

Editorial Introduction

In the scientific visualization community the motivations are primarily technical, related to the types of graphic that technology allows us to generate. Only seldom do users consider the fundamental design aspects of their work and then usually this involves a passing mention of the well known works of Tufte. The result is often a graphic of either terrifying complexity or else one that is a poor representation of the data it is supposed to reveal. Yet almost all of the graphics produced would benefit from some attention to the various principles of map design that have been assembled by map makers (cartographers).

In this Case Study, Dykes and Unwin take what at first sight seems to be a simple graphic, the so-called choropleth map for area aggregated data, and show how different practice in the data to be visualized and symbolism used can lead to totally different maps and, hence, conclusions about the underlying geography. One solution would be to counsel perfection, that there is only one monosemic graphic for any one data set. The other is the approach they adopt, which is to use a series of graphic devices, most notably multiple linked representations and the ability to brush subsets, to present a multiplicity of views of the same data. A critical notion here is that of map stability, it being argued that any patterns that seem to be the same, irrespective of the display techniques used are almost certainly real effects. This seems to us to be an idea that could equally well be used in any visualization work.

Maps of the Census: a rough guide

Jason Dykes & David Unwin

*Department of Geography, University of Leicester
Leicester, LE1 7RH, United Kingdom*

&

*Geography Department, Birkbeck College,
University of London, 7-15 Gresse Street,
London W1P 2LL, United Kingdom*

e-mail: jad7@le.ac.uk
 d.unwin@geog.bbk.ac.uk

Web Page <http://www.midas.ac.uk/argus/>

Abstract

This Case Study describes the considerations that are needed to produce maps of data from the Census of Population. The 'area value' or choropleth map is the standard means of displaying such information on paper. It is a very imperfect visualisation device. First, it is necessary to be careful about the numbers that are mapped and, in particular, never to map absolute numbers. Second, choropleth maps are very sensitive to the mapping zones being used. To produce maps that do not distort the underlying distributions it is necessary to understand how the zones were defined and the effects of their varying sizes on the mapped pattern. Third, there are a series of strictly cartographic considerations related to how these maps are classed and the symbolism used. All of these issues are illustrated using data from the 1991 Population Census for Leicestershire, UK.

These problems lead to a consideration of the need to develop new mapping tools. Dynamic maps can take advantage of an interactive software environment to overcome some of the limitations of the static map. The possibilities which they provide for interactive engagement with data make them appropriate tools for exploratory analysis, or visual thinking. A mapping tool is introduced, which exemplifies this form of map use and examples of the techniques that might be used to visualize the UK Census of Population are provided.

An annotated bibliography is provided to refer map makers and users to relevant literature.

1. Introduction

Visual methods of displaying enumerated data, such as those collected in the UK Census of Population, have long been attractive to those engaged in geographic enquiry. Symbols positioned to represent sample locations can be visually configured in order to display recorded attributes in a map containing both the Gestalt to allow geographic patterns to be interpreted at a regional scale, and the detail for local variations to be detected and data values perceived. In this *Case Study* we address some of the issues involved in mapping such data and describe the development and

use of a tool that uses maps in a highly interactive environment for exploratory data analysis, and resources which promote the approach. In cartography such displays are referred to by the word *choropleth*, in which *choro* means area and *pleth* indicates value.

Choropleth maps are used to visualise data that have been enumerated, or summed, over some defined areas of the Earth's surface. Typically in UK these are derived from the ten yearly Census of Population, and the scientific interest lies in the spatial pattern they exhibit. The role of visualisation in this application is to allow such patterns to be detected and explored in an essentially exploratory approach to



analysis.

Figure 1: A choropleth map of population density for the 187 Census wards of Leicestershire in 1991.

Figure 1 shows a choropleth map of the overall density of population in census wards drawn from 1991 census data for Leicestershire (UK). The stippled areas are zones for which the data are suppressed for confidentiality reasons.

The main elements making up this map are

- The data themselves, in this case the head counts for each of 187 smaller areas (wards). These have been converted into areal densities of population in numbers per unit area.
- Data defining the spatial extents of these small areas. These give limits for each zone on the map.
- To simplify the visualisation, the set of density values has been classified into five categories that span the data range. We have used five shade categories with class limits defined by inspection of the frequency distribution of values so as to maximise the visual effect. Each category is assigned a shade in a graded

sequence from dark grey to white. Conventionally, the darker the shade, the higher the data value.

In effect, what has been created is a two dimensional histogram in which the individual areas have the same role as the *bins* (classes) and the mapped density values are the same as the heights of the histogram *bars*. This type of map is described in virtually every cartography textbook (our favourite is that by Borden Dent, 1985) and has become the standard way of representing area-value data on paper. Many computer cartography programs in particular have been developed to produce such maps on screen and paper. Almost all geographical information systems (GIS) have a choropleth mapping capability. Note that in our figures we have not included any of the standard *base detail*, such as a scale bar, frame and north point, that is usually added to maps for publication. Our intention here is to use the maps that we produce to demonstrate variations in map design, thus they do not follow many of these usual cartographic conventions which would in this context be superfluous. Figure 1 shows that there is a concentration of population in the central city (Leicester) and the other main towns of Loughborough, Hinckley and Melton Mowbray. Although this type of map is a faithful reflection of the data (head counts and zone boundaries) used, it can often be a very poor representation of the underlying patterns it represents

Thinking of it as a two-dimensional histogram helps us see why. In statistics histograms are used to summarise sample data and provide a graphical estimate of some unknown, underlying probability density function (for example, the normal distribution). This function is often added using appropriate parameter estimates, such as when we fit a normal curve to a histogram using the estimated mean and standard deviation. When drawing a histogram the analyst has total control on the size of the classes into which the data are grouped and in most cases these bins are of equal width. Not only this but, since there are just two sources of variation, the value range shown along the x-axis and the frequencies of occurrence shown on the y-axis, the two-dimensions given by a sheet of paper or computer screen are sufficient for the required visualisation.

On a choropleth map the zone boundaries create *spatial bins* over which we have no control and that usually have unequal sizes and shapes. Secondly, we have three sources of variation, the (x, y) co-ordinates that make up the pattern of zones and the value assigned for each zone. The need is to represent three dimensions of variation in two, and this is accomplished by use of a classification of the zone values into categories, each of which is given a grey scale shading. Clearly, the choice of both the categories and the shades used to represent them is arbitrary, but plays an enormous role in determining the look of the finished map. The result is that our map may be as much a result of the zones used and the choice of categories and symbols as it is of the underlying geography. Although we develop these ideas for this specific type of map in his book *How to Lie with Maps*, Mark Monmonier (1991) provides similar examples for virtually any type of map that might be drawn. Early in the days of computer cartography attention was drawn to many of these problems by Baxter (1976) and an excellent recent summary is provided by Kennedy (1994).

