is hidden in the reams of statistics now issued about Britain's people, households and dwellings. We want to be able to see what is happening to all the people of this country, to those living in tower blocks as much as those living in rural estates. We do not want to map the use and nature of British land, but to portray the employment and nature of the people in Britain. Here we are not concerned with 'census mapping':

> "Census maps, and related data, are used for three purposes: to locate census areas in relation to other places; to display census data in an effective and parsimonious way; and, to carry out analytical operations such as calculating the population characteristics of the catchment population of a major store."
>
> Barr 1993 p.250-251

The Office of Population and Census Statistics has released over a billion statistics derived from the 1991 census. We could whittle these down to a few numbers: 54,156,067 people, an unemployment rate of 9.5%, a few measures of statistical deviation; a table of sixty "key statistics", or the results of a hypothesis test. But that would not do justice to understanding the

wealth of information that has been collected. The desire of visualization is to grasp as much complexity as possible, to see what all these figures have to show, not to simplify but to comprehend. As a result very different axioms to conventional mapping emerge:

> "Detail cumulates into large coherent structures ... *to clarify, add detail.*"
>
> Tufte 1990 p.37

Social Structure often has a fine texture to it in which the microcosm reflects the macrocosm, like the patterns of fractal mathematics now so easily revealed by the computer (Mandlebrot 1983). There is so much detail that it would be almost impossible to draw a single picture by hand which fairly reflected the spatial composition of our society. The best method for visualizing a set of data will depend on the patterns that exist in that data which, in turn, cannot often be found without visualization. So a recursive search for a successful image is often needed, and in this a multitude of different techniques can be used. This chapter demonstrates the use of a few of these for visualizing census data in the hope that this will inspire others to experiment further.

The human eye-brain system has evolved to search for, and analyse, spatial patterns (Marr 1982). At speed it can spot order, repetition, grouping and inconsistencies in a picture. These are the very things researchers are looking for in a society. If we can turn these numbers into pictures we may be able to use the census to understand society better than we have been able to before, to learn something which we did not know, rather than merely to confirm our expectations.

2 Census Geography

> "The practices through which social structure is both expressed and reproduced cannot be divorced from the structuring of space and the use of spatial structures."
>
> Pred A. 1986 p. 198

Visualization of census data does not have to be spatial — nevertheless it is that aspect which is concentrated on in this chapter. The reason for this is that the greatest advantage of census information over other social surveys is that it provides us with, in effect, thousands of identical simple surveys which were taken at the same time in every locality. Most other surveys are restricted in only providing information for a single place, or for a sparse national sample, or for only a few regions or districts of the country — even though they might provide much more detailed data about their more limited subjects.

The finest spatial level of output of the 1991 census in England and Wales is the enumeration district. Each enumeration district contained, on average, four hundred people and consisted of just a few streets or a single block of flats in towns and cities, part of a village or a hamlet in the country. There were over 131,000 enumeration districts in 1991 in England and Wales (SEE GEOGRAPHY CHAPTER for how Scotland differs). We have been given information

about very many small groups of people and all their neighbours. It is therefore the spatial aspect of this data which gives it is value.

For each enumeration district a national grid reference to an accuracy of one hundred metres is given in what are called the *small area statistics*. This is supposed to identify the centre of population of the area. The points were chosen by hand and in the first release of the statistics many hundreds were found to have been misplaced! However, for the first time, digital boundaries of the enumeration districts have been produced, in theory delimiting precisely where 'on the ground' people were counted (these are due for release during 1993). Enumeration districts are likely to range in size from less than one thousandth of a square kilometer (the base of a high rise block) to over one hundred square kilometers (encompassing empty moorland in the most remote areas). It would not be surprising to find the largest being more than a million times the area of the smallest — whilst still containing fewer people.

How might we go about visualizing the characteristics of people in so many areas, which vary so much in size? **Figure 2** shows a grid drawn over a "normal" equal land area map of the British Isles. The grid is not uniform, but is violently stretched and twisted. It is made up of almost two thousand squares which each (on the mainland) contains 30,000 people. There are many such squares in the large cities, but few in the more rural areas. The picture is more complex than this, however; for in East Anglia, for example, the effect of Norwich can be seen, like a weight pulling the fabric of some giant net inwards. Each square contains the same number of people, yet some cover great swathes of land while others are barely visible. Perhaps it would be more sensible to pull all the lines straight — so as to form a rectangular grid — ensuring that each square would be of equal size so that their populations are more fairly represented than on a traditional map. An *equal population cartogram* would have been created (Tobler 1973a), but more on that later.

In 1981 digital boundaries were available only for wards — of which there were over 9,200 in England and Wales, and into which enumeration districts were nested. The electoral wards varied in population from 498 (Lower Swaledale) to 41,502 (Birmingham, Weoley), and in land area from 19 hectares (Skipton Central) to 44,789 hectares (Upper North Tyne). Thus they exhibit a skewed distribution in which the majority are small in area where people huddle together on the land. On a conventional national grid projection only a small proportion of the wards are visible. **Figure 3** shows these wards' boundaries for the whole country, with an inset focusing in on London. At almost every level of magnification there are some wards which cannot be seen and others (usually containing the least people) which dominate the image. Simply because the census provides geographical information using only one projection does not mean that that projection must be maintained (although almost all *mapping* does this). Similarly, we do not have to use the raw counts of the people it lists, but can transform these to use more meaningful scales.

What satellite imagery has shown us most vividly is that people do not appear on photographs of the earth's surface — they are not like fields, rivers or mountains. People are tiny, houses are tiny: something that may appear obvious, but which is all too quickly forgotten in conventional census cartography. More importantly, people are very unevenly distributed over most of the land surface of Britain, so that if a traditional equal area map projection is chosen for anything

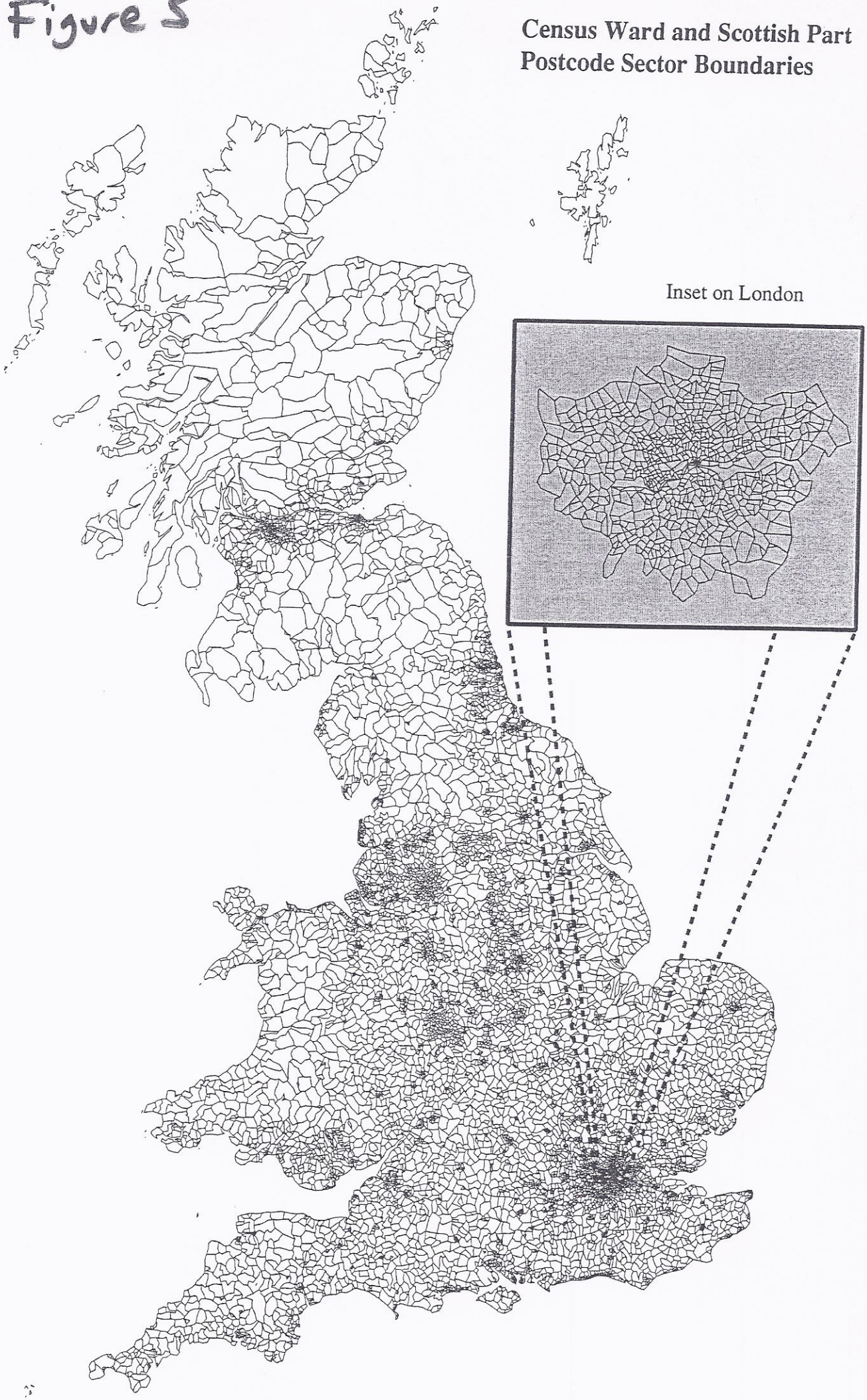
Figure 2

1981 Equal Population Grid Squares

From Smoothed Enumeration District Cartogram 30,000 people squares

Figure 3

Census Ward and Scottish Part Postcode Sector Boundaries



larger than a small part of a city, gross over- and under-representations of different social groups will be presented in the image. It should never be forgotten that:

> "mapping can be extremely informative but it can also mislead"
>
> Rhind, 1983, p.171

The information in the census concerns people and households. In visualizing these, a primary aim should be that each person and each household is given equal representation in the image. Current computer versatility allows this, so a principal challenge of visualizing the 1991 census is to achieve this. In the User's Handbook of the 1981 census (figure 6.5) a handmade cartogram of nineteen districts was presented. **Figure 4** shows an equal population cartogram of the more than ten thousand wards of Britain based on the 1981 census. In effect, the lines in Figure 2 have been straightened (for more details see Dorling 1991). Theissen polygons have been used to make the projection appear continuous and to create the 'stained glass window' effect as the populated areas meet the sea. Clipping a coastal boundary would result in a neater if less interesting image (SEE GIS CHAPTER for an explanation of these terms). The cartogram represents one level of abstraction beyond the land (choropleth) map — one level closer to the maps of social landscape we are aiming to see.

On the type of cartogram from which Figure 4 was derived, every ward is represented by a circle. Each circle's size is proportional to the population of that ward. Traditionally these cartograms have been used in mapping election results, and Figure 10 (described in the section on visualizing wards) illustrates this. In this approach, most wards on the cartogram have been placed adjacent to wards that they share common boundaries with, although occasionally this has not proved possible (thus it is a 'noncontinuous area cartogram' — in the terminology of Olson 1976). Even when reproduced to the page size of this book each ward is made visible in a single image. Here is a projection upon which it is possible to show the fortunes of every group of people at something akin to the neighbourhood level, nationally, without a particular bias against those groups who happen to live in the most densely populated areas.

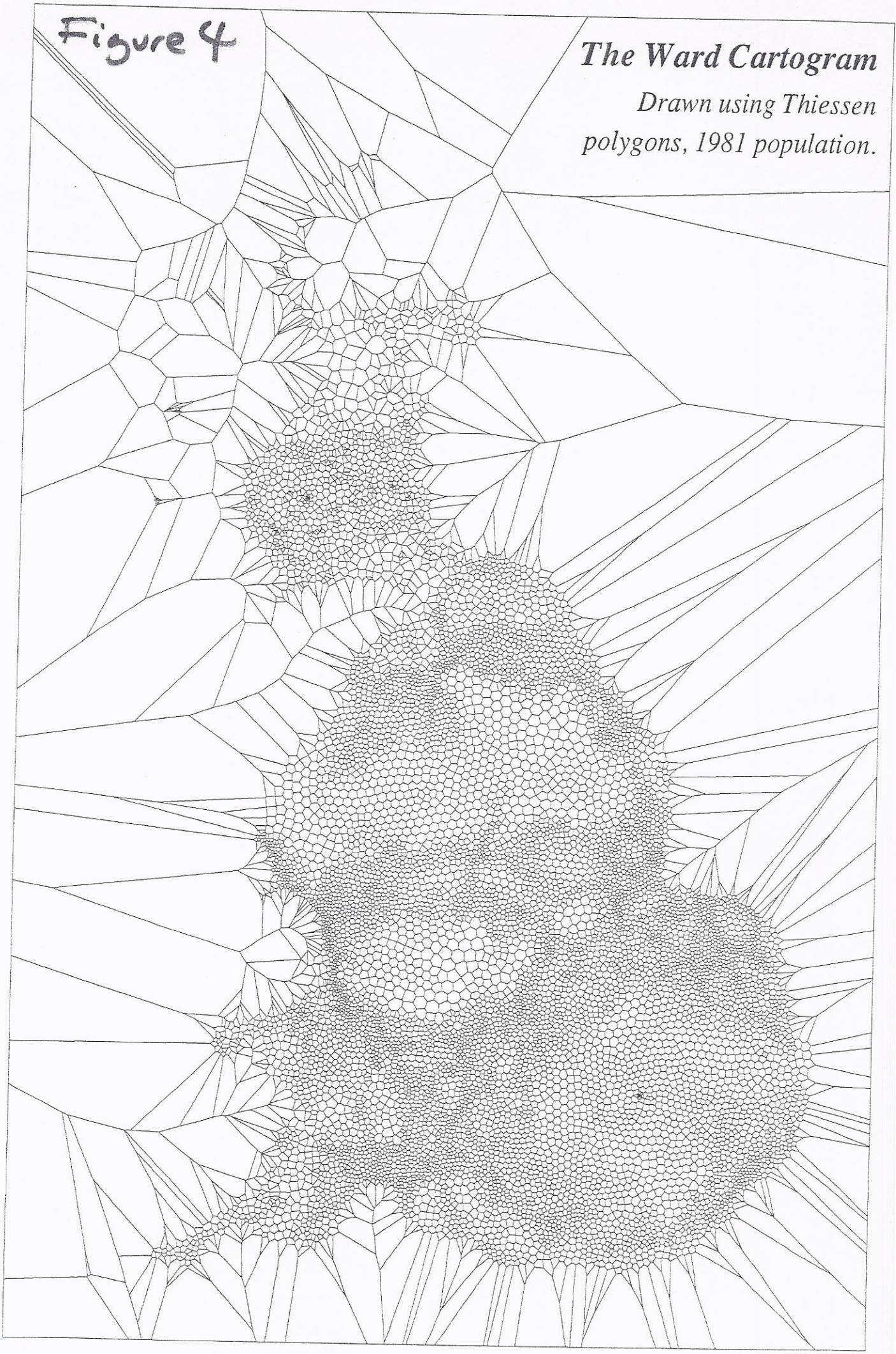
At what spatial level, however, should we be looking at society? Now that the creation of high resolution area cartograms is practical we are not restricted to using the lowest administrative level (Hunt 1968) or the smallest artificial areal units (Census Research Unit 1980) that are visible. We can use whatever level we choose. There is, however a wealth of good advice to be followed in making the choice of scale:

> "Even given these resources, it is essential, in general terms, to use the most detailed data available for mapping ... "
>
> Rhind 1983 p.183

We can even begin to go below the enumeration district level and redistribute census statistics in a dasymetric fashion using the locations of unit postcodes — if we so choose (see for example Figure 6, which is explained later — dasymetric mapping meaning only to shade the pieces of land upon which people live).

Nevertheless, the choice of how many places to adopt and show on paper does affect the

Figure 4



The Ward Cartogram

Drawn using Thiessen polygons, 1981 population.

techniques we can use for the visualization of census data. It also affects the number of independent variables we can investigate simultaneously, and the degree of detail with which we can investigate each one. Basically, the more areas we wish to see at once, the less we can see about each — there is a finite (if large) amount of detail that a single image can contain — and remain comprehensible. To decide how much *spatial* detail to see, necessitates deciding how much *social* detail to omit. To do that, we must first know how much detail there is in the census variables.

3 Portraying Variables

The 1991 census asked just over only a dozen basic questions; however each question has a number of possible answers (SEE VARIABLES CHAPTER). For example, "where were you born?" is divided into over one hundred possible replies (nation-states). "Where do you work?" is categorised into possible flows to over ten thousand wards, and is then cross-classified by "what do you do?". To begin it is often simplest to narrow down to a single yes/no answer, which can be converted into a proportion when reported for the group of people living in an area. This proportion can be represented in an image using a shade of colour. Enhancement using different colours, like the colours for different heights on a traditional physical geography map is possible, but runs the risk of creating patterns that are artefacts of where the colour boundaries lie.

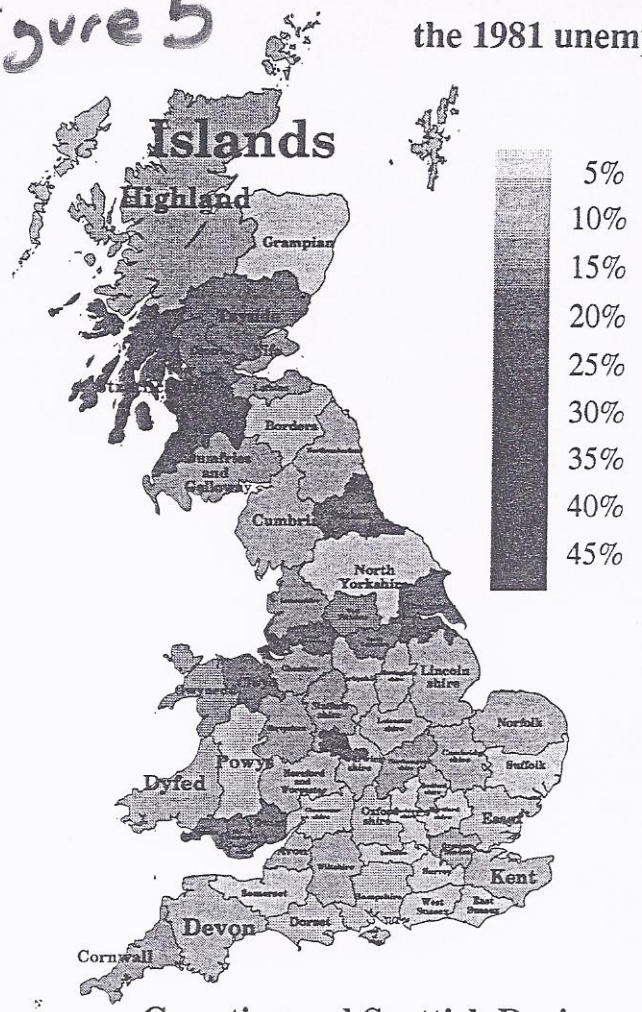
A multitude of statistical methods (SEE UNIVARIATE STATS CHAPTER) can be used to match variable proportions to shades of colour. The areas can be ranked by their proportions and grouped by above or below average, quartiles, or even percentiles. However there is a lot to be said for portraying detailed information in as direct and simple a form as possible (see Tobler 1973b). The images will often be difficult enough to interpret without having to bear in mind that the figures have been manipulated in a complex way before being converted to shades of colour. Continuous shading of areas is in the spirit of visualization because the aim is not to produce maps from which individual values can be accurately read (a table or database is better for this), but to produce images in which the overall distribution is accurately portrayed.

Even if we do not impose arbitrary categorical boundaries for shading (according to the variable depicted) we will still be using arbitrary spatial boundaries for calculating the value of the area in the first place. **Figure 5** illustrates what is called the Modifiable Areal Unit Problem (Openshaw 1982). In the figure, four conventional choropleth maps of the same unemployment measure from the census are shown shaded according to a single continuous scale. All that differs between the maps are the areal units — the boundaries used — to calculate and portray the variable. The figure shows which choice of boundaries has the greatest influence for this distribution, and where they make little impact on the impression gained. For instance, the high unemployment rate of the town of Corby is quite resilient to the different ways in which that place can be delimited. The problem is that, in general, this effect will be different for each distribution.

The Modifiable Areal Unit Problem can be handled in a number of ways. The simplest is merely to illustrate it by using multiple boundaries, and with interactive visualization it is

Figure 5

Annotated choropleth maps, shaded by the 1981 unemployment rate



Counties and Scottish Regions

Local Government Districts

Parliamentary Constituencies

Amalgamated Unemployment Office Areas

possible to redraw images instantly using different boundaries, to see the effects of these choices. A more sophisticated approach to this problem is to think more carefully about what areas are meaningful to visualize. A set could be specifically defined (eg Housing Market Areas), or 'fuzzy boundaries' could be used — for instance by employing kernel mapping methods (see Brunsdon 1991 for an explanation).

Arbitrary boundaries, however, have their greatest influence on the impression gained in their use for portraying statistics, not in their use for calculating them. If the whole of a county is shaded dark grey because levels of unemployment are particularly high in one of its towns, is our image accurately reflecting reality? Just because proportions are calculated for one area does not mean they have to be shown using a scaled down replica of the boundary of that area (ie by choropleth mapping). David Martin (1991), for instance, makes a strong case for virtually never using traditional choropleth mapping.

Dasymetric mapping, where only the points where people actually live are shaded, has long been advocated to produce more realistic pictures (Figure 20 shows an example from the 1851 census). The most detailed digital data source available to the academic community to achieve this "automatically" is the Postcode Address File which gives a ten or one hundred metre grid reference for each of over 1.6 million postcodes in Britain. Each postcode has been linked to a 1991 enumeration district (from the addresses on the census form) so it is possible to map dasymetrically below the finest level of enumeration. This method could be enhanced further by introducing a degree of chance into the graphics by "spraying" colour slightly at random around each postcode to reflect the fact that dwellings are spread out around these points, and that we can not to sure of their reliability (Barr 1993 p.251)!

But what do we see when we visualize using these methods? **Figure 6** takes Figure 3 a stage further, zooming into central London by postcode. Unfortunately, as we more accurately portray physical reality we are left with less and less space in which to show social reality. Even in the most densely populated part of Britain, what dasymetric mapping shows most clearly are those areas which it leaves colourless. The parks, rivers and roads of the capital are what stand out clearly on a dasymetric map of any variable for a city like London, and then most prominent are the isolated places, which usually contain least people and for which averages and proportions are least meaningful.

The severity of the Modifiable Areal Unit Problem for conventional thematic mapping is due to the fact that greatest emphasis is given to those places containing fewest people — where the arbitrary movement of boundaries can have the most severe effect on the values calculated. If the same set of unemployment statistics as used in Figure 5 are mapped on population cartograms as in **Figure 7**, the arbitrary boundaries can be seen to have much less influence on the impression gained. Our methods of visualization must be robust if one arbitrary choice or another is not to lead us to misinterpret what is important.

Already the number of things to be kept in mind has multiplied and we have so far only considered visualizing a single simple proportion over a few areas. Researchers soon tire of this simplicity and wish to look at several variables to see how they are related. It is obviously possible to produce several univariate pictures side by side for comparison, but it rapidly

Figure 6

The Distribution of Households by Postcode in 1991

Inset of 25km² from the heart of London. The area of each circle is in proportion to the number of households at that location.

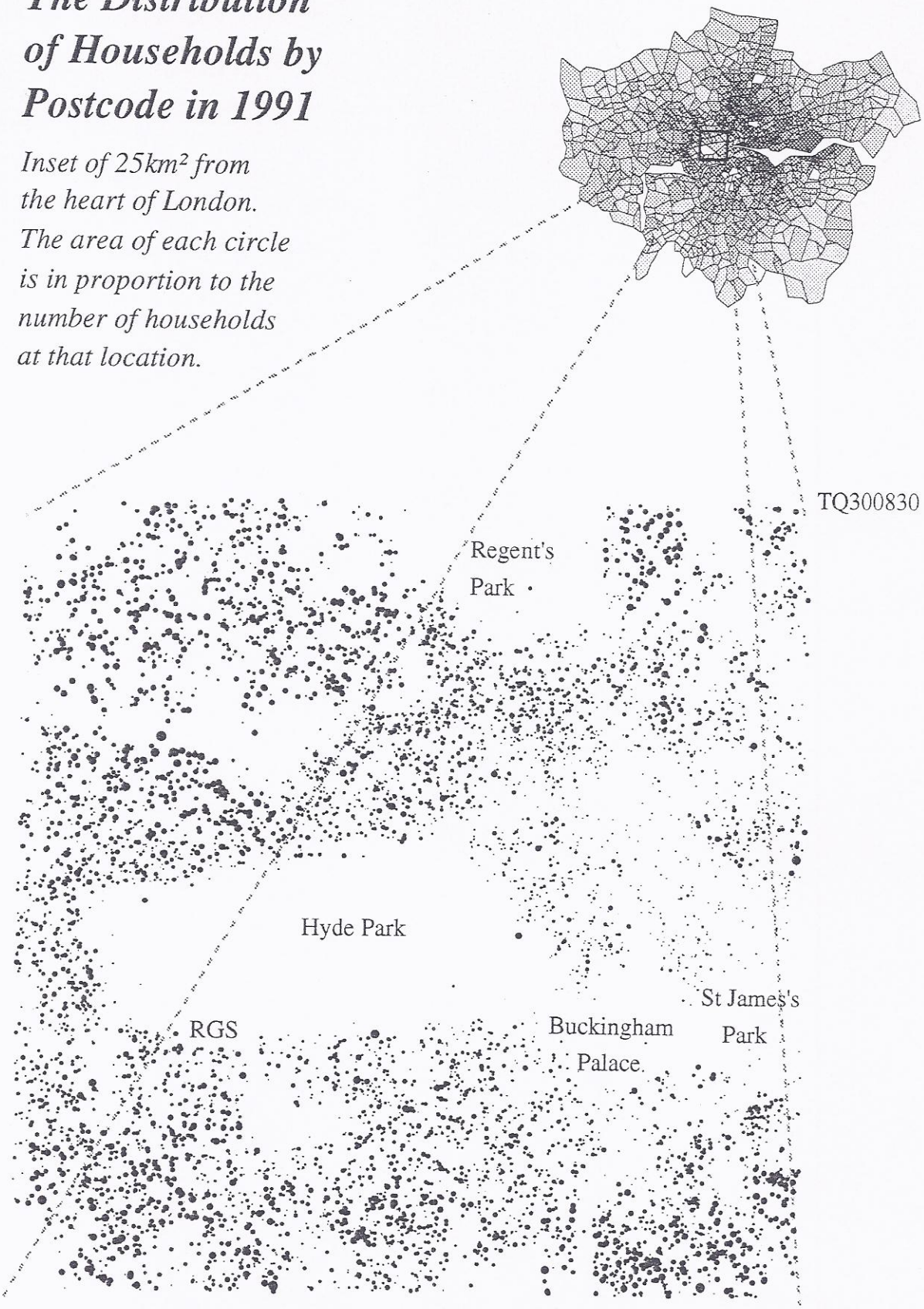
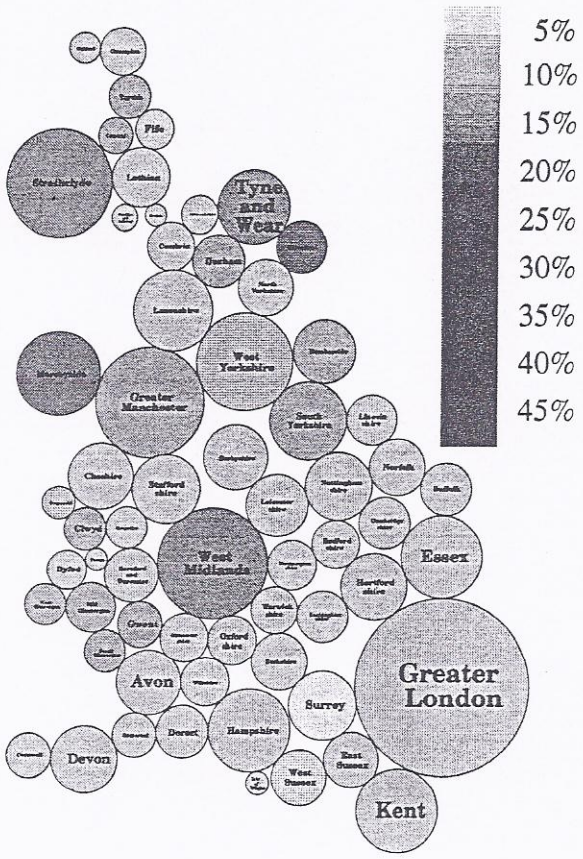
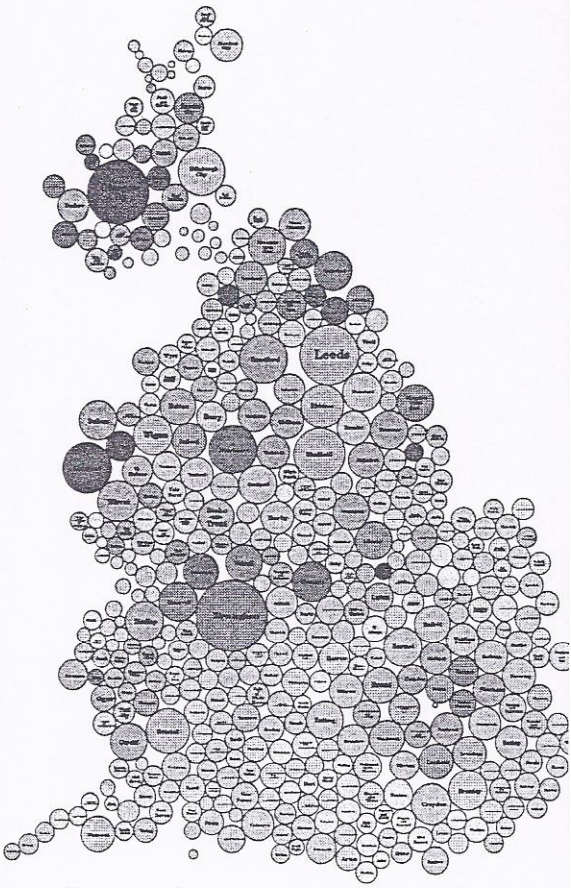


Figure 7

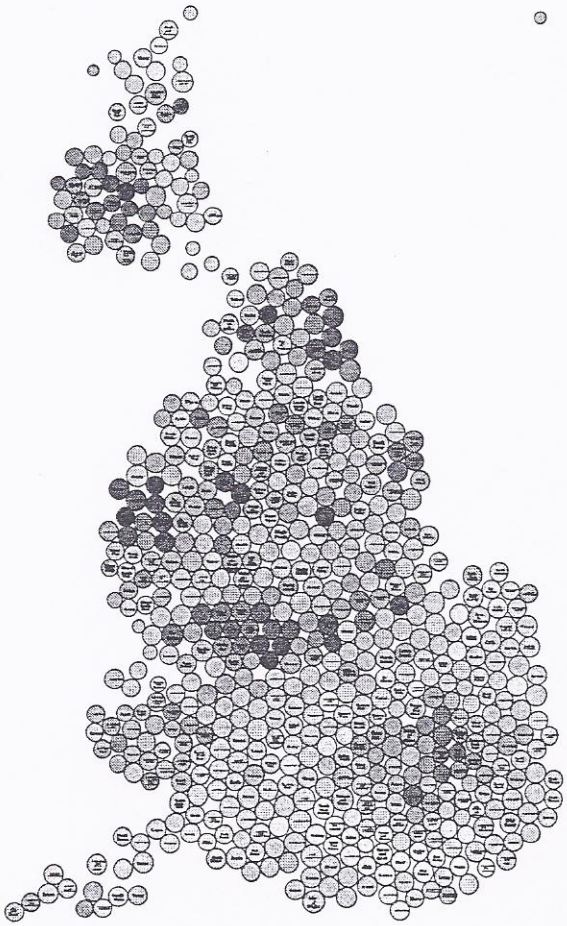
Annotated equal population cartograms, shaded by the 1981 unemployment rate



Counties and Scottish Regions



Local Government Districts



Parliamentary Constituencies

Amalgamated Unemployment Office Areas

becomes apparent that only the simplest of cross-references can be made between them (for example: 'are the same areas shaded lightest?'). What is required is the inclusion of several variables in the same image. Methods of doing this are discussed later; what is important here is to stress the basic decisions which need to be made about the variables before the drawing can begin. It must also be realised that as we increase the complexity of our design, so too we decrease its ability to be more generally understood. At one extreme we are producing pictures for our exclusive use and need not worry as to whether they are meaningful for others — at least not until the results have to be demonstrated.

Variables can be mutually exclusive; for instance it is not possible to have been born in two different places, or to be both in and out of work. Variables can also be called exhaustive when they cover all eventualities: for instance, "born in Britain" or "born outside Britain". Visualizing two mutually exclusive and exhaustive variables only requires the portrayal of one independent variable (as the second can be derived from the first) and is thus a much simpler problem than simultaneously visualizing two independent variables such as the proportions of people born outside Britain and the proportions unemployed (in each of thousands of places).