2. Issues to consider when area value mapping

In this section we outline and illustrate some of the standard ideas from cartography about good and bad practice in choropleth mapping. These have been

grouped into sections on the numbers mapped, the zones used, the classification into class intervals and the symbolism employed.

2.1 Issues concerning the numbers

Figure 2 shows the same data as were used in Figure 1, but instead of mapping numbers that express the areal density of the population we have simply attempted to visualise the total population.

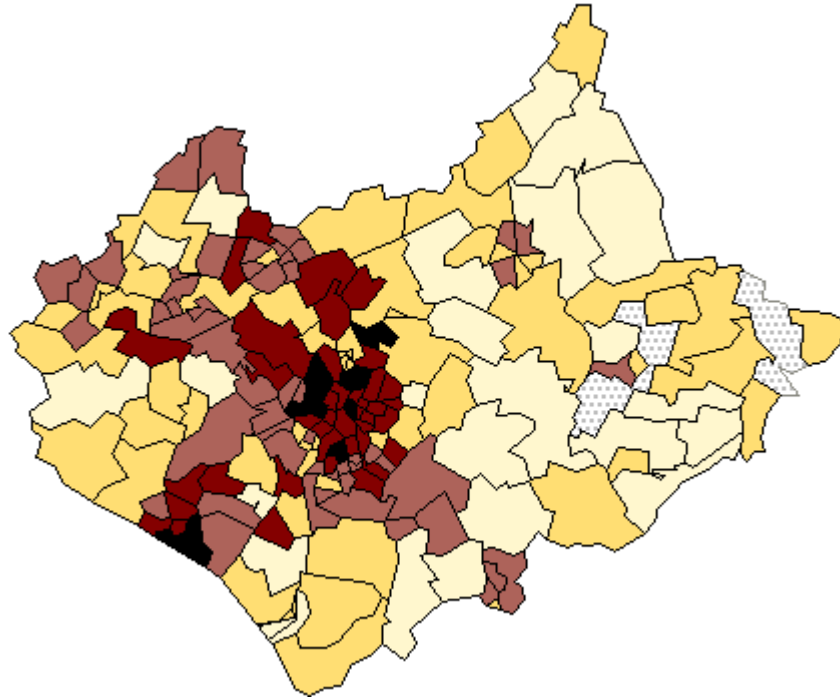


Figure 2: Total population of Leicestershire at the 1991 Census.

In this example the inner city wards that have the highest populations also have small spatial areas, so that the pattern remains similar to that shown in Figure 1. In other cases large zones may tend to have large population totals and *vice versa*, so that the effect of the zones used is that the map tells us very little about the underlying distribution. The distribution of population is better revealed by mapping a ratio, in which the total count is expressed relative either to the area over which it has been aggregated (giving an *areal density*) or to some population total (giving a *population ratio*). Figure 3 shows a population ratio, the percentage of the population aged between 0-15. The transformed variable mapped is now a dimensionless quantity whose value is not dependent on the areas of the zones used, but the map still contains a dependence on the areal units in that physically large areas tend to dominate the display.

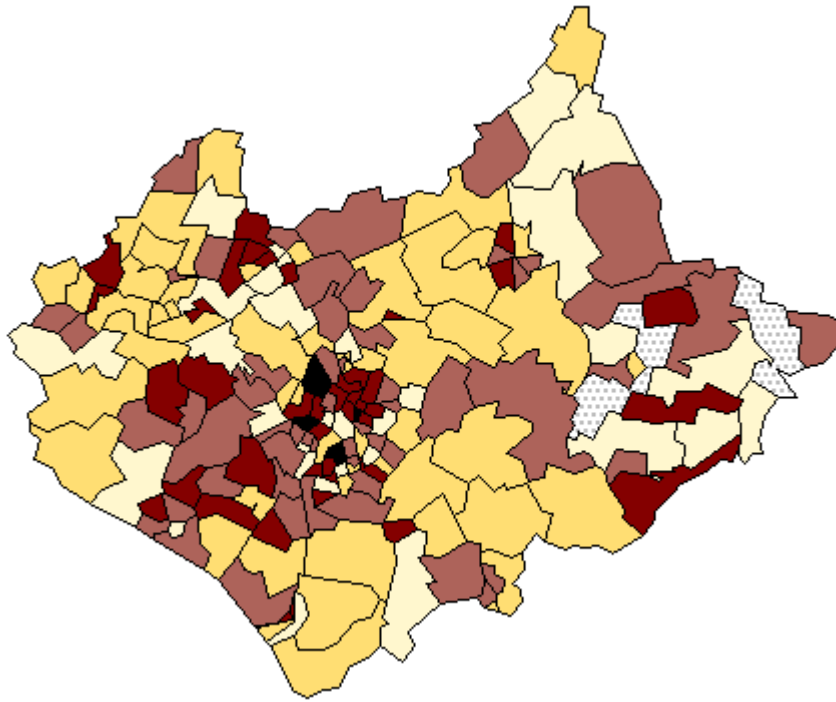


Figure 3: Total children aged 0-15 expressed as a proportion of the total ward populations, Leicestershire 1991.

Depending on the problem being studied, it is possible to refine the numbers used in several ways to provide more revealing maps. Some options to consider are:

1. Standardisation of the denominator in ratios mapped in some problem relevant way. In spatial epidemiology, for example, use is often made of *Standardised Mortality Ratios (SMR)* which are the ratios of the number of deaths in each area to those expected on the basis of some externally specified (typically national) age-sex specific rates. More generally, in a Census Atlas (CRU 1980) produced using the 1971 Population Census of UK, a team based at Durham University mapped a number of variables using as a measure the so-called *signed chi-square statistic*. This was defined as the squares of the differences between the actual numbers in the zone and those expected if the variable in question were uniformly distributed over the entire area divided by this expected total as in a conventional chi-square statistic. Use of the square of the differences necessitated each chi-square value to have its positive (more than expected) and negative (less than expected) signs to be added after the calculation. A simpler alternative is to map the square root of these quantities $(O_i - E_i) / \sqrt{E_i}$ which automatically takes care of the sign. Figure 4 shows the distribution of number of cars in each ward expressed as a chi-square relative to the expected number *per capita*. The calculation clearly draws out attention to greater than expected car ownership in the larger rural areas in a clear commuter belt stretching south-eastwards from Leicester and the smaller inner city wards with less than expected car ownership. The chi square value automatically emphasises the greater significance of variation in zones with more cases.