The visualization of independent variables also raises the risk of invoking the "Ecological Fallacy" (Openshaw 1984). The fact that in a particular area there are significant proportions of two groups of people should not imply that there is any evidence that the two groups include the same people (at least not until the sum of their proportions rises above one hundred per cent). Often there are cross-tabulations in the census statistics which show apparent simple correlations to be false. A classic example is that areas that tend to have high proportions of people born outside Britain often also tend to have the lowest average levels of qualifications, however it is usually not the 'lifetime migrants' who have the lowest number of degrees or HNDs. But, how best to visualize the pertinent features of a table of statistics such as these, for and in, each area?

When faced with such complex problems, the temptation is to reduce the variables down to categorical values. For instance, when investigating the spatial relationship between unemployment and, say, twelve social groups, each social group could be assigned a colour, and the social group with the highest or lowest unemployment rate in each area would determine that area's colour. A dozen variables, each with a practically continuous range, will have been reduced to a single variable which can only take twelve values. It is important to be aware of how much information is being thrown away by such a process (which is also heavily dependent on the arbitrary boundaries drawn between social groups: 'who is a manual worker and why should we use that category or any other?').

Let us suppose, however, that a single image is required to display the relationships between social group and tenure, where we have a matrix of perhaps twelve variables by seven for each area. One possibility would be to subtract the national matrix of proportions from each local matrix and choose to categorise that area by the greatest deviation shown — e.g. to show 'an unusually low number of professionals buying their homes'. Whether it is possible to do this depends on the structure in the data and the number of categories which emerge. Often we have to admit defeat and use several images (or amalgamate many of the variable's distinctions). We must also take account of the inherent variability of such statistics (particularly if taken from

the 10% census), so as not to identify an insignificant change as important (Cole 1993).

Finally it must be pointed out that the census does not only contain information about people in areas, but also information about people moving between areas. Flows of people migrating and commuting are some of the interesting datasets which have been derived, and the most challenging and rewarding to visualize; challenging, because instead of a single set of areas, we now have a matrix of areas to represent, ie the flows between each area and all the other areas — squaring the number of "objects" which might need to be rendered visible. To do so, however, would be rewarding, because there is so much to see in flows which are the most spatial, of spatial data (unemployment can by studied without looking at its geography but not, by definition, migration).

Before we look at how matrices of variables and flows can be visualized, it is necessary, in an attempt to be systematic, to look at the visualization of the most simple of variables — proportions — over the greatest number of places possible (without dasymetric mapping) — this is, all enumeration districts.

4 Visualizing Enumeration Districts

> "...within any town as in the region as a whole there is a pattern. The poor housing, schools and levels of unemployment will tend to be concentrated in certain districts — as they are concentrated in inner city areas of large conurbations of this country. At the level of the region, too, there is a pattern, increasingly clear and changing."
>
> SEEDS 1987 p.6

Enumeration districts represent the finest level at which census data are released in England and Wales (in Scotland the situation is slightly different — SEE CHAPTER ON CENSUS GEOGRAPHY). As has been detailed in section 2 above, there are over one hundred thousand enumeration districts in Great Britain, with many thousands of statistics being reported independently for each from the 1991 *small area statistics*. How can that national pattern across this finest level of resolution be seen?

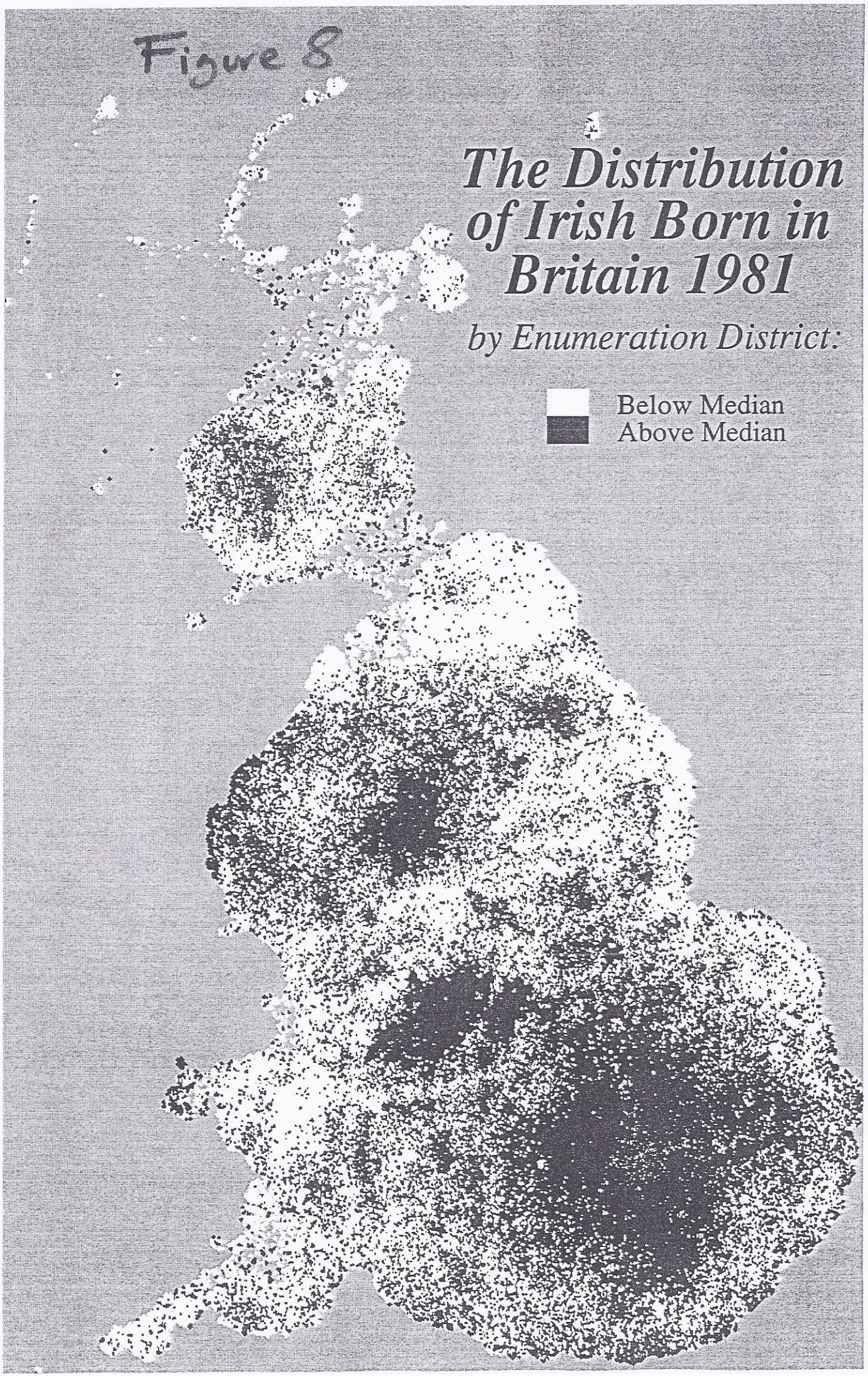
An A4 sheet of paper (or an average computer monitor) is roughly 50,000 square millimetres in area, excluding a little space for margins and borders. If each person were to be given equal area then the scale would be just over 1000 people per square millimetre (irrespective of the areal units used). Three average enumeration districts would then have to be squeezed into each square millimetre. At this size they can only be shown by a spec of colour (although still larger than the dot on the top of this letter 'i'). What would an image, which purported to show the characteristics of so many communities, look like?

At the simplest level each enumeration district dot is shaded either pure black or pure white. **Figure 8** shows the distribution of Irish-born people living in Britain by enumeration district in 1981. An enumeration district was shaded black if it had more than the median average proportion of Irish-born residents. Thus half the country (on the cartogram) is black and the

Figure 8

The Distribution of Irish Born in Britain 1981

by Enumeration District:



other half white. The major metropolitan centres appear as dark clusters, while a more speckled appearance is given to those areas where the concentrations are less clear cut. Although only black and white are used here, at this scale the dots can merge to create shades of grey at the fringes.

Many more elaborate shading schemes than this can be envisaged. For instance, only those enumeration districts which held the most highly concentrated half of the Irish-born population could be coloured. Alternatively, a series of grey-scales could be used; the darkest for the most highly concentrated ten percent, down to the lightest for the enumeration districts with the lowest proportions of Irish-born. Svegö (1984) has suggested that the border (in the case of Great Britain, the sea) be coloured a neutral grey rather than the usual white. It might even be more appropriate to use a colour not on the grey-scale. The fact that it can be claimed that such subtle changes can affect our impression of an image is not a weakness of visualization but a strength, as we learn that there are many ways of seeing the same pattern, each literally in a slightly different light.

The mixing of shades to create textures, too, can be seen as an advantage and can even be enhanced by smoothing the image to achieve some desired level of *generalization*. **Figure 9** shows the result of using ten passes of the most basic binomial filter (through which, in each dimension, each point is given a new value equal to half its old value and a quarter of its two closest neighbours — Tobler 1989). In the figure the very complex distribution of occupational groups has been smoothed, where three-colour mapping (see below) is being used to the illustrate subtle trivariate interactions of people by class. Generalization very much simplifies the image and is particularly useful when reductions in size are required, but; as with all these decisions, it changes the impression of the distribution. Watching the generalization take place as an animation can be very enlightening — but that subject too is reserved for later discussion.

While black, white and grey-scales are most appropriate when looking at a single distribution (as we can best distinguish shades without colour), if the wish is to study more than one variable at once we must consider the use of colour. On almost any microcomputer screen colour is now taken for granted, but high resolution, high quality printing and publishing can cause difficulties (this chapter for instance was limited to four colour prints). Much can be seen with colour which could not be visualized in black and white.

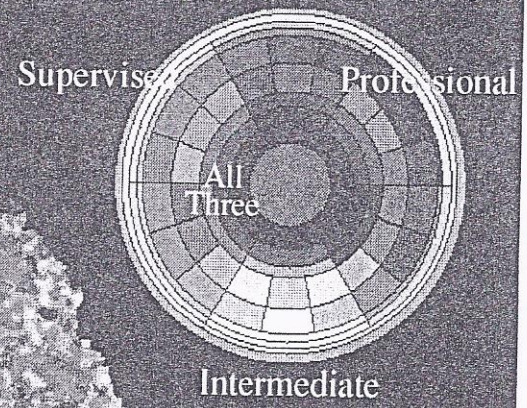
Colour can be used to make grey-scale images appear less drab; although this does not make the best use of the opportunities which colour offers, and may result in the picture being more poorly understood (as fewer levels of distinct hues can be distinguished than can shades of grey). For the presentation of categorical data, however, colour is almost a necessity, as distinctly different shades are what is required. The most challenging use of colour in visualization comes when an attempt is made to maximise both the advantages of the distinction between hues and the continuity of saturation (and/or lightness). This strategy is needed when we wish a single colour to represent the levels of three dependent — or two independent — variables, through the mixing of colours, as is explained below.

Two-colour, or bivariate, mapping now has quite a strong tradition in cartography. It can be done in many different ways but here just one possibility is outlined. Suppose enumeration

Figure 9

The Distribution of Occupation in Britain 1981

Quartile Levels by Enumeration District of:



After 10 Passes of Binomial Smoothing

districts were coloured according to the proportion of people born in Ireland — using red when this statistic was in its highest quartile and white when in its lowest, and with two shades of pink in between. Overlay this distribution with that of people born in Asia, coloured similarly but with shades ranging from blue to white. Where there were high concentrations of both groups the enumeration district would be coloured purple, white for neither, red for mainly Irish and blue for mainly Asian and with various shades in between. The basic geographical distribution of these two groups across the whole of Britain could be made visible at the highest level of resolution on one (admittedly garish!) page. A pure red dot (representing an enumeration district with the proportion of people born in Ireland in the top quartile) by a pure blue dot (proportion of Asian born in the top quartile) might look to the eye as purple, but would that be a disadvantage? After all, the dots were only so different because of the imposition of an arbitrary line drawn for the benefit of enumeration. If the line had been drawn another way both areas might have been coloured pure purple. Shading of this kind, at this resolution, can make up for some of the difficulties of the modifiable areal unit problem (see above) by producing images which are to an extent free from the arbitrary effects of the precise delimitation of enumeration districts, because of our eyes' natural ability to generalize the image (Arheim 1976).

The next obvious step after two-colour shading is three-colour, or trivariate, colouring. This is the most sophisticated possibility when there is only a pinprick to shade. Three primary hues of varying intensity will produce a unique colour for every possible combination (while more than three colours do not). When images use three-colour shading, the patterns formed are usually much more complex. It is, however, often the case that the more complex a picture is, the more there is that can be learnt from it. This is not necessarily easy; but can be very rewarding in terms of gaining new insights.

The example of generalization (Figure 9) shows clusters of enumeration districts nationally where the occupation of 'heads of households' is predominantly 'professional' (blue), 'intermediate' (yellow) or 'supervised' (red). Where there is a concentration of professional workers, for instance around Greater London, the area will appear bright blue. Where there are a high proportions of people in both intermediate and supervised occupations (but not in professional occupations), for instance around central Birmingham, areas of orange emerge. Superficially, there is not a great deal of difference between this kind of image and a CAT (Computer-Aided Tomography) scan, with the latter showing a slice through the human brain and the former, a slice through society. One important difference is that a great deal of training is needed to interpret a map of the brain, with which most of us are unfamiliar, whereas a map of society can be much more quickly understood (as we constitute it) and so can show more complex structures.

Because the three distributions of occupations spatially 'repel' each other (people in different classes of job tend not to live in the same streets), the primary colours chosen, of red, blue and yellow dominate this image. If the variables chosen had been birthplace (subdivided by, say, Irish, Asian and Afro-caribbean), then a great many places would be coloured either black (high proportions of all three) or white (very low proportions) because these groups tend to cluster together, although many more subtle mixes of colours would also be visible. In Figure 9 the sharpest divide is shown to be in inner London, where a white buffer appears between the

blue centre of the capital and its red core. This is where the spatial generalization procedure failed to merge the colours because social divides along this boundary are too great.

One final word of warning about colour. The apparent colour of an object is affected by the colours surrounding it. Isolated spots of unexpected colour stand out in otherwise uniform areas far more than they do where there is already great spatial variation. The colour of each object you see is, in fact, the result of a mixture of colours on the page that is unique to every distribution shown. This is not so grave a problem when visualizing geographic space as it is with other subjects. To live in the only (hypothetical) enumeration district in Northumberland, which has more than a quarter of its population born in Asia, is a very different experience to living in an enumeration district in London with a similar proportion of the population born in Asia (where this proportion is not at all unusual). We must just understand that our pictures can be as subjective as the conclusions we draw from them — and the decisions which led to their creation — despite the explicit consistency of their design. Each apparent weakness of a visualization technique may also be argued to be a strength, although the apparent strengths of such positivism can easily be turned about to list obvious weaknesses.

4 Visualizing Census Wards

> "Human beings make a strange fauna and flora. From a distance they appear negligible; close up they are apt to appear ugly and malicious. More than anything they need to be surrounded with sufficient space — space even more than time."
>
> Miller 1934 p.318

There are many ways in which enumeration districts could be amalgamated to form larger areas for study. We must always have a reason for discarding so much information in amalgamation. The reason for using wards is that they are the lowest level of official British administrative geography for which non-census information about our society is widely disseminated. They are also the smallest areas for which political officials are elected, and so have some purpose other than enumeration. There were over 9,500 census wards in England and Wales in 1991. Again the situation is somewhat different in Scotland and so the CHAPTER ON CENSUS GEOGRAPHY should be consulted for those interested in the nomenclature more generally. Here, purely for convenience, Scottish Part Postcode Sectors are considered alongside English and Welsh wards for the purposes of visualizing the census.

The 1991 census contains much more detailed information at the ward level in the *local base statistics*, allowing far finer subdivisions of population characteristics to be made than are available in the small area statistics. How though should we begin to investigate the distribution of people by, for example, nearly a hundred countries of birth, at the level of the ten thousand wards in the country? This can obviously not be adequately achieved with a single map: although if that were the goal (for instance to produce a ward based atlas of the whole census) how might that goal be approached? One option, which has been alluded to before, is to start using colours to represent categories, rather than representing variables.

With areas larger than pin pricks (see Figures 3 and 4) we can be more sure of recognising particular shades of colour. Thus, a more sophisticated statistic could be chosen than the usual averages or quartile ranges. Each ward could be categorised by the place of birth which most deviates from the usual proportion. So a ward with every single one of its residents born in England might be sufficiently unusual as to be labelled "English-born", whereas one with just ten residents born in Trinidad and Tobago might qualify, according to whatever statistic was devised, as being unusually "Trinidadian and Tobagan — born". Each continent (of birth) could then be given a primary or secondary colour, shades of which would be used for many of the individual countries. Thus a choropleth cartogram, like a physical geographer's soil or vegetation map, would be created highlighting those towns of greatest diversity of birthplace and those of most homogeneity.

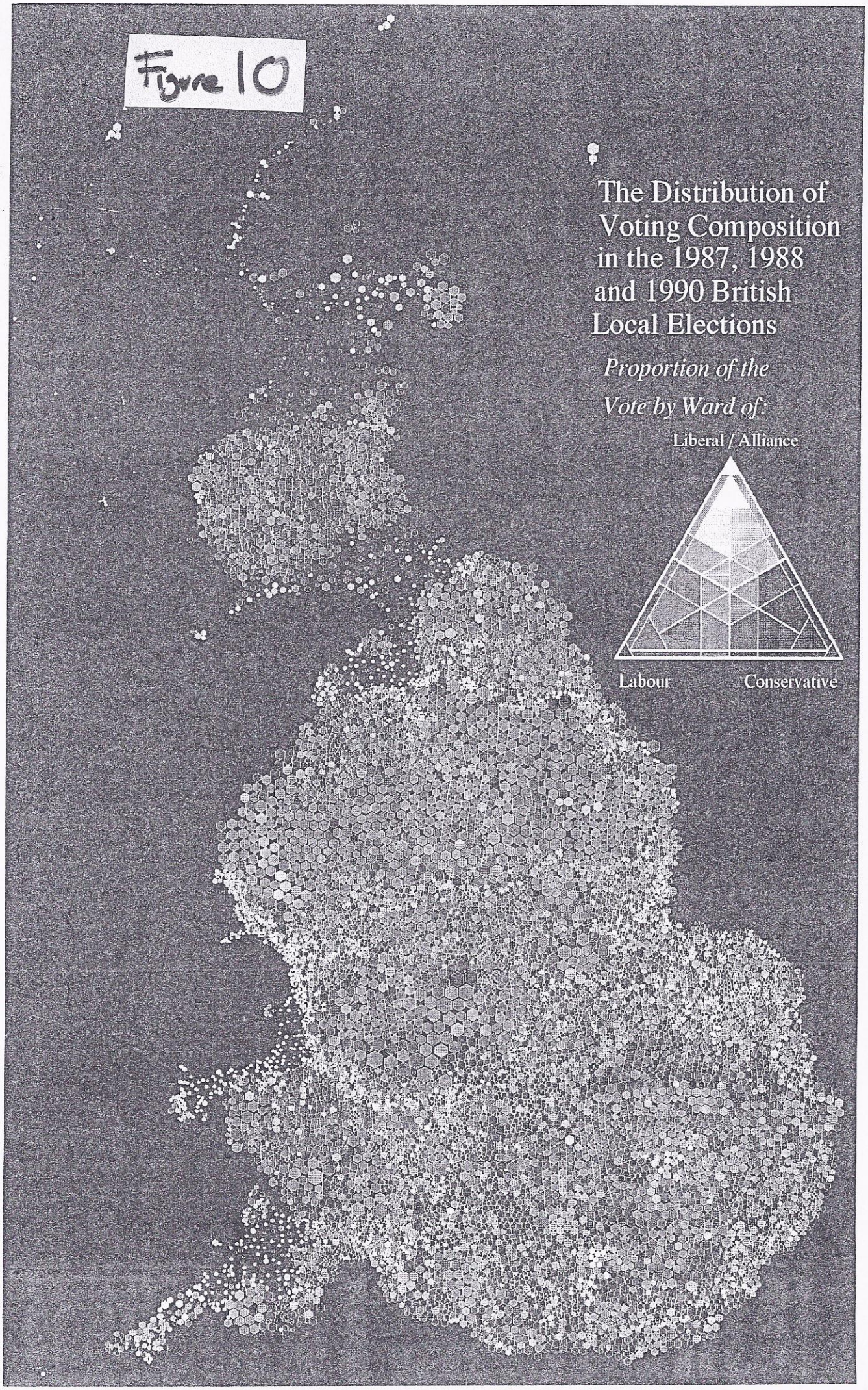
Some statistical sophistication may be needed to avoid just highlighting "quirks" in the figures. Estimates of sampling error may have to be made if, for instance, data from the 10% sample of census forms was being used — particularly when measuring the numerical change over time of quite small groups. It might be found that their proportions of the whole population would have to alter substantially for a change to be significant enough to be worth showing. The reliability of even the "100%" results may at times also need to be questioned (SEE CHAPTER ON CENSUS QUALITY). For instance, in mapping the change over time of the location of young men (the category which is known to be least well enumerated in the 1991 data). These problems are not, however, reasons for avoiding mapping at the ward level. The aggregation of data increases "accuracy" only in the most trivial of senses.

Figure 10 shows a cartogram based on the night time populations of all wards in Britain, used to visualize the results of the British local elections of 1987, 1988 and 1990 combined (all three years have to be included to get a complete coverage of the country). Each ward has been coloured according to 36 categories of election result, ranging from cyan for a Conservative marginal where Liberal came second and Labour third to blood red where a ward where only a Labour candidate stood. The standard electoral triangle (Upton 1991) is used as a key in which the colours of all the possible categories are displayed to indicate the proportions of the vote they represent. To explain adequately just this one diagram would take several pages of text. That is not the purpose here: what is of interest for this chapter is that this is possible.

The lace-like patterns of local voting tally very closely with those of occupation which were shown in Figure 9 (as has been known for a long time, but can now be demonstrated visually). Features such as the concentration of Liberal seats on the edge of Inner London, and the Liberal's general high propensity to appear where there is a greater degree of social mixing, might have been missed in a conventional statistical analysis (where the 'answers' have to be known to some extent before the 'questions' can be posed). The fact that everywhere there are small clusters of bright blue ('safe' Conservative seats), surrounded by purples (Conservative/Labour marginals) encroaching on reds with almost infinite repetition might well also be missed without visual analysis — as too can the stark message of left-wing support from most of the electorate in recent local elections.

Returning to the visualization debate, more flexibility of depiction could be introduced by altering the ward's shape, for example to arrows or teardrops (such as is discussed in the next

Figure 10


The Distribution of Voting Composition in the 1987, 1988 and 1990 British Local Elections
Proportion of the Vote by Ward of:
Liberal / Alliance
Labour
Conservative

section). Such techniques can be effective when there are marked regional variations, but can be confusing when the pattern is more diffuse. The effectiveness of using changes in the shape, size, orientation and colour of each ward's depiction depends on both the actual pattern of the variable portrayed and on how easy it is for the mind to relate that particular variable to each mode of representation. Can we easily *think* of that variable as *dark* or *light* (eg unemployment 'black spots'), as one particular colour (eg party affiliations as red, yellow and blue), up, down, left or right (commuting), wide or thin (range of available tenures perhaps), large or small (economic status)? In the ward cartogram shown earlier (Figure 4), size is used to represent total (night time) population, but could be used to show the daytime population after the effects of commuting have been included. Thus in urban areas the circles would overlap whilst rural areas would appear empty. The circles could then be coloured by, say, predominant "mode of transport to work", so showing how the transformation took place.

Much more elaborate graphics schemes than this could be devised, but with each ward having less than 5 square millimetres in which to *express itself*, even on a cartogram, it is best to restrict the options to (albeit more elaborate) shading and colour schemes. In the above example, the overlapping in city centres would be problematic, on the other hand if all wards were scaled so that even the largest didn't overlap, then very few would be visible. There is still more information to be shaded, however, because as opposed to most enumeration districts (for which there is no comparable data over time), we can amalgamate the enumeration districts of previous censuses to fit approximately into the ward boundaries and so calculate — and hence visually express — approximate changes over time (even when the ward boundaries have altered). It is then possible to ask: has the total population gone up or down, or how have the characteristics of the population changed? Almost inevitably, the image will have to depict the areas, as they existed, at one particular date. But how best to present this information?

If there were 200 people out of work in a ward of 2000 "economically active" adults in 1981, and 224 seeking work out of 2800 in 1991, then the number of unemployed has risen by twenty four while the percentage has fallen by two percentage points (from 10% to 8%) or by a fifth. Which change statistic we choose to show will often dramatically alter our final image, as will the way in which we choose to show them. The purpose for which we are visualizing the information has to be the prime motivation for which statistics we choose to portray. If the interest is in planning the provision of benefit offices then it may be the absolute change that is of most relevance. If it is in the state of the labour market, then the shift in proportions may be more useful; alternatively, if it is in the changing chance of individuals being unemployed, then the relative changes could be chosen.

If we are not just interested in the changes between two dates but between 1971, 1981 and 1991, we would probably draw separate pictures for changes between pairs of dates, but could then go on to compare the 1981-1991 change with the 1971-1981 change. Is the change accelerating or decelerating, reversing or continuing? However, to understand a picture of 'change in change' requires simultaneous understanding (and preferably visual depiction) of the absolute change, and of the original and current levels of the variable. Complex analysis must remain comprehendible. The rate of growth or decline of a newly immigrant population makes little sense without having an idea of the initial size of that population. With more than three time periods included, there may be no option but to use separate maps — because the temporal

patterns will tend to break up the spatial ones. More elaborate methods of depicting change are discussed in the next section.

Even greater challenges are posed by the portrayal of what in many cases actually causes the changes — flows of people. Flow is much more than change; flows describe how things change — for example, how the night time population becomes the daytime population (through travel-to-work, travel-to-school and so on). A greater degree of magnitude of information is contained by a matrix of flows than in a vector of change, and hence the difficulties (and benefits) of visualization are similarly magnified. The prevalence of, say, commuting flows to a single city centre could be depicted by shading other places so that the higher the proportion of its resident workforce commuting to that centre then the darker it is shaded. If we have over ten thousand places (as in the case of wards nationally) then ten thousand different maps could be drawn! The *Special Workplace Statistics* of the census do, in fact, give us this level of information, further broken down by many of the characteristics of those commuting. The challenge for visualization is to transform this huge matrix into condense visible form to see the structure and the disparities.

Luckily, there is inherent structure in travel-to-work flows which makes their spatial depiction, even at this level, not too problematic. Between the ten thousand wards in Britain there could, theoretically, be as many as one hundred million movement streams. Rather unsurprisingly, less than one percent of these actually occur in an average year. However, the flows are often heavily biased in one particular direction (towards towns in the morning when commuting), so much so that we do not need to show this. If we plot only those flows where more than 2% of the employed population of a ward work in another ward, our diagram will show the daily movements of just over half the workforce in space (at an almost individual level). **Figure 11** shows how these look on an equal area projection, and illustrates how much more is now possible than was even thought plausible ten years ago:

> "In short it is unlikely that more than thirty flows can be shown successfully on one map and planning the design of a flow map is singularly difficult — 'try it and see' is an essential element of such map creation."
>
> Rhind 1983 p. 176

Whether any number of flows can be visualized successfully depends on the inherent structure of those flows. Twenty million individual flows, converging within only a few hundred groups, are not difficult to understand; whereas ten thousand flows going in all directions would look much like any other set of ten thousand lines. With migration there is less spatial conformity to the movements because the subjects are no longer constrained by travel-time. But even changing the way we determine what is a significant flow to plot can dramatically alter the image created. **Figure 12** shows the effect of measuring the propensity of movement (by dividing the numbers of commuters by the workforce living in the wards of origin and destination). Drawn on a cartogram base for clarity, London dominates the image with its low central night time populations. Manchester creates a distinct cluster because it has a single "Central Business District" ward, whilst there are almost no flows to be seen in Birmingham using this measure (because it has a much more dispersed pattern of employment). Without seeing this picture, how can you know what effect a particular choice of statistics might have?

Figure 11

Daily Commuting Flows Between English and Welsh Wards in 1981.

Flows of more than 2% of the employed population at the area of residence are drawn as thin lines to their workplace.

50% of all commuters included - 10,319,230 people.

Figure 12

Daily Commuting Flows Between English and Welsh Wards in 1981.

All flows which satisfy the following inequality are drawn as thin lines-

$$\frac{m_{ijst}}{p_{is}p_{js}} > \frac{1}{25000}$$

Flows of over 1000 people drawn as thick lines.

Where-

m_{ijst} : The number of people moving from place i to j between times s and t.

p_{is} : The number of people at place i time t.

i: Place of residence.

j: Place of work.

s: Nighttime.

t: Daytime.

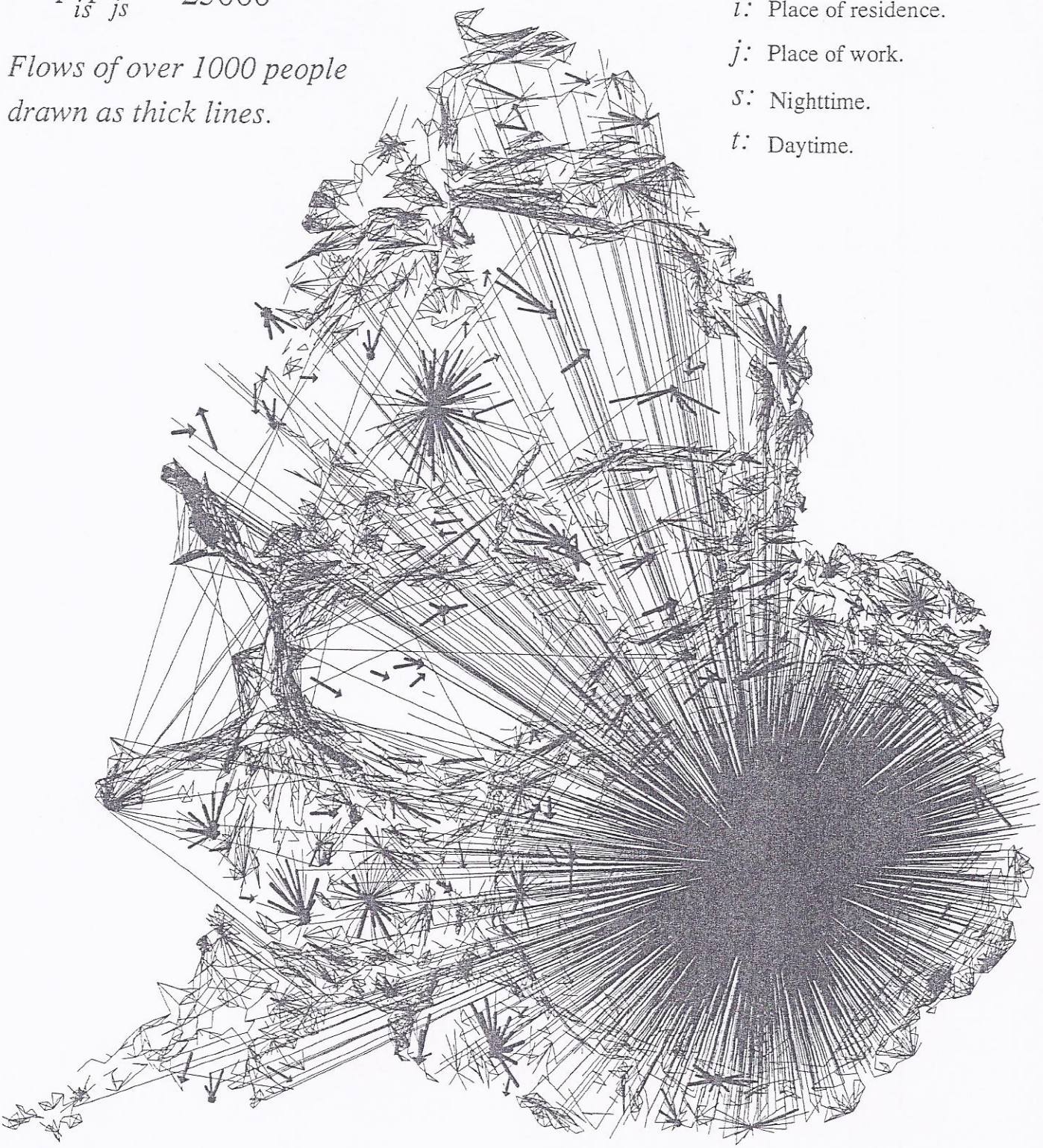


Figure 12 also highlights some of the serious difficulties in depicting flows between a large number of areas. The modifiable areal unit problem is rife in flow mapping because the size of an area influences the rate of flow (as smaller areas will have less flows within them, and hence a higher proportion of flows out of them — one further reason for Birmingham's absence in Figure 12), and the *shape* of the area can also have a drastic effect. Long thin wards are more likely to have higher rates of in and out flow than more compact ones (because, even if people were moving at random in a long thin ward, they would be closer to a boundary and therefore more likely to cross it and thus be registered as a 'flow'). The problem does not diminish when large areas are used. Between standard regions, most of the migration or commuting is due to flows of people who live near the boundaries of these regions and commute or migrate just a few miles across the border. Flows between non-contiguous areas may have more significance, so charting those alone could partly overcome the problem of the influence of each area's size and shape.