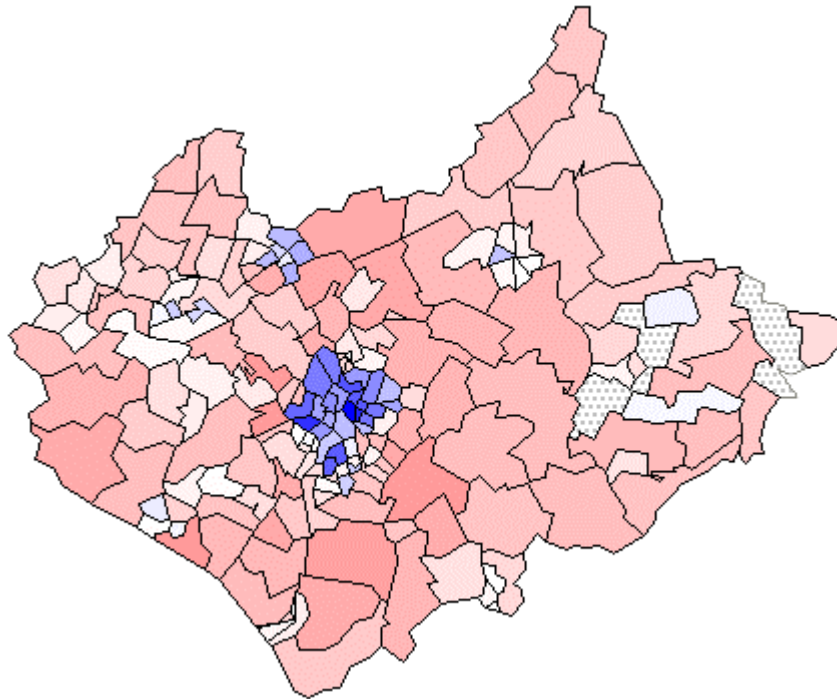


Figure 4: A chi square map. Total numbers of cars relative to the numbers expected under the assumption of uniform distribution throughout the population. The large map shows the full range with low values dark and high values light. The smaller maps use the full range of grey shading to focus on the areas with fewer (top) and more (bottom) cars than expected given their population totals.

2. A common problem in any such mapping is where either the numerator or denominator in the ratio mapped is very small and this often occurs when the cases being mapped are rare or of very high spatial variability. The result is that the ratios become very unstable to small changes in these data. In such circumstances it is sensible to map quantities that are more robust to such changes. Instead of mapping the ratio of observed to expected counts, the *probability of getting a count more extreme than that actually observed* can be mapped. Such a map must be based on some probability model and that usually employed is the Poisson, as is appropriate for an assumption of a random distribution of the individuals. Details of the approach and its assumptions can be found in Bailey and Gatrell (1995, pages 300-302).
3. Another alternative where there are small numbers involved is to adjust the estimated ratios in each zone either away from, or towards an overall global value for the rate according to some prior measure of our confidence in these ratios using concepts from Bayesian statistics. For a very clear introduction to this method see either Marshall (1991) or Langford (1994).
4. A number of recent workers have suggested that in many cases it is appropriate to visualise using what have been termed *local statistics*. Typically, in using local statistics we attempt to learn more about each individual zone by relating it in some way to the values in its neighbours. Several local statistics have been suggested and their use illustrated. For example, Getis and Ord (1992; see also Ord and Getis, 1995) define a G-function which gives an index of spatial clustering of a set of

observations over a defined neighbourhood as $G_i(d) = \sum w_{ij}(d) \cdot x_j / \sum x_j$. Here x is the regional variable and $\mathbf{W}(d)$ is a symmetric 0/1 matrix of weights with 1's for all the areas defined to be within distance, d , of the given area, i . All other elements are zero, including the link of i to itself. Computation of $\mathbf{G}(d)$ requires a suitable data structure from which to determine the $\mathbf{W}(d)$. The vector of values for each region, $\mathbf{G}(d)$, shows how locally anomalous the region is with increasing distance for the given variable, x . Each area has its associated G -function that can be mapped for given d , or plotted as a function of distance. A restriction on this statistic is that as defined it is only useful if the variable has a natural origin. It is thus inappropriate for the study of change variables or variables that have negative values. Second, Anselin (1995) has shown that global coefficients for measuring spatial dependence, or spatial autocorrelation, can be decomposed into local values. These include Moran's index of spatial autocorrelation, I , and Geary's contiguity ratio, C . The local form of Moran's I is the product of the zone value and the average in the surrounding zones, $I_i(d) = z_i \cdot \sum w_{ij}(d) \cdot z_j$ and the local variant of Geary's contiguity ratio is $C_i(d) = \sum w_{ij}(d) \cdot (z_i - z_j)^2$. Almost any of the classical statistics can be calculated as a local value (for example the mean, standard deviation and correlation). This idea has recently been exploited by Fotheringham *et al.* (1996) who compute maps of how the estimated parameters of a regression model vary spatially over a fine grid of cells. Maps of these estimates provide additional information about the spatial stationarity of the phenomenon being mapped. Figure 5 shows a map of the values of each zone's local departure from the neighbouring values (expressed as a local z-score) again using the data for children aged 0-15.

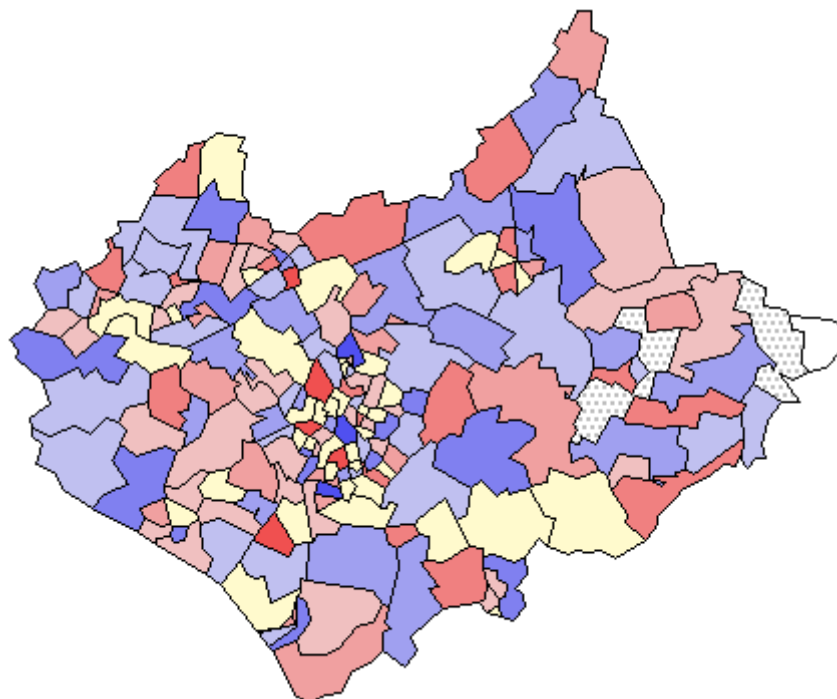


Figure 5: A map of a local statistic. Maps show percentage of the total population aged 0-15. Variation from the local mean is calculated as a z-score. The main map shows local lows and local highs ranging from light to dark. The small maps show positive and negative outliers mapped with darkness to represent magnitude in the top and bottom maps respectively. Localities are defined by first order adjacency.

2.2 Issues concerning the areas

Whatever the numbers used, it should be obvious that the resultant map depends just as much on the pattern of the zones employed in their collection, their shape and areas, as it does upon the phenomenon under investigation. This is one manifestation of the so-called *modifiable areal unit problem* (MAUP) and it is particularly significant in the analysis of census data because the enumeration areas are almost always relatively arbitrary (see Unwin, 1981, 116-119). The MAUP means that virtually any numbers collected over such areas, including the ratios that are displayed on choropleth maps, will depend greatly on the precise boundary definitions used. In the USA most of the state counties are roughly the same spatial area, and this removes some of the worst effects of the problem when we visualise data for aggregations across them. However, the large spatial extent of each state means that there are problems relating to a map projection component in any display of these data as familiar projections do not display equal areas with equivalently sized zones.

In the UK census enumeration districts, which attempt (very unsuccessfully!) to keep a roughly constant population in each unit, vary considerably in areal extent and this can have a very marked effect on the appearance of our maps. Figure 6 is a simple choropleth map of the electoral wards of Leicestershire at the 1981 census shaded such that the shade of grey used in each zone is inversely proportional to its spatial area. It can be seen that there is a pronounced spatial variation in these values, with the zones, or *resolution elements* (resels?) significantly smaller in the more densely populated areas such as the cities of Leicester and Loughborough. In fact, although it contains absolutely no census population count data at all, this map is very similar to Figure 1, our population density map! Any map of census data using zones of equal population will always have a bias towards the representation of the larger spatial units in the less densely populated rural areas rather than the possibly more interesting intricate detail within the cities. Quite apart from any other statistical considerations, all visualisation of area value data needs to be informed by knowledge about how the zones were defined.

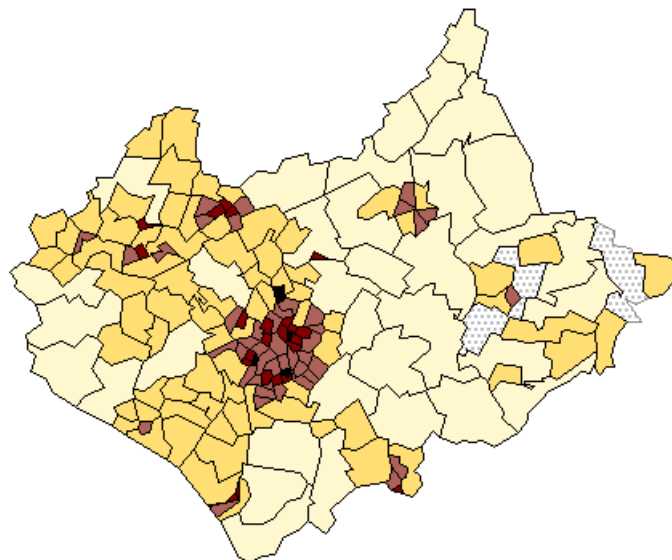


Figure 6: A choropleth map of the electoral wards of Leicestershire at the 1981 census shaded such that the shade of grey used on each zone is inversely proportional to its spatial area.

This area dependence is a fact-of-life in almost all geographical visualisation. To escape it, three radical options have been suggested:

1. The first is to display the census data on what is called a *cartogram* base. As symbol size is frequently used to signify relative attribute importance it seems contradictory to use a map base upon which symbol size equates with geographic area. This is a largely arbitrary attribute appropriate only if focusing on rural processes, using zone shape to identify the geographic location or assessing distances. In population cartograms the zone areas are themselves expanded or contracted according to the total number of people (the denominator in any population ratio). At the cost of creating what can be a very unfamiliar and seemingly distorted 'geography', this magnifies small areas with large populations at the expense of areas that have very few people. Typically densely populated inner city areas are magnified at the expense of more sparsely inhabited rural ones. It can be argued that the resulting cartography is not only better able to show greatly varying spatial detail, it is also in some sense more socially equitable. In recent years, cartograms have been used to great effect by Dan Dorling to visualise many aspects of the social and census geographies of Britain (for examples, see Dorling, 1995) and code (BBC BASIC, and C) published as part of a very full technical monograph (Dorling, 1996). Figure 7 shows an area cartogram based on the population values for the census zones of Leicestershire created using Dorling's 'circle growing' algorithm.

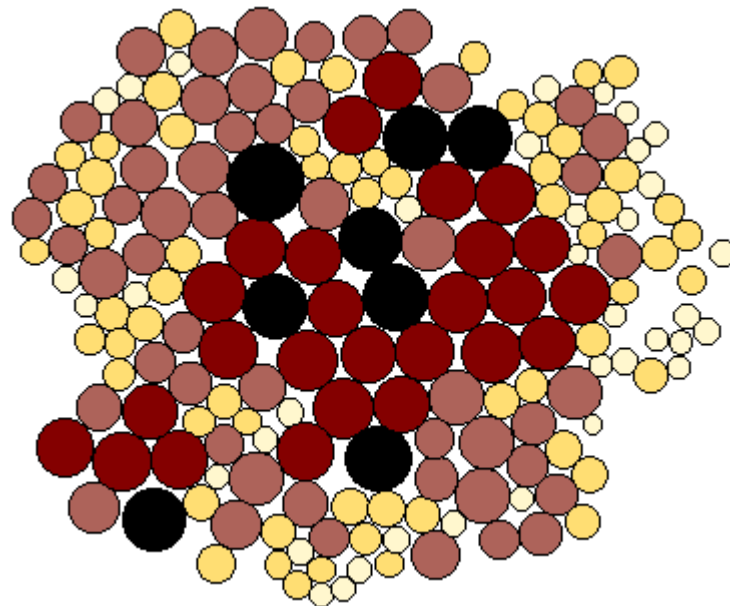


Figure 7: A non-continuous population cartogram based on the population values for the census zones of Leicestershire created using Dorling's 'circle growing' algorithm. Both the circle areas and shades are proportionate to the zone populations.