The 1991 census data will allow us to draw new maps of both commuting and migration flows. Suppose we were interested in seeing how the flow maps (at this level of detail) had changed: how the static structure of change had itself changed — the spatial distribution of the differences between two matrices! To illustrate the difficulties of this problem it is best to use an example. Suppose 1,000 people commuted in 1981 from town A to town B while some 800 moved in the opposite direction daily. Between 1981 and 1991 the working population of town A rose from 20,000 to 25,000, while that of town B fell from 50,000 to 40,000. In 1991 1,100 workers commuted from A to B while only 700 moved from B to A. What, in a nutshell, had happened? Now imagine trying to describe the situation when A and B are represented by several thousand small places. The solution of this visualization problem is left as an exercise to the more ambitious reader!

5 Visualizing Places

> "Images are only images. But if they are numerous, repeated, identical, they cannot all be wrong. They show us that in a varied universe, forms and performances can be similar: there are towns, routes, states, patterns ... which in spite of everything resemble each other."
>
> Braudel F. 1979 p.133

The sorts of problem described in the previous sections become progressively easier to handle when there are fewer places involved and hence more space to show the characteristics of (or interactions between) these places. A further advantage is that people at least have an idea where districts, towns and regions are, and so can better transcribe what they see back to what they know as reality, better than they can from a map of over one hundred thousand areal units (within which they have to learn to navigate in a new kind of space). However, the disadvantages of only visualizing census information at the level of the district (constituency, labour market or whatever) is that much of the fascinating spatial detail on society which the census provides, is jettisoned. Further, many of the tentative solutions that arise from using

spatial detail to overcome difficulties such as the modifiable areal unit problem and the ecological fallacy, which become possible when the maximum detail of the information is retained, also have to be abandoned.

This section has been left towards the end because once the objectives are widened, it becomes more difficult to provide useful prescriptions. Once you have the scope to vary the shape, size and other features of your visual representations, then the design constraints loosen. Instead of trying to define the limits of what is possible under these conditions, it is therefore more appropriate to give a few examples of what is plausible. There is, however, one overriding maxim: if the design is to show the overall pattern as well as that within localities: the representations chosen must be able to *meld* into a single image, to be composed of patterns but also to be able to form an overall pattern. They must be able to create what is called a *gestalt* — an organisation which is more than the sum of its parts, a picture which can reflect a whole facet of our society (for many more imaginative examples see Szegö, 1987, or Tufte, 1990).

Beginning with the most simple embellishments, the circles that have been used in so many of the examples in this paper can be turned into 'pies' or concentric 'rings' to show the relative sizes of subsections of the population (by, say, country of birth) or substrata of a phenomenon through time (immigration over time, for instance). **Figure 13** illustrates rings being used to show the changing spatial distribution of county level unemployment differentials over time. Influential practitioners have rarely found either of these methods to be effective because the "superimposition of several images destroys each of them" (Bertin 1981 p.182). The use of numerous pie charts on a map is even more of a pet hate of many cartographers, because they incorporate multiple sins of graphic design. People are unable to judge angles easily and can certainly not compare them in different sized circles across (what often ends up being) a garishly coloured map. Pies and rings are simple graphics, and often suffer from not being well thought out in terms of how they might tessellate.

If we were interested in showing multivariate information about a single place we would probably not use pie charts, but would draw charts. The simplest chart is made up of bars, one bar for the level of each variable. Thus we could show, for instance, the number of people employed in eight types of industry simultaneously. If we divided the bars we could also show the proportions of workers who are male or female. By having the bars horizontal and extending both right and left our chart begins to look like a 'population pyramid' — one side for either sex. By reflecting the pyramid again (this time in the horizontal) the population could be further subdivided, by say part-time and full-time workers — a 'population cross'. Draw one of these crosses in each of hundreds of places and the interactions of some thirty-two variables can be depicted across space.

Figure 14 shows how the visual realisation of such a process might look, showing the variance in employment by industry for parliamentary constituencies. The dominance of areas in the centre of London and other major cities is clear, creating mushroom shaped pyramids where many more people work in administration, finance and services than in manufacturing. The widths of the bars of the pyramids have been scaled to reflect the national proportions of people working in each industry. The coalmines of the East Midlands and the dearth of jobs in many parts of Scotland (at that time) can both be identified. This is not an example of thematic

Figure 13

The Space/Time Trend of Unemployment in Britain, 1978-1990.

(On a 1981 County population cartogram, outer rings being most recent years - scale indicating deviation from space/time independence.)

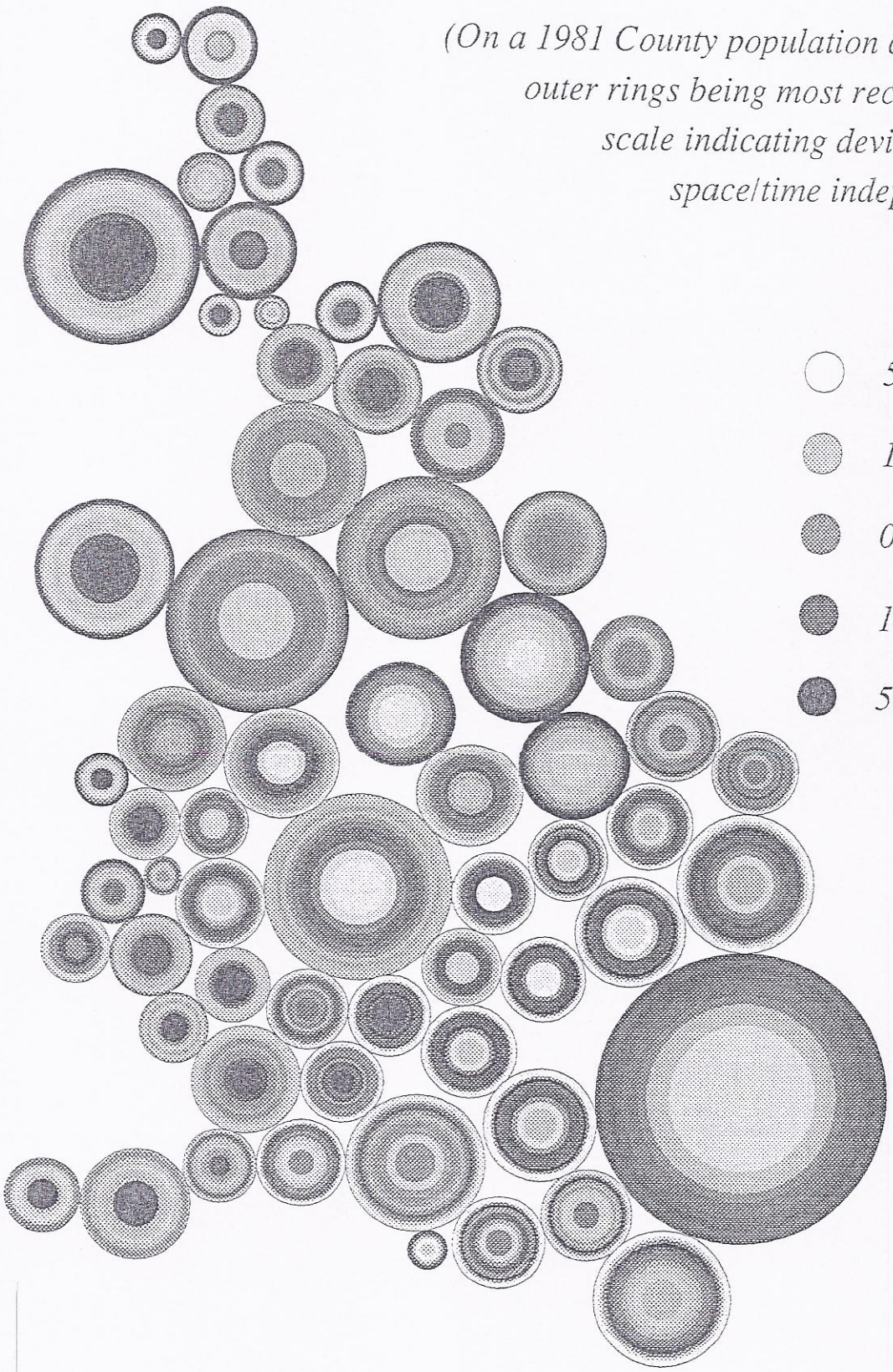
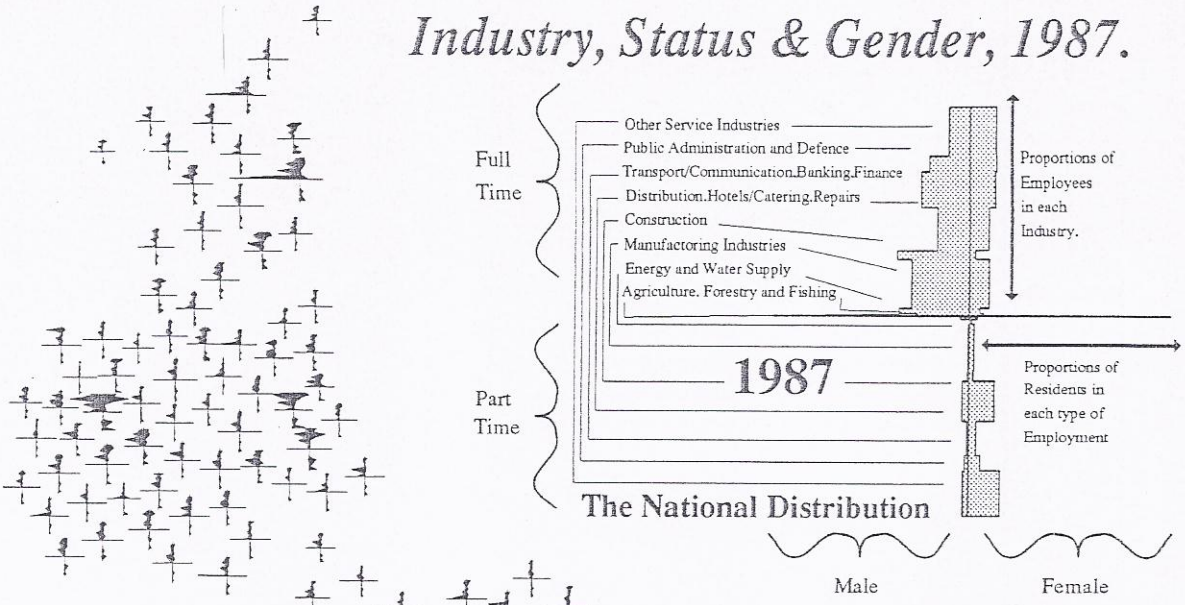
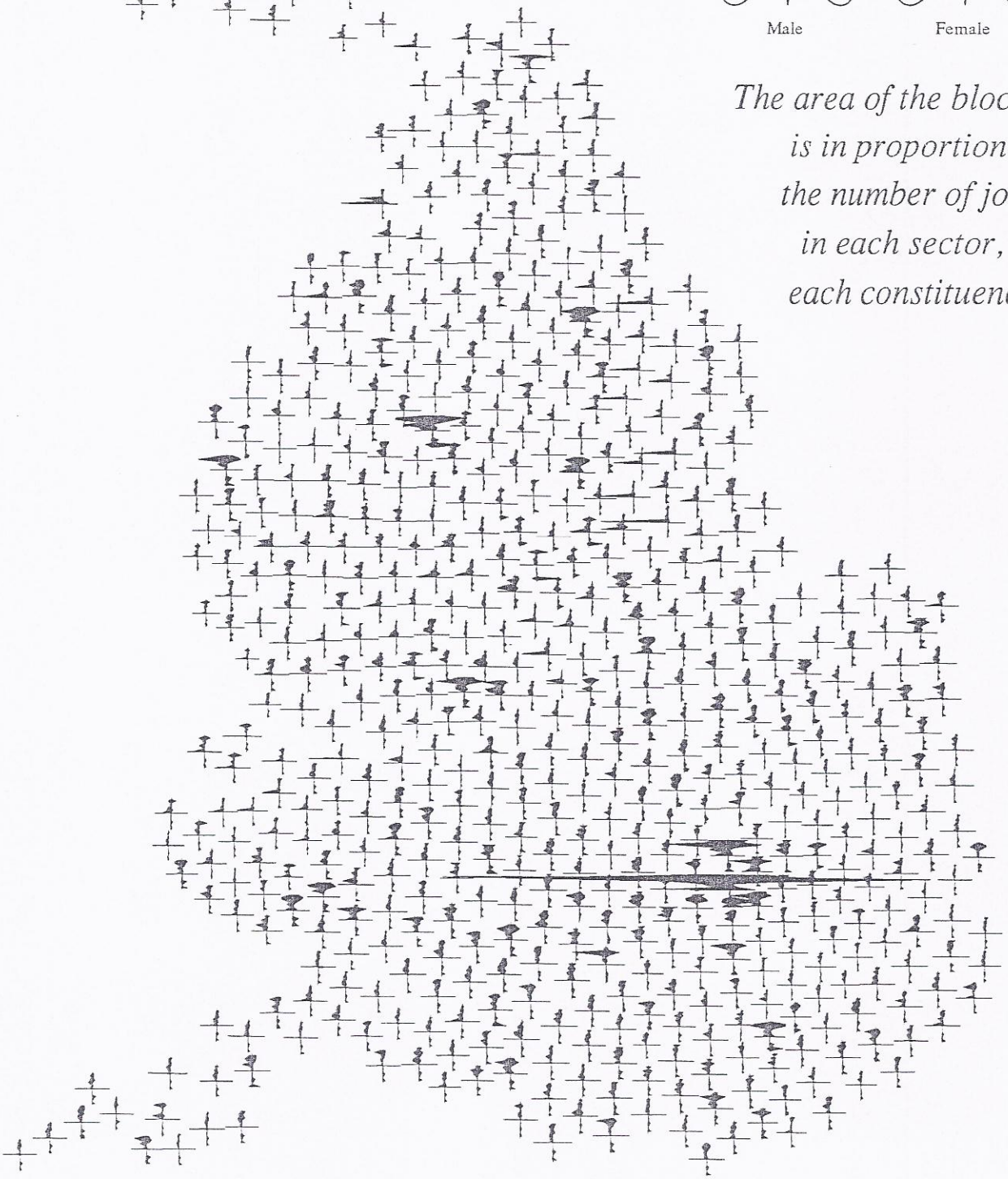


Figure 14

The Distribution of Employment by Industry, Status & Gender, 1987.



The area of the blocks is in proportion to the number of jobs in each sector, in each constituency.



mapping for two reasons. Firstly, it is too complex to expect people to understand easily at a first glance (indeed interactive magnification on a computer screen is needed to fully appreciate the structure it shows). Secondly, it is less 'representing political features of the earth's surface' (from the concise Oxford dictionary) than a visualization of (making visible) what just happens to be related to the earth's physical surface. The details of this point are somewhat prosaic, but in essence the figure fails as an illustration because you already have to know what you are looking at — to know what you are looking at.

One problem with bar charts is that the order in which variables are placed along the chart greatly influences the visual impression given, and that order is arbitrary. If the order of the industries, say, were made the same as their national rankings, then charts where a gradual rise was broken would show areas where the industrial mix was at odds with what would be naively expected if all places mirrored the national distribution. What is most important about these symbols (as with all *glyphs* — "sculptured characters or symbols") is that they create a recognisable shape. It is the outline of the bars that is important, so colouring the whole symbol a single shade creates a single object — admittedly with thirty two appendages — which is simple enough to be relatively easily compared across space.

However, the fundamental difficulty remains. Bar charts, graphs and pyramids were originally designed to stand alone, and thus often contain enough complexity and detail as single entities, let alone when many are displayed simultaneously as map symbols. To be used in a spatial context, glyphs must generalize and simplify the information if the overall patterns are to be understood, particularly if more than a few dozen areas are to be compared. As the number of areas increases, so too do the differences between areas. The industrial structure becomes less predictable and the population structure more varied. Unfortunately as the symbols get smaller and comparison becomes more difficult. We must design simple glyphs which do not require a great deal of space, and which the eye can quickly comprehend without excessive examination.

The most basic, traditional glyph which can satisfy the above criteria is the arrow. The use of lines to show flows between thousands of areas has already been covered. Once we have the space to give the arrows width, size, shape and colour they can show much more details of the flows they depict. The visual advantage of arrows is that their aggregate expresses a form of its own. Like a flock of birds in flight, groups of arrows pointing in a similar direction appear to be *going* that way; they form a single image. This is exactly the impression we wish to create.