Colour Figure 4b uses the cartogram base to show the distribution of Figure 4, highlighting the relatively small numbers of people in the identified commuter belt which is over-represented by the land area map due to the large areas which the commuters occupy.

2. The second approach is to transform the census data onto a regular, high spatial resolution grid such as that provided by the Ordnance Survey National Grid, as suggested by Martin (1989). The details are beyond the scope of this case study, but involve interpolation of the census totals from the irregular units onto which they are aggregated by the then Office of Population, Census and Surveys onto a very fine grid mesh of 200m side pixels. These data are readily available for research use by academics in higher education in UK from the MIDAS service at Manchester University. They enable displays of the spatially continuous areal density of population (people per square km) to be created. Figure 8 shows such a surface view of the population of Leicestershire created using these estimates. The view is from the north west, so that we see the towns of Shepshed and Loughborough as 'hills' of relatively high population density 'in front' of the large 'mountain' that is the county town of Leicester.

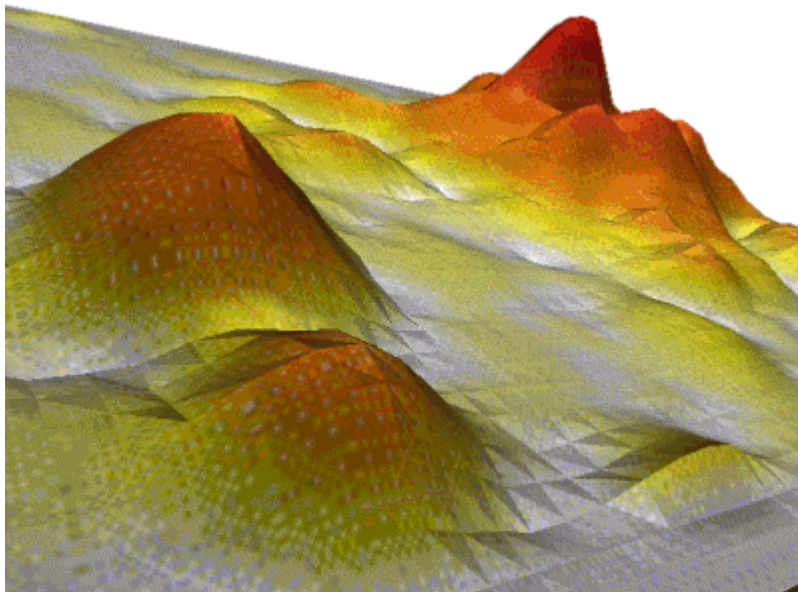


Figure 8: A surface view of the population of Leicestershire created using the Martin 200m population estimates. The view is from the north west. (Visualisation by courtesy of Kath Stynes, Ravenscroft College)

3. A third approach is to use external information to isolate only those areas within the wards that have residential housing and then calculate the areal densities using these areas rather than the total zone areas as denominator. This is called *dasymmetric* mapping and it was originally proposed by Wright (1933). In his original study Wright used inspection of a topographic map as a basis for determining the residential areas, but in a recent study Langford and Unwin (1994) used satellite remote sensing to achieve the same result.

2.3 Issues concerning the symbolism:

Prior to the introduction of computer cartography all choropleth mapping had an additional step between the assembly of the data values and their visualisation using symbols on the map. This was to group values into a limited number of discrete classes. Each class was then assigned a shading category on the resulting choropleth map. Not only did this greatly simplify the labour involved in drawing these maps, it was also justified on the grounds that the map reader could not readily distinguish

between more than a limited number of shade categories (Dobson, M, 1973). It also overcomes the problem of simultaneous contrast, where the appearance of grey values is influenced by surrounding shades. Classification provided scope for enormous variation in the resulting displays given by the choice of both the number of classes and the class limits, or bins. Most texts in cartography have extended discussions on these topics (see for example Dent, 1990, pages 153-178) and it is relatively easy to show that different choices can lead to very different displays. Figures 9 and 10 show the data for children aged 0-15 using two differing, but traditional classification schemes. It can be seen that the *message* they give might also differ. Figure 9 has a five class scheme using equal value intervals to define the bins, whereas Figure 10 has the same number of classes but an attempt has been made to get an equal number of cases into each class.

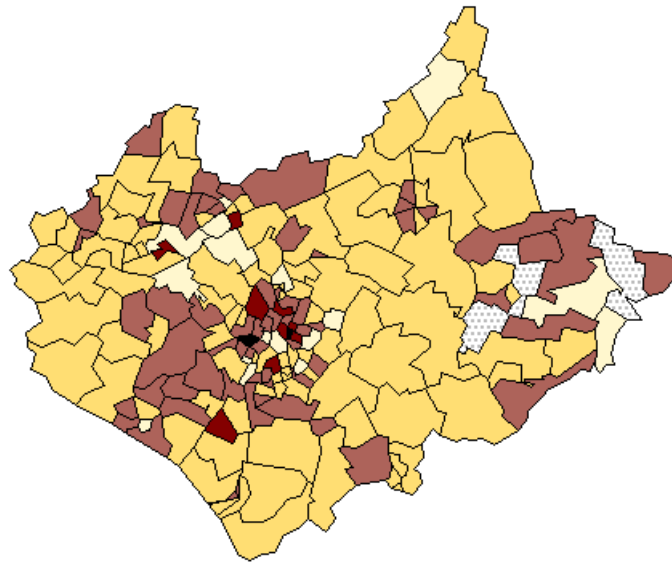


Figure 9: A choropleth map of the proportion of children aged 0-15, using five classes based upon an equal bin interval classification scheme.

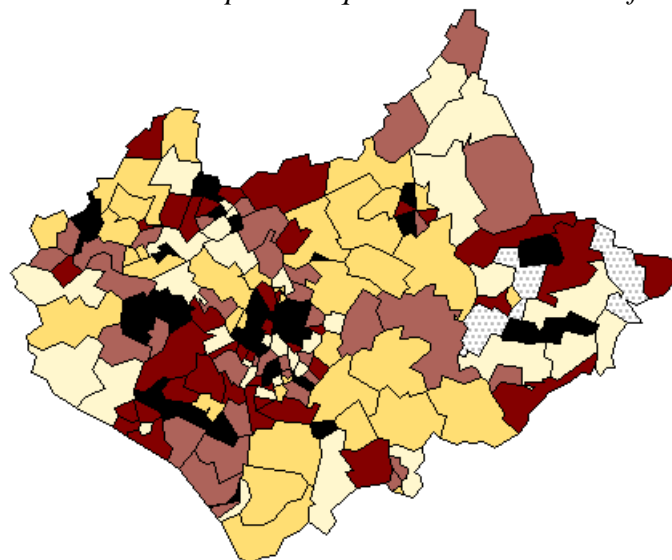


Figure 10: The same data visualised using the same number of classes but a different classification scheme. Here equal numbers of wards are classified in each bin.

Most authorities agreed that if a choropleth map is to have classes, then it is best to use a limited number, typically 5 or 7, but the choice of classification scheme is by no means simple. In a classic paper on the problem, Evans (1977) lists over twenty possible alternatives and his basic recommendation is that the selection of scheme should always be based upon an inspection of the overall statistical frequency distribution of the zone values.

Choropleth maps produced with computers allow very gradual and subtle variations in colour. In 1973, the American geographer Waldo Tobler suggested that it ought to be possible to create shade patterns that would vary continuously in relationship to the variable being mapped (Tobler, 1973). Tobler's software did this by varying the spacing of the lines in a shading pattern. Others experimented with FAX technology to provide a 255 shade grey scale (Muller and Honsaker, 1978). Modern PC screens support 16 million colours, with 256 shades of grey and so effectively unclassed maps are easily produced. Figure 11 shows such a display using exactly the same data as were used in Figures 9 and 10 but filling each zone with a grey shade in exact proportion to its value. That it is visually very similar to the five-class equal interval scheme used to create Figure 9 should be obvious, and there is a literature on map user's perceptions of such continuous choropleth shade maps (see Muller, 1979; Petersen, 1979).

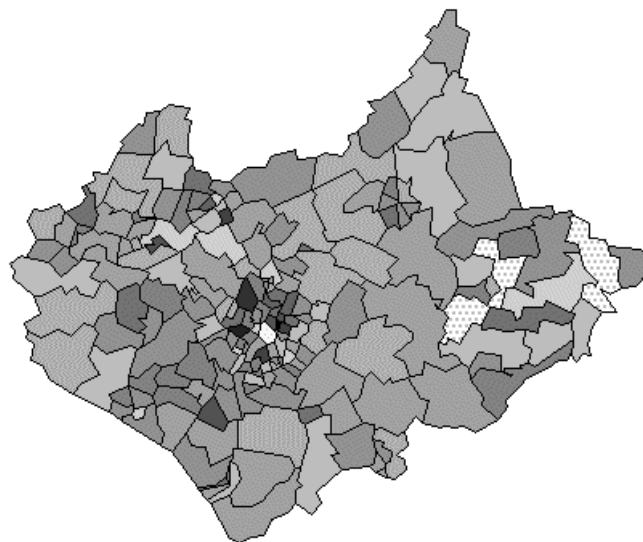


Figure 11: A 'classless' choropleth of these same data using a continuous grey scale.

A final problem in conventional census mapping is the symbolism used. Most of the illustrative figures we have used have a grey scale applied uniformly over the zones mapped with the darker shades to represent higher values. It should be clear that in an ideal world this symbolism will be 'read' as a linear sequence in which 'twice' as dark would indicate a doubling of value and so on. The psychological evidence is that this is difficult to achieve. An alternative is to abandon shades of grey as indicators of value and instead use symbol size. Figure 12 shows the same data as Figures 10 and 11, but we have used a *located proportionate symbol* to signify value. The symbolism is identical to that used in the non-continuous cartogram (Figure 7) but locational information is retained at the expense of symbol overlap.

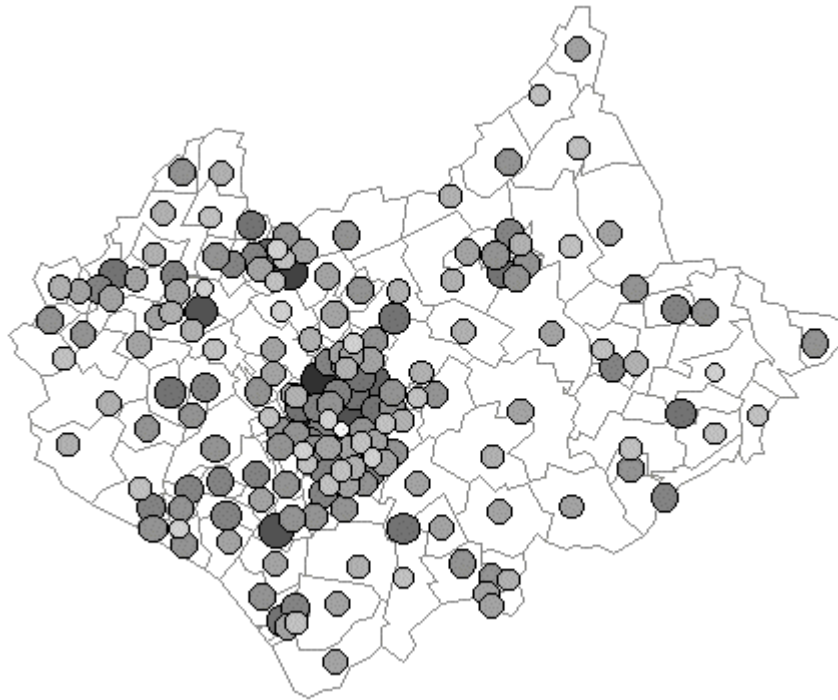


Figure 12: A graduated circle map of the same information. The circle areas and shading are proportionate to the zone values and the circles are located at an assumed centroid within these areas.

This has two major advantages over the shaded area maps. First, the location of each circle at some assumed central location within the area for which the data have been aggregated makes the area dependent nature of such data extremely clear. Second, it is possible to shade the centre of each symbol by a grey shade or colour to visualise an additional variable. The disadvantages are that the perception problem remains and again there is a considerable literature on map user's perceptions of this type of symbolism (see Crawford, 1973; Flannery, 1977). It has been shown that by and large users tend to underestimate the true values represented by larger circles. It has also been shown that the design of the legend can affect this (Dobson, 1974). Figure 12 again demonstrates the trade-off between accurately depicting both location and value simultaneously. Addressing this balance is the crux of area-value mapping.