One subject which is particularly amenable to the use of arrows is the depiction of change over time in three mutually exclusive variables. Although this technique has been used in Britain mostly to show three-party swings between general elections (eg Upton 1991) it could be used to depict change in any manner of census variables, the change from 1981 to 1991 in the proportions of three groupings of socioeconomic groups falling into social-classes could be a good example. The direction of the arrow can represent the direction of change as a vector with, for example, 'up' meaning an increase in the 'top' group, 'right' indicating an increase in the 'middle' group and an orientation towards the 'left' showing a rise in the 'lower' group (see the discussion below on Figure 19 for further explanation). The length of the arrow would indicate the magnitude of the change. The size of the arrow would be in proportion to the size of the groups, while the position of the arrow is determined by the location of the place it

represents. The colour of each arrow could illustrate the proportions of people in each class in 1991, to give more meaning to the changes in direction (see Dorling 1992 for examples).

In one sense, nine dimensions would be seen in this relatively simple picture — two for position, two for direction, three for colour and one for each of length and size; but that would be a gross exaggeration. The position of the place is shown by two *dimensions*, while the image is representing seven very closely knit *variables*. It is the strength of the relationships between the variables that allows so much to be depicted. The arrows would work well in this example because direction would be meaningful, and the three variables which made up direction were really one — change in class composition. If the image worked well it would be because the spatial relationships between the classes chosen were strong enough for discernible patterns to emerge.

Glyphs more complex than arrows have been specifically designed to allow quick comparison of the overall pattern of multivariate information where less simple structures are expected and where the change in direction of more than three mutually exclusive variables may be of interest. In the literature the best known of these are *trees* and *castles*, where various aspects of a basic shape are altered to produce many variations of an underlying structure which aids their comparison. It is the maintenance of this basic structure which aids our visual assimilation of the objects into an image. What distinguishes these symbols from many of the objects described above, is that they have been specifically designed to be used as glyphs — to be compared quickly with one another.

Castles have various parapets, which alter in height and aspect as the values of the variables change. In many ways they are simply an embellishment of the bar chart, altered so as to allow the mind to form an impression of the general 'shape' of the place more easily, using a more familiar symbol. Bar charts can only go up or down, have a peak here or there, but they are still charts. Castles appear more as single objects, and so it is hoped that an overall impression from a whole 'country's worth' of these can be obtained. Kleiner and Hartigan (1981) used castles to represent the varying fortunes of 15 companies over 25 years, grouped recursively into clusters of similar companies forming separate 'towers' of a single castle. The census variables provide numerous complicated classifications which could benefit from similar forms of visual expression.

Kleiner and Hartigan (1981) also introduced the use of 'trees' in order to show changes in variables which can be recursively grouped (because a tree recursively branches). The image of a tree is also visually familiar and thus relatively easy to interpret. Just as castles grew out of bar charts, trees have grown out of the simple graphic symbol of the weather vane (used to show average wind directions through the lengths of eight radial 'spokes'). Rather than order the spokes as a wheel, they have become the branches of a tree. This works because we are used to seeing tree-like objects which vary in shape but have a rough symmetry about them. The order in which the variables are assigned to the branches and twigs is crucial for the impression gained. It is usual to place the total population variables at the base, put the smaller subgroups to the sides and so on. Whether this works or not depends on the variance in the information being depicted. A relatively convincing 'wood' can be created; thickets, copses and spinneys of different species can be identified (i.e., overall tendencies for trees to have a certain

combination of features in certain parts of the picture, and for other combinations never to occur). **Figure 15** shows the use of trees to depict different aspects of the British housing market. The branches show different categories of housing; their widths being the number of sales in any one year — their lengths, the average prices. Their areas, then, are in proportion to the value of the sub-market and they sum to the total value of the housing market in each place. This is not census information but it could easily be combined with statistics such as the number of people buying their homes to enhance that information.

The most contentious glyphs created to date are based on human faces, drawn by Herman Chernoff. Faces, it is argued, are the visual image we are best equipped and experienced to decipher. We naturally combine their features to interpret moods such as happy or sad, sly or dull. What is more, we can easily compare faces to look for family resemblances or the mood of the crowd. Faces maintain a basic structure in which even a slight variation often holds meaning.

The original Chernoff faces aimed to show the values of as many as eighteen variables simultaneously (Chernoff 1973). These have been criticised because certain different combinations of variables could result in the same visual expression and because the order in which so many variables were assigned to features had a drastic effect on the impression gained. The example shown (**Figure 16**) portrays only five variables, using just eleven *Bezier* curves to construct the faces which reflect their values. The image combines variables from the last two Figures and introduces some more, to show how these glyphs can be used to compare statistics which do not fall into simple groups or categories.

Figure 16 was developed as part of a study of social change between the 1983 and 1987 British general elections, with each parliamentary constituency being represented by a face. The size of the face is drawn in proportion to the electorate, the shape of the face by house price change (fat if rising more than average — thin if less), the shape of the mouth by employment (smile if rising), the width of the nose by electoral turn-out (big if rising) and the eyes by industrial structure (large low eyes for more service employment, small and high on the face where other industries predominate). The 'disappointment' of parts of the north is clear in the signs of gloom in those constituencies 'faces'! Scotland's faces' are thin (where house prices rose slowly), while many inner city noses are small as turnouts are low. Colouring the faces by the change in the vote would allow the spatial connections between these social changes, and electoral change, to be seen where and if they occurred. There is, of course, a great danger here of evoking the ecological fallacy in quite a spectacular way. These are the 'faces' of places, not people.

Strong local relations in space are perhaps the clearest message to be formed by the images shown here. Sharp divisions are also immediately apparent, as are more gradual changes. The faces can be seen to be reflecting places' *reactions* to a changing social situation. In this sense they are the ultimate form of spatial reification — giving a place lifelike characteristics and opinions which it can not actually have. Chernoff faces are contentious for the very reason that they are seen as useful. Peoples' reactions to faces are much stronger than their reactions are to more neutral objects, which are claimed to depict information more objectively. In perception, there is a continuum from personal likes and dislikes of certain colours in maps, to individuals' strong reactions to cartoon faces. Visualization of social information, however, is all about

Figure 15

The Changing Distribution of Housing by Price, Attributes and Sales, 1983-1987

The area of the branches is in proportion to 5x the total housing inflation by each set of attributes, in each constituency.

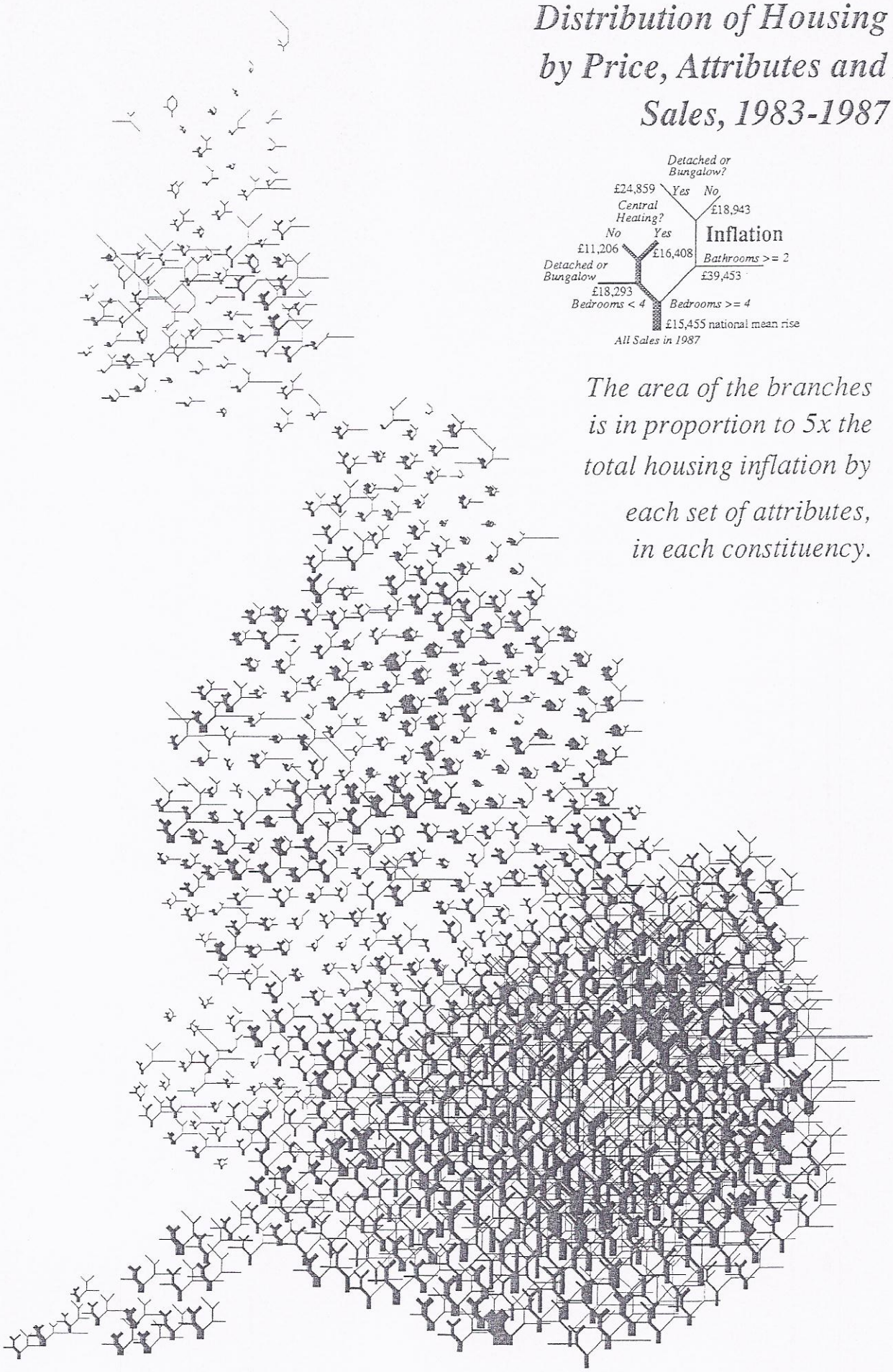


Figure 16

The Distribution of Voting, Housing, Employment and Industrial Compositions in Constituencies in 1987.

Facial features are in proportion to the relative values of the social and economic characteristics of the areas they represent.

engaging our imaginations and emotions. The emotional response to Chernoff faces can therefore be thought of as a strength as well as a weakness.

A more serious criticism of the use of glyphs is that they can overload the viewer with information: too much is being asked of the eyes and the mind. In this section it has been shown how badly designed symbols are impossible to decipher spatially, while better thought out images can help the viewer form higher level structures from the simple pictures of collections of places. It is a mistake, however, to think that these symbols can add another dimension to the two we have on paper. Glyphs show multivariate structure, not multidimensional form. We can look at a great many categorical aspects of many places simultaneously, but varying the features of an object is not a good substitute for varying its position because different combinations of colour and shape can only imply simple relationships, such as correlation over time.

Just because we have the space to use more sophisticated symbols does not predicate the use of more simple graphical techniques to tackle complex problems. **Figure 17** attempts to depict the changes in population over the three decades from 1961. Three-colour mixing on a cartogram is used. Places are black when their populations have declined faster than the national average during each of the three periods 1961-1971, 1971-1981 and 1981-1991, and white when they have consistently risen more quickly. The rural belt of the home counties is white while most of inner London is black. More subtle shades of colour characterise other changes. Much of the South West (and Tower Hamlets) is green because the last decade has seen a reversal in its long term population decline. The urban/rural fringe districts are orange as their populations grew fastest in the 1960s, after which stricter planning controls were applied and demographic factors came into play.

The wealth of information which can be gained from using sophisticated techniques such as these appears great. But there is a good deal we are not able to see in the structure (which in the last example is derived from less than two thousand statistics). This is because we cannot show change over time and space simultaneously in two dimensions (although the graphs below the figure's title try to give an idea of the kinds of temporal movements involved). Features such as the colour of an object have no geometry and so are limited to conveying only a few values. To attempt to see into a real third dimension we must begin to think in terms of volume, discussed in the final section of this chapter, a possibility which becomes feasible when we only wish to view a few regions.

6. Visualizing Regions

> "The world is complex, dynamic, multidimensional; the paper is static, flat. How are we to represent the rich visual world of experience and measurement on mere flatland?"
>
> Tufte 1990 p.9

This, penultimate, section begins to explore the possibilities when the spatial distributions over just a few dozen areas are the interest. These are the numbers involved with, for instance,

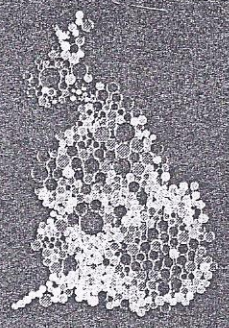
Population Change in Britain by District, 1961 - 1991.

Figure 17

Individual Distributions of Change:

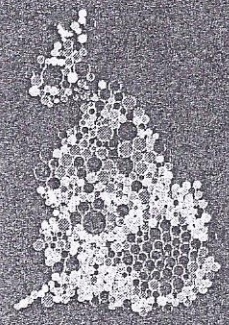
The area of each circle represents the total population of a district. The circle's colour shows how that has changed:

Each colour shows the population change trajectory of a group of Districts. Low 61/71 growth is coloured blue, 71/81 - yellow and 81/91 - red. Mixing to make many different shades.



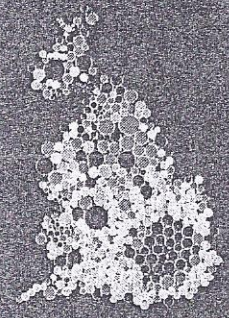
1981 / 1991 Change

Median 0.9% increase - light above, dark below.



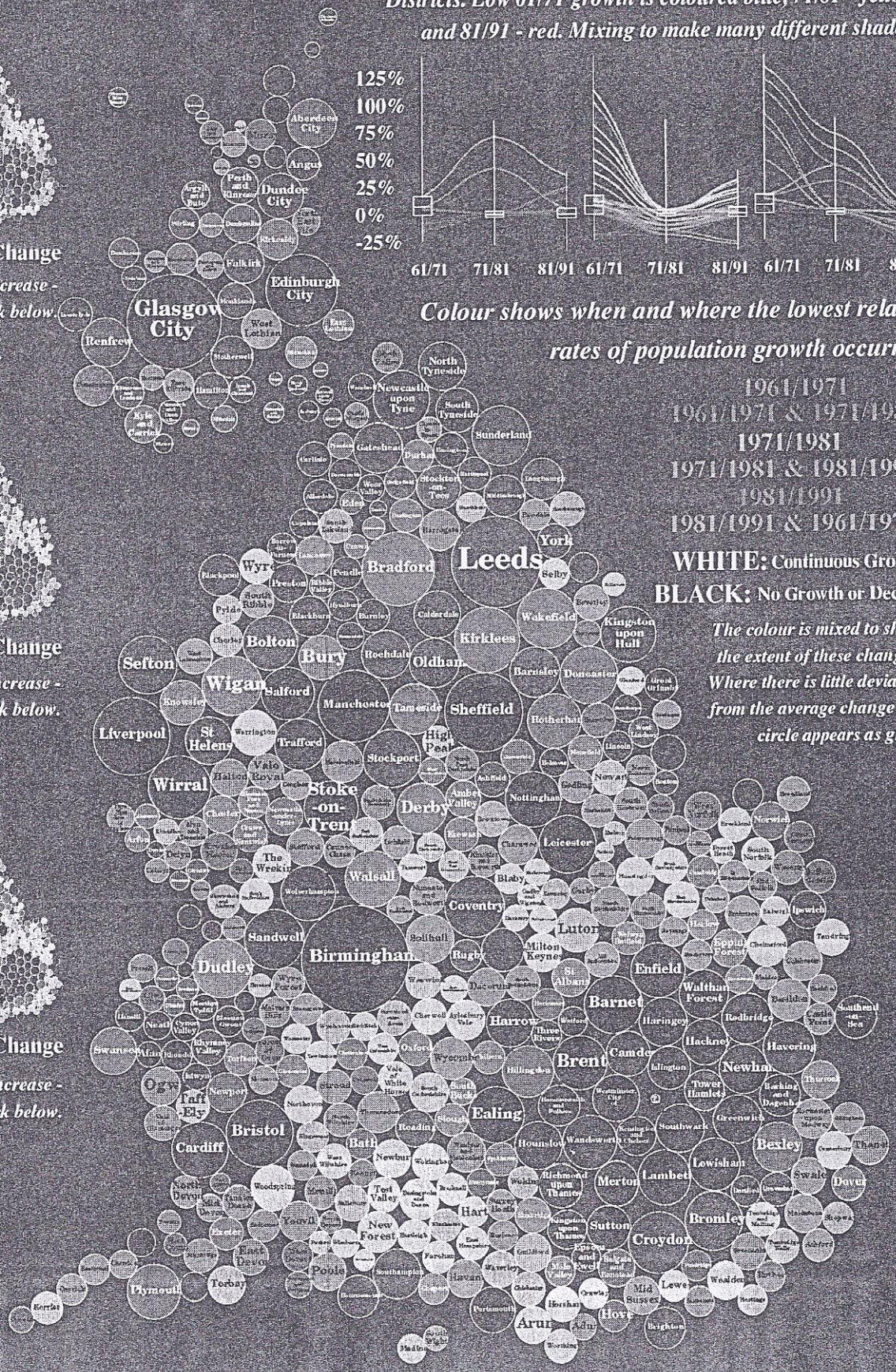
1971 / 1981 Change

Median 4.1% increase - light above, dark below.