Faced with these problems and sat at a workstation capable of producing many millions of colours, the temptation might be to try to use a graded series of colours in choropleth mapping. Our straightforward advice is that you use a scheme that has a natural progression of hues and lightness as recommended by Brewer (1994), such as the scheme we have utilised in the colour versions of these pages. As was shown many years ago by Jacques Bertin in his classic book *Semiologie Graphique* (*Semiology of Graphics*, English Translation of 1983) colour is best used to show qualitative differences, not the quantitative intensity that we need when mapping census data, and the main reason for this is that what we see as 'colour' has more than one dimension of variation. Although we might specify a colour using the RGB system, what appears is generally regarded as having three components of variation, the *hue* (the colour itself) its *saturation* (its intensity or brilliance) and its *value* (lightness or darkness). Not only this, but different hues produce differing responses in the eye brain system. Some, such as yellow, are dominant, others, such as blue, are recessive. Add to this the fact that everyday use gives certain colours special significance (such as red for danger, at least in western societies) and you have a

recipe for what in Monmonier's words is 'a cartographic quagmire' (Monmonier, 1991, page 147). This is not to say that colour isn't useful in choroplethic mapping. It is particularly useful in showing a bi-polar scale with positive and negative values as, for example, might result from a calculation of a percentage change in population. Figures 4 and 5 demonstrate the difficulty of representing a bi-polar distribution in monochrome and a coloured alternative. If you really feel that it is necessary to colour census maps of single variables, then it is vital that you first read some of the literature in the field (Mersey, 1990; Monmonier, 1991, 147-156; Brewer, 1994).

3. Dynamic mapping tools and cartographic data visualization

This section will stress highly interactive visualisation as a way to use the issues noted in (2) above to the map user's advantage. Current technology permits rapid map production, and so single, or more often multiple, transient maps can be developed as requested to highlight very specific user-led queries. Graphics are now available instantly at the press of a button, and they embrace new dynamic properties that can add enormously to their functionality and fitness for use. In consequence, a trend has developed away from a single optimum means of representation towards multiple alternatives, where symbols appear in different ways at different times to stress particular messages. This has had an impact on cartography. Rather than being restrictive shackles that must be searched through for an aesthetic, multipurpose, or optimal solution, map design parameters can be altered in real-time to assess confidence in patterns through the extent to which they are influenced by changes in the method of representation. Muehrcke (1990) referred to this property as *map stability*. Map users can focus their attention on different aspects of a spatial data set at different times with maps that vary as required. Local variations and minor fluctuations can be afforded the visual power of the full range of symbolism available, if and when requested.

Maps can thus be used an integral part of the research process, as dynamic visual interfaces to a collection of spatial information which respond when they are interrogated for information. This shift in map use was recognized by DiBiase (1990) who identified private map use as distinct from that occurring in the public realm and termed the cyclical map-based process of visual exploration and synthesis of ideas *visual thinking*. It corresponds with a use of graphical tools early in the research process across the sciences termed *visualization* and acknowledged by McCormick, DeFanti & Brown (1987). The use of dynamic maps for visual thinking is called *cartographic visualization*.

Early examples of transient private maps which embraced dynamic techniques were provided by Ferreira and Wiggins (1990) whose *density dial* allowed users to slice through a variable's range at a level determined by a dial and view the resultant classification on a map. Such a dial enables users to determine the sensitivity of the classification, and by varying the scheme they can determine the level at which a variety of patterns occur.

Such transient symbolism can be used to simply highlight symbols in a nominal manner when they are selected for investigation. Stuetzle (1988) was an early pioneer of what is called *brushing*, where corresponding symbols in two or more views are identified by identical distinct symbolism. Monmonier (1989) extended the technique to link choropleths with statistical plots, adding a geographic component to scatter plots, and vice versa. Just as transient maps eliminate the constraints associated with producing single representations, linking effectively removes the restriction of having

to use the spatial dimensions of the page to show geographical locations. Maps can use the spatial arrangement of symbols on a page to reflect geographic relationships whilst at the same time using locations on linked views to show statistical variation.

Dynamic maps enable the focus of interest to be varied spatially by changing scale, or location, with zoom and pan functionality, and statistically by selecting subsets from statistical views of data. For example, the problem of simultaneous contrast can be overcome by selecting a data sub-set to extend the variation in symbolism and by highlighting symbols on a statistical dot plot which uses location to display the same values.

In order to publicise and promote these techniques and changes in map use a JISC New Technologies Initiative Project, ARGUS, produced software, documentation and data sets for cartographic visualization as demonstrators and in research and teaching. The resources developed constitute a *Visualization Toolkit for Teaching and Learning in the Spatial Sciences* (see <http://midas.ac.uk/argus>) which consists of:

- A series of tutorials on cartographic techniques
- Example data for a specified region in the public domain
- Software for visualizing two-dimensional data sets
- Software for visualizing three-dimensional data sets
- Research papers on methods and techniques

Two-dimensional census data sets were visualized with a Cartographic Data Visualizer, *cdv*, specifically developed by the project to demonstrate a variety of new methods and opportunities for interacting with enumerated data, such as those outlined above. The following section introduces some of these techniques.