1961 / 1971 Change

Median 8.5% increase - light above, dark below.



1961/1971
1961/1971 & 1971/1981
1971/1981
1971/1981 & 1981/1991
1981/1991
1981/1991 & 1961/1971

Colour shows when and where the lowest relative rates of population growth occurred:

WHITE: Continuous Growth

BLACK: No Growth or Decline

The colour is mixed to show the extent of these changes. Where there is little deviation from the average change the circle appears as grey.

Daniel Dorling 1991

English counties, or with standard regions. While it is generally bad geographical practice to look for patterns in society at such an aggregated level, there are a few advantages to so doing. First, much non-census information is only available at these levels, including many of the results of annual official surveys. Second, people are even more likely to be familiar with the concept of regions than, say, districts and so have less to learn in order to begin to understand the image (not always a good thing). Most importantly, however, once we become interested in changes over time and other subjects which involve a more detailed geometry — the simplicity of regions compensates for the complexity of the subject.

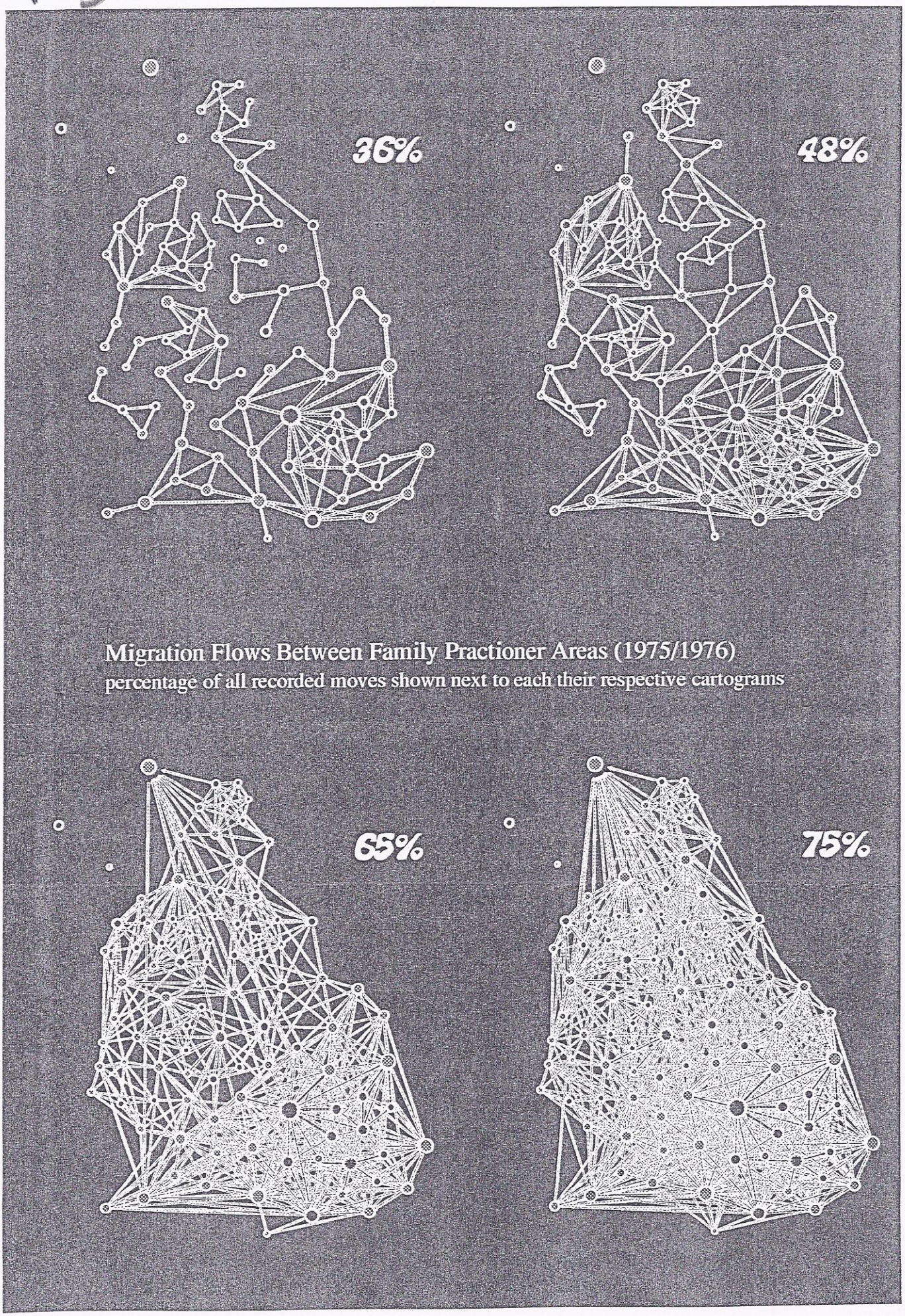
There are some new forms of census information which are only available at these relatively crude level of aggregation. The 2% Sample of Anonymised Records of individuals is available only for aggregations of districts where more than 120,000 people live. The 1% sample of individual households is geographically disaggregated by only twelve standard regions. The Longitudinal Study, which links the records of a 1% sample of individuals between censuses traditionally provides information only down to the district level, with many analyses suppressed for areas smaller than regions. Thus there is a great deal of interesting census information for which very little spatial detail is available, but which we still wish to visualize (partly because it is often of a very complex nature).

A second reason why we might wish to use such crude geography is to look at change over time in the years between censuses. A practical example is in the monitoring of migration flows, where the most spatially detailed information available between censuses is for the one hundred or so Family Health Service Authorities of England and Wales. The National Health Service Central Register of patient movements provides a full migration matrix at this level of aggregation since 1975 on a quarterly basis. The advantage of having so few areal units is that more sophisticated graphic techniques can be employed.

Figure 18 shows the migration streams comprising the significant 65% of all the recorded moves between 1975 and 1976. A population cartogram is used to ensure some clarity; two concentric circles are placed at the centre of each area being represented, their size in proportion to the numbers of in- and out-migrants, with a black circle uppermost indicating an overall loss of people, and the net flow being proportional to the gap between the circles (and vice versa for white). An arrow is drawn between each pair of areas with a significant flow, with its width in proportion to the numbers of people moving, and the arrow-head pointing in the direction of net flow. If the reverse flow is also significant, it is placed on top of this (so as not to be obscured), with its arrow-head at the other end of the line. A white border is then placed around each arrow to distinguish it from the general tangle of lines. One graphical subtlety is the sorting of all these arrows so that the ones between contiguous areas are drawn uppermost, the ones between second-order contiguous areas below them, and so on. The most important decision to make, however, is always which flows to show and which to ignore, by determining what is a significant movement.

By raising and lowering the level of significance you choose, the picture becomes more or less cluttered as lines appear and disappear. A sequence of images showing how the structure builds up can be quite revealing (Dorling 1991). The Pennines appear as a formidable barrier to movement, while the North West cluster of areas is one of the last groups to connect to the

Figure 18



Migration Flows Between Family Practioner Areas (1975/1976)
percentage of all recorded moves shown next to each their respective cartograms

national network of migration streams. Animation of the image, as the level of significance changes, highlights this quite well and helps understanding of a very complex diagram. The much higher general rates of migration in the south east can be readily distinguished from the Figure, with a Wash-Exe line dividing that dense mass of lines from the much lower levels in the North. Use of the population cartogram as a base gives the density of lines (and the areas of the circles) some meaning. The tightness of flows between the areas of the North East is clearly apparent, represented by the 'hexagon' of areas in the top right. The strong connections between the single northernmost area representing Scotland, with parts of London, is also very obvious: Scottish migrants flow right the way across the country, generally avoiding the places in between.

It is, however, much more difficult to explore graphically the spatial nature of changes to the flows over time. This can be demonstrated by taking a simpler problem than the migration example. Unemployment data is now available at very frequent intervals at a highly disaggregated level. In the past, this detail was not available and the further we go back in time the less detail there is. Suppose we wished to investigate how the levels of unemployment had changed annually at a fine level annually from the end of the 1970s to just before the latest census was taken. If, for each of over 800 (unemployment office) areas we want to show twelve levels of unemployment. How could this be done?

Nationally, the level of unemployment varies dramatically over time. If we were just to plot the crude rates it would be very difficult to distinguish between different areas. What is of interest here is the time a particular level of unemployment is 'unusual' for a particular place and a particular time. To measure this, the conventional technique is to calculate an expected level (given the place and year) and subtract that from the actual level. How then can we show these twelve deviations? Using an area cartogram in which the places are represented by circles, the simplest option is to show each deviation by colouring (different greys) each one of the twelve rings in each circle. Figure 13 has already illustrated this technique. Using an analogy with the cross section of a tree, the outer rings can represent the most recent years and the inner core the first year in the series. Each ring should have its area in proportion to the total workforce of that year. Where a place such as London is light in its centre and gets darker around its edge, that signifies that the level of unemployment there is rising from below the national trend to progressively further above it. Many places on Merseyside, Tyneside and Clydeside have a dark core and dark circumference and are paler in between, indicating that these places did "relatively well" when unemployment levels were at their highest nationally, because they were then not so unusual in their high levels (this also illustrates a problem of using such simple measures). Oxfordshire is the only county which clearly is improving consistently against the national trend, at least up until 1990.

If phenomena are found to be too complex for representation in this way, one form of escape is to look at even fewer cases or places. With just twelve regions it is possible to stop mapping the geography entirely and just use charts. Classically, a line chart is used with one line for each region, the horizontal axis being time and the vertical axis being the value of some variable. For visualization purposes we could begin to break the rules a little: only show lines when they were interesting, choose in which order they were to overlap, give them a thickness which altered along their length, hang other symbols off them, or even let them bifurcate (ie split) if it

was appropriate to divide a region into two at some point. All we are doing is exploring and embellishing from a basic graphical model, but freed from the restrictions of having to represent the geography of the situation in some form. The graphical possibilities then very quickly become limitless. Because we are not trying to illustrate the numbers, but to understand them better, we do not need to stick to accepted templates.

A good example of what becomes possible when you are no longer trying to show geographical patterns is provided by triangular graphs which use space to show the relationship between three dependent variables across many places when there is no wish to show places' spatial relationships with each other. **Figure 19** demonstrates this with local election results collected for the 'divisions' which county councillors represent (areas somewhat larger than wards). Over 3,000 places are shown, again as circles, their areas in proportion to their electorates and their positions within the triangle determined by the share of the vote of each of the three major parties (where one party was not represented, a histogram is built up on the side of the triangle opposite the missing party's apex). The places form a shape in the 'variable space' of the triangle. In this case they demonstrate one stage in the polarisation which was taking places as the cluster of divisions formed a 'V' shape — seats either aligning on the Labour-Conservative apex, or along the Democrat-Conservative line. Changes in the shapes formed on these triangles over time can produce very interesting patterns.

One method increasingly being advocated for studying change over time (and other changes) in visualization, is animation. In studying elections it was with changes between triangular graphs that this writer found animation to be most useful, because drawing lines between points on the graphs quickly created a mess (Dorling 1992). The other major uses for animation are in panning and zooming over a very detailed image, or in illustrating very simple movements to a large audience — for instance the movement of the "centre of population" of a county over time (Monmonier 1992). In general, flying over landscapes created by raising the ground by the level of some social variable may not be a particularly useful way of understanding its distribution — but others might disagree. Where animation and perspective viewing might be most useful, however, is in studying the geometry of the more complex abstract worlds we might create — a landscape where distance is drawn in proportion to travel time for instance; but here we are quickly departing from what it is possible to visualize with census data.

When creating the illusion of a surface on a two dimensional medium, depth cues have to be used — perspective and shading. But if that surface, created by the values of one variable over space, is to be coloured according to the distribution of another variable, the information is lost and confounded. It it lost because it cannot be seen (being behind part of the surface), and it is confounded because shading and shadow intermingle. To really appreciate three dimensional geometry you must be able to fly around that surface, or at least be able to 'rock' the image from side to side. In doing this, however, you are never quite able to get a hold on what you are looking at (Kaufman et al 1990 p.162). A surface is not an alternative for a choropleth map, because on the one hand it contains more information, and on the other hand it contains less. A surface expresses geometry — distances between objects on the surface are defined — so 'shortest' routes can be seen which are not always straight lines. This is an advantage when a geometry is being defined explicitly, but if you are not trying to measure distances in terms of, say, years of unemployment, then the advice which applied to the precursor of the perspective

Figure 19

1989 County Council Elections: English Voting Composition

Every electoral division won by one of the three major parties is shown by a circle on the diagram.

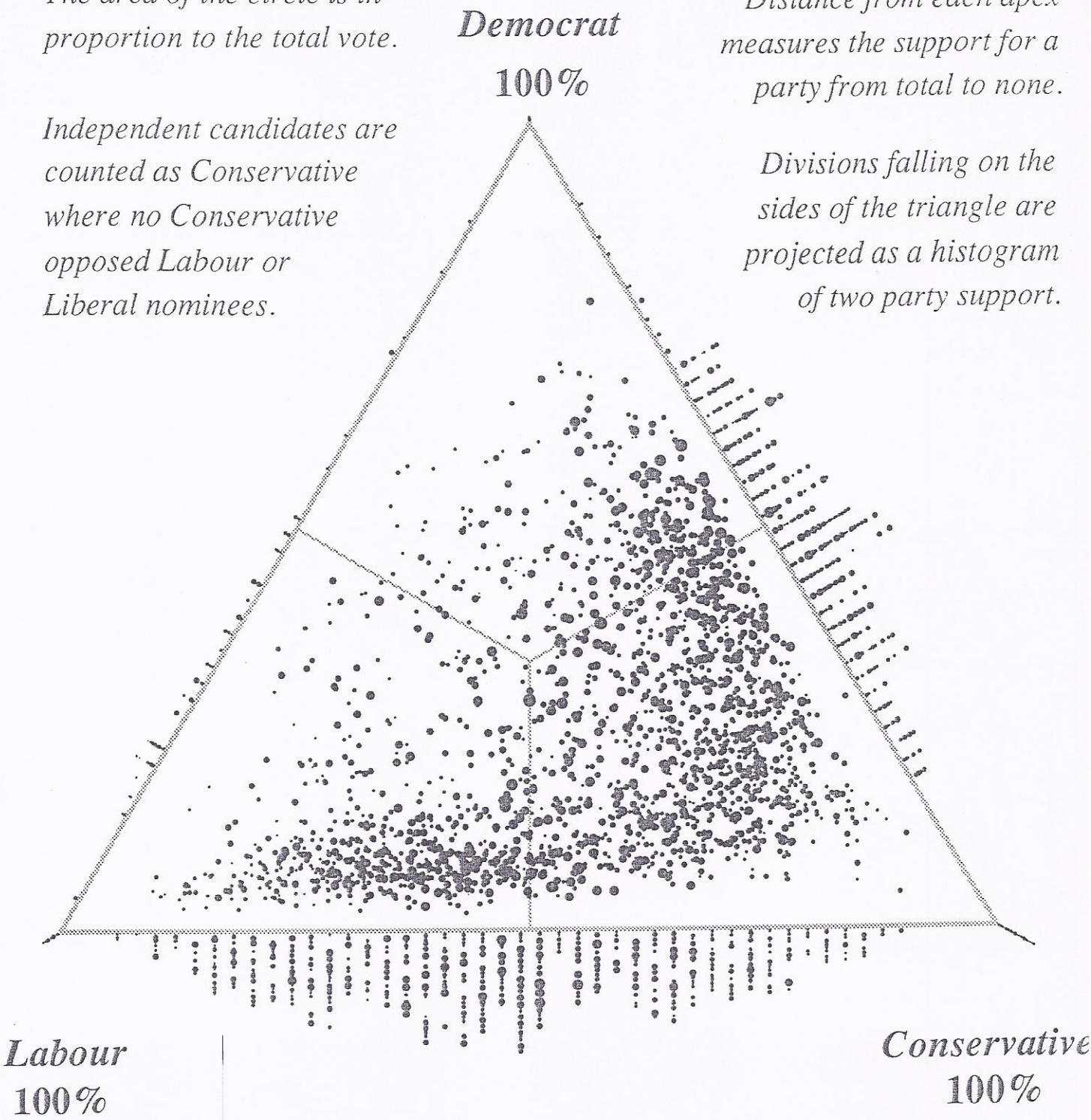
The area of the circle is in proportion to the total vote.

Independent candidates are counted as Conservative where no Conservative opposed Labour or Liberal nominees.

The position of the circle indicates the composition of votes in that division.

Distance from each apex measures the support for a party from total to none.

Divisions falling on the sides of the triangle are projected as a histogram of two party support.



view still applies today:

> "— it is extremely wise to avoid contour mapping of census data."
>
> Rhind 1983 p.190

We do not think in a three dimensional geometry — many tests have shown this (Parslow 1987). The geometry of visual thinking is essentially two dimensional. Many people cannot actually perceive depth (Young et al 1988 p.419). Another disadvantage to be remembered is that we have poor visual memory (there is just too much to remember). Of course, there are problems with even the most basic of visual techniques: the emotional overtones of colour are perceived differently by different people; the colour blind cannot see the full trivariate range; and, of course, some people cannot see at all. Visualization is not a panacea. With information as rich as that from the census you can quickly exhaust the possibilities before you are able to study the features in which you are most interested. You can, however, always then invent new possibilities.

7. Conclusion

Visualization of the census is not an easy subject to cover in just one chapter because there is so much more that could be said and drawn. I have not discussed the practicalities of visualization — the hardware and software — because these are changing so quickly. All the illustrations were produced by simple programming and the in-built software of a standard 'Acorn' British school microcomputer. If you pay a significant amount of money for a "visualization package" it should be able to do a great more than this. There is still an advantage to programming from scratch, however, because totally new ideas cannot be implemented using other people's applications. Programming languages are also becoming simpler to learn. Geographical Information Systems (GIS), on the other hand, have severe limitations in this field:

> "In fact few of the graphical operations feasible in a GIS are going to be appropriate for census data."
>
> Barr 1993 p.267

Traditional problems of visualizing the census can be superseded if a new methodology is adopted. Difficulties such as boundary change need not be a problem since, as people move, so do the collating boundaries move around them, eternally attempting to encompass them. An animation of changes in the population can also incorporate change in the boundaries. If cartograms are used, places 'disappear' and are 'born' as they shrink and grow in line with their populations and geographical definitions. Use of cartograms also avoids the age old problem of wasting most of the paper showing the characteristics of the fewest numbers of people at the expense of the majority — who often are those about whom we know the least, who tend to be poorer, less visible in our cities and on our maps.

On the subject of flow mapping, which this chapter has dealt with at several stages, it should be pointed out that not all flows are spatial — flows between tenure and social groups can be

measured from the Longitudinal Study; flows on and off the unemployment register can be seen, as can flows of votes between parties measured by electoral surveys. All these aspatial forms of flow represent challenges for future new methods of visualization. Depicting how these matrices of flows have changed over time is an even more challenging problem waiting to be addressed, and after that, showing the spatial component of these changes!

Graphic depiction has come in and out of fashion in cycles over the last two hundred years (Beniger & Robyn 1978). Its resurgence is usually caused by the invention of new printing technologies and the availability of more abundant information, so we should not be surprised to see a rise in visualization as, internationally, data from the 1990 and 1991 censuses are released and high quality graphics workstations become widespread. **Figure 20** shows dasymetric mapping being first used in the British census of 1851. Now we can do this using the postcode/enumeration district lookup file in a matter of seconds. Our methods may be quicker, but how much of an innovation do they represent?

A new challenge to census visualization is to see how time and space can be transformed to represent more clearly the patterns within them, on paper or in animation (Rucker 1984). A landscape where distance is travel-time has been mentioned above and should be possible to create. Ultimately we ask: can we compare the evolution of one thing with the flows of another and the distribution of yet others, without collapsing reality into dimensions which cannot contain its complexity? One dream of this writer is to make possible the interactive visualization of a single dynamic map of Britain which can be asked to show whatever variables you wish to see in a given form, from whichever census, for any areal units, using any projection that can be defined.

Pictures of spatial social structure should have the power to reflect the complex tapestry and delicate lace-work of the relations between people, through the places in which they live and the spaces they create. As we are inundated with the greatest volume of information about our society that we have ever had in digital form, new methods and methodologies will inevitably be created to study it. In all this technological excitement we should always remember to ask — why, for whom and for what are we looking?

Acknowledgements

Thanks are due to Tony Champion, Mike Coombes, James Cornford and David Dorling who all commented extensively on the first draft of this chapter. The research this chapter is based on was funded by the Strange Doctoral Studentship of the University of Newcastle upon Tyne between 1990 and 1991.

Figure 20

ENGLAND AND WALES.

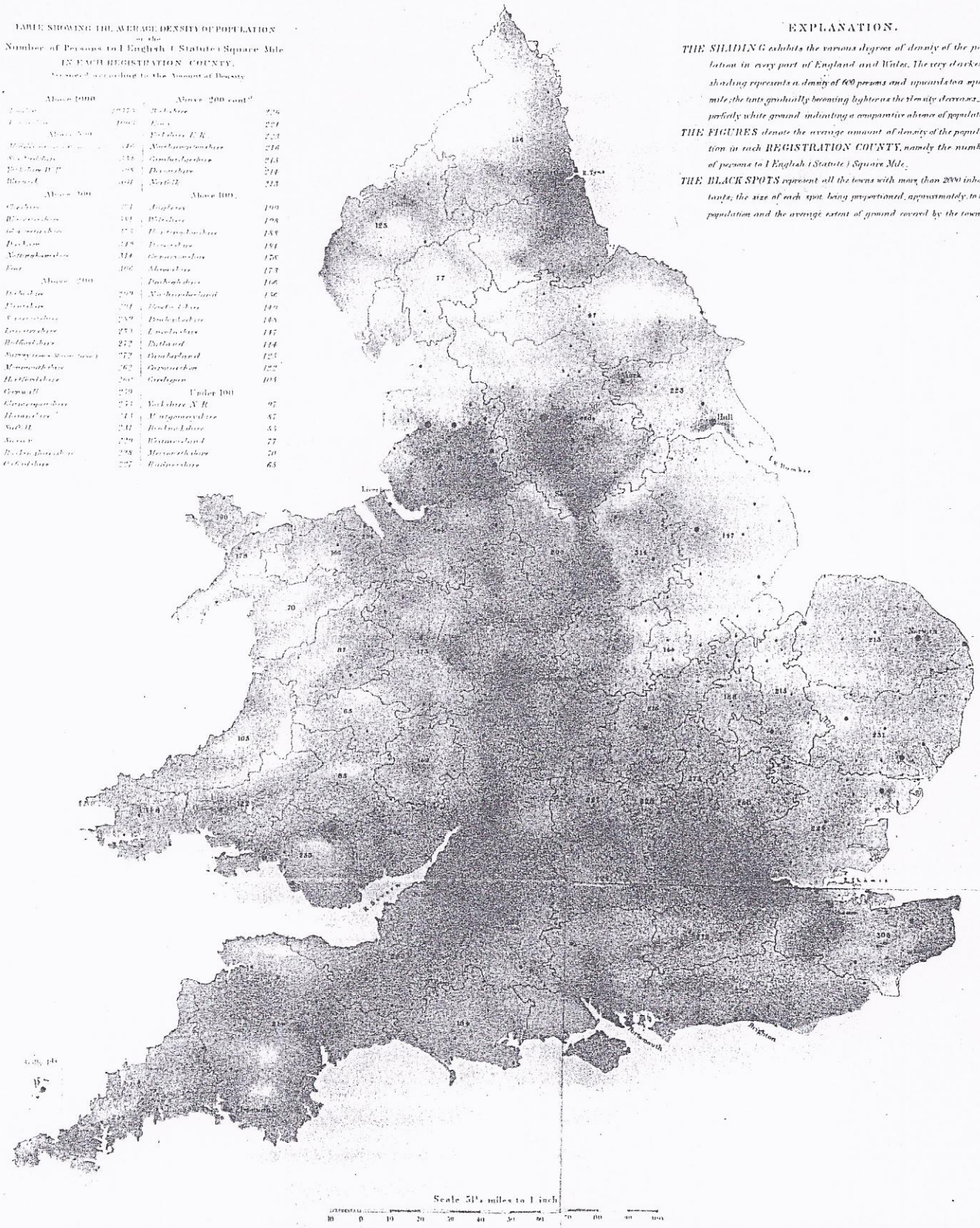
DISTRIBUTION OF THE POPULATION.

CENSUS OF 1851.

Designed by Augustus Petermann, Phys.l Geographer to the Queen.

TABLE SHOWING THE AVERAGE DENSITY OF POPULATION
or the
Number of Persons to 1 English (Statute) Square Mile
IN EACH REGISTRATION COUNTY,
Arranged according to the Amount of Density.

Above 1000		Above 200 cont.d	
[illegible]	[illegible]	[illegible]	226
[illegible]	[illegible]	Essex	221
Above 500		Yorkshire E.R.	221
[illegible]	[illegible]	Northamptonshire	216
[illegible]	[illegible]	Cambridgeshire	213
Yorkshire W.R.	[illegible]	Devonshire	211
Warwick	[illegible]	Norfolk	213
Above 300		Above 100	
Cheshire	[illegible]	Anglesey	199
[illegible]	[illegible]	Wiltshire	198
[illegible]	[illegible]	Hertfordshire	189
Derbyshire	[illegible]	Pembrokeshire	181
Nottinghamshire	314	Carmarthenshire	176
Kent	306	Monmouthshire	173
Above 200		Pembrokeshire	166
Berkshire	299	Northumberland	156
Flintshire	291	Brecknockshire	149
Cambridgeshire	289	Buckinghamshire	148
Leicestershire	253	Lincolnshire	147
Bedfordshire	242	Rutland	114
Surrey (inc. Metr. Sect.)	232	Cumberland	125
Monmouthshire	262	Carmarthenshire	122
Hertfordshire	260	Cardigan	104
Cornwall	259	Under 100	
Gloucestershire	253	Yorkshire N.R.	97
Hampshire	244	Montgomeryshire	87
Suffolk	241	Brecknockshire	83
Sussex	239	Westmoreland	77
Buckinghamshire	238	Merionethshire	70
Oxfordshire	227	Radnorshire	65



EXPLANATION.

THE SHADING exhibits the various degrees of density of the population in every part of England and Wales. The very darkest shading represents a density of 600 persons and upwards to a square mile; the tints gradually becoming lighter as the density decreases—the perfectly white ground indicating a comparative absence of population.

THE FIGURES denote the average amount of density of the population in each REGISTRATION COUNTY, namely the number of persons to 1 English (Statute) Square Mile.

THE BLACK SPOTS represent all the towns with more than 2000 inhabitants, the size of each spot being proportional, approximately, to the population and the average extent of ground covered by the town.

List of Figures

References

Arheim, R., 1976, The perception of maps, The American Cartographer, Vol.3, No.1, pp.5-10.

Barr, R., 1993, Mapping and spatial analysis, in Dale A. & March C. (eds) The 1991 Census User's Guide, London: HMSO.

Beniger, J.R. & Robyn, D.L. 1978, Quantitative graphics in statistics: a brief history, The American Statistician, Vol.32, No.1, pp.1-11.

Bertin, J., 1981, Graphics and graphic information processing, Walter de Gruyter: Berlin.

Braudel, F., 1979, The wheels of commerce, Fontana Press: London.

Brunsdon, C., 1991, Estimating probability surfaces in GIS: an adaptive technique, EGIS '91 Proceedings, pp.155-164, Brussels, Belgium.

Census Research Unit, 1980, People in Britain — a census atlas, H.M.S.O.: London

Chernoff, H., 1973, The use of faces to represent points in k-dimensional space graphically, Journal of the American Statistical Association, Vol.68, no.342, pp.361-368.

Cole, K. 1993, Sampling error and the 10% small area statistics, Manchester Computer Centre Newsletter No. 22: University of Manchester.

Denes, A. 1979, Isometric systems in isotropic space: map projections: from the study of distortion series 1973-1979, Visual Studies Workshop Press, New York.

Dorling, D. 1991, The visualization of spatial social structure, Phd thesis: University of Newcastle upon Tyne.

Dorling, D. 1992, Stretching space and splicing time: from cartographic animation to interactive visualization, Cartography and Geographical Information Systems, Vol.19, No.4, pp.215-227, 267-270.

Harley, J.B., 1989, Deconstructing the map, Cartographica, Vol.26, No.2, pp.1-20.

Hunt, A.J., 1968, Problems of population mapping: an introduction, Institute of British Geographers, Transactions no. 43, April 1968, London: George Philip & Son.

Kaufman, A., Yagel, R., Bakalash, R. & Spector, I., 1990, Volume Visualization in cell biology, Visualization '90, Proceedings of the first IEEE conference on visualization, pp.160-167, October, San Francisco.

Kleiner, B. & Hartigan, J.A., 1981, Representing points in many dimensions by trees and castles, Journal of the American Statistical Association, Vol.76, No.374, pp.260-276.

MacEachren A.M., 1992, Visualization, Geography's Inner World: Pervasive Themes in Contemporary American Geography, R. Abler, M. Marcus and J. Olson (eds), Chapter 6, Rutgers University Press: New Jersey.

Mandlebrot, B.B., 1983, The fractal geometry of nature, W.H.Freeman & Co.: San Francisco.

Marr, D. 1982, Vision, W.H.Freeman & Co., New York.

Martin, D. 1991, Geographic information systems and their socioeconomic applications, Routledge: New York.

McCormick, B.H., Defanti, T.A. & Brown, M.D. (eds), 1987, Visualization in scientific computing — a synopsis, IEEE Computer Graphics and Applications, Vol.21, No.6.

Miller, H., 1934, Tropic of Cancer, London: Grafton (quote from the ninth edition, 1965).

Monmonier, M., 1992, Directional profiles and rose diagrams to complement centrographic cartography, Journal of the Pennsylvania Academy of Science, Vol.66, No.1, pp.29-34.

Olson, J.M., 1976, Noncontinuous area cartograms, The Professional Geographer, Vol.28, No.4, pp.371-380.

Openshaw, S., 1982, The modifiable areal unit problem, Concepts and Techniques in Modern Geography No.38, Geo books: Norwich, England.

Openshaw, S., 1984, Ecological fallacies and the analysis of areal census data, Environment and Planning A, Vol.16, pp.17-31.

Parslow, R., 1987, A new direction for computer graphics, Computer Bulletin, September, pp.22-25.

Pred, A., 1986, Place, practise and structure; social and spatial transformation in Southern Sweden: 1750-1850, Barnes and Noble Books: Totowa: New Jersey.

Rees, P., 1992, The data of the decade, paper presented to the Censuses Analysis Group meeting, University of Leeds, 6-7th April.

Rhind, D., 1983, Mapping census data, in D. Rhind (ed) A Census User's Handbook, Methuen: London.

Robinson, A.H., 1989, Cartography as an art, in D. Rhind & D.R.F. Taylor (eds) Cartography Past, Present and Future, pp.91-102.

Rucker, R., 1984, The fourth dimension: towards a geometry of a higher reality, Houghton Mifflin Co.: Boston.

SEEDS, 1987, The south-south divide, The South East Economic Development Strategy Association: Harlow, England.

Sternberg, B., 1993, Military and civilian GIS work in tandem, GIS Europe, Vol.2, No.3, pp.16-17.

Szegö, J., 1984, A census atlas of Sweden, Statistics Sweden, Central Board of Real Estate Data, Swedish Council for Building Research and the University of Lund: Stockholm.

Szegö, J., 1987, Human Cartography: mapping the world of man, Swedish Council for Building Research, Stockholm.

Tanur J. M. et al (eds), 1985, Statistics: a guide to the unknown, Wadsworth & Brooks/Cole: California.

Tobler, W.R., 1973a, a continuous transformation useful for districting, Annals of the New York Academy of Sciences, Vol.219, pp.215-220.

Tobler, W.R., 1973b, choropleth maps without class intervals?, Geographical Analysis, Vol.3, pp.262-265.

Tobler, W.R., 1989, An update to numerical map generalization, in R.B. McMaster (ed) Numerical Generalization in Cartography, Cartographica: University of Toronto Press.

Tufte, E.R., 1983, The visual display of quantitative information, Graphics Press: Chesire, Connecticut.

Tufte, E.R., 1990, E.R., 1990, Envisioning Information, Graphics Press: Chesire, Connecticut.

Tukey, J.W., 1965, The future processes of data analysis, Proceedings of the tenth conference on the design of experiments in army research development and testing, report 65-3, Durham N.C., U.S. army research office, pp.691-725.

Upton, G.J.G., 1991, Displaying election results, Political Geography Quarterly, Vol.10, No.3, pp.200-220.

Young, F.W., Kent, D.P. & Kuhfeld, W.F., 1988, Dynamic graphics for exploring multivariate data, W.S. Cleveland & M.E. McGill (eds) Dynamic Graphics for Statistics, Wandsworth: California.