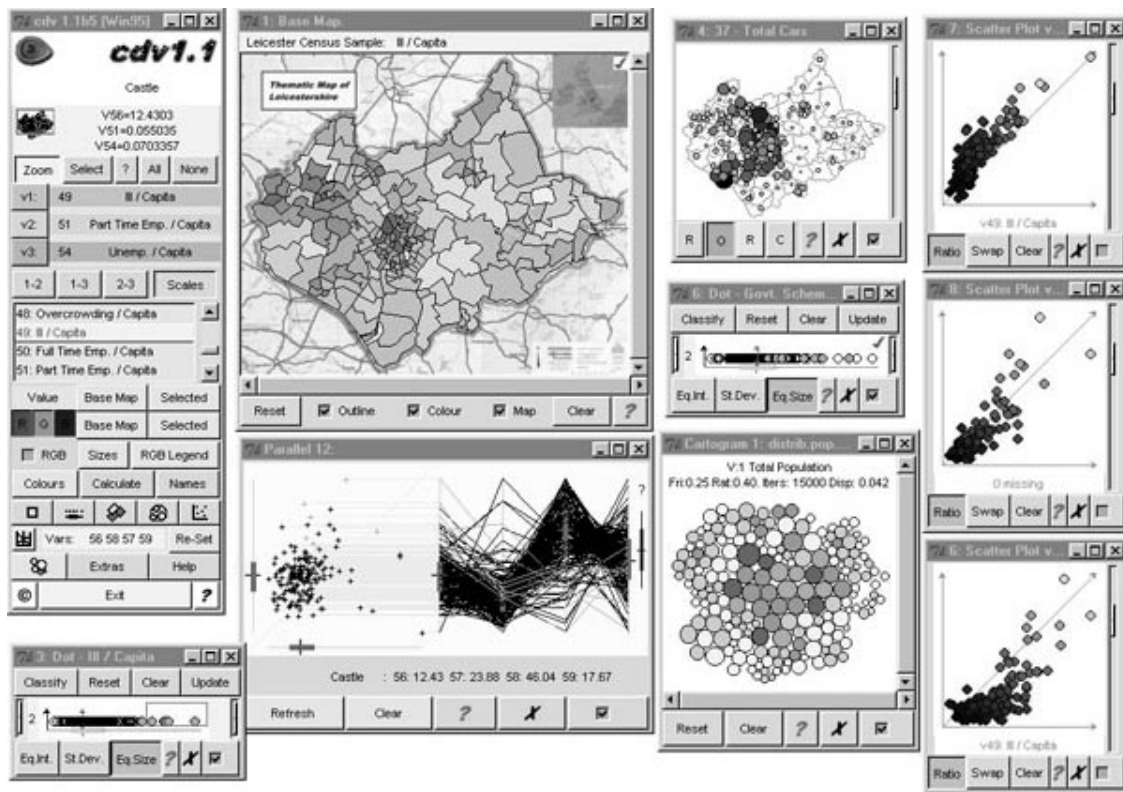
4. The use of *cdv* to visualise the Population Census.

The notion behind *cdv* was that data exploration required the kinds of speed of response common to most desktop point and click graphical user interfaces (GUIs) and expected by most computer users, abstract graphics that were suitable to help identify patterns, and the flexibility to create novel and particular views of data sets. Despite considerable advances, software dealing with geographic information still tends to produce graphics that mirror traditional production cartography with slow, intricate maps, and an emphasis on realism and beautiful reproduction rather than symbolism. It was envisaged that a graphics front-end to a spatial data set could take advantage of the GUI features without being slowed down by GIS baggage.

Tcl/Tk (Ousterhout, 1994) is a GUI-builder particularly appropriate for rapid development and which contains a *canvas* widget on which symbols can be plotted at specified locations. The types of symbolism that can be programmed compare well with the elements identified by Bertin (1983) and others, and give excellent scope for creating abstract views of spatial data (see Dykes, 1997a). The capabilities for changing symbolism and updating views by linking events to cursor/mouse combinations are varied and powerful, meaning that Tcl/Tk is extremely suitable for dynamic mapping for exploratory map use (Dykes, 1996, 1997b).

Cdv takes two ASCII files, one containing a series of unique identifiers followed by enclosing co-ordinates, and another containing rows beginning with an identifier and followed by ratio-scaled attribute fields. The range of each attribute is calculated and a polygon plotted using the enclosing co-ordinates. The polygon is filled with a grey shade from 100% intensity (white) to 0% (black) based upon the data value within the range. The result is a classless choropleth map of the type shown as Figure 11. Polygons can be interrogated for information about their location and

value and a point and click GUI allows users to select data elements and a variety of



views of the data.

Figure 13: A look at cdv - Linked geographic and statistical views. Here long-term illness / capita is shown on a classless choropleth and population cartogram. Those on government employment schemes are shown by a dot plot and scatter plots show the relationships between three age categories. A dot plot showing illness/capita has been used to identify locations of the extreme values in the other linked views.

As Figure 13 shows, these include dot plots, box plots, scatter plots, circle maps, and parallel co-ordinate plots. In a cdv session, the emphasis is on transience and the user is encouraged whenever possible to adapt the cartography used to highlight specific information. Re-scaling, zooming and panning of views, and user-defined data classification, colour schemes and symbol scaling are supported. All views are dynamically linked, meaning that when a symbol is interrogated with the cursor, or when a group of symbols are selected with a lasso, symbols corresponding to those chosen in one view are highlighted in all others. This means of alleviating the traditional cartographic enigma of showing three dimensions of variation on a two-dimensional plane is illustrated in Figure 14.

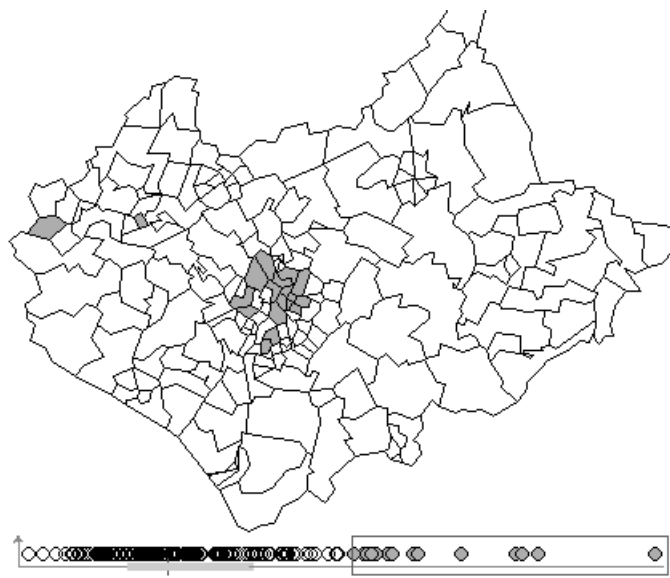


Figure 14: Long-term illness / capita. A dot plot uses locations to show the range of data values, with a box plot identifying the median and quartiles. Extreme values are selected interactively with a lasso and highlighted on the map view.

Data sub-sets can be selected, either spatially to focus attention on particular geographic areas, or by attribute values to re-scale the data range for more precise mappings between symbolism and value as shown in Figure 15.



Figure 15: Long-term illness / capita. Here the extreme values highlighted in Figure 14 have been selected as a data sub-set and new views created to show the variation within that data range. The map has been interactively re-scaled and panned to show the area where most of the extreme values occur.

The vast opportunities for symbolism in Tcl/Tk and interpreted nature of the scripts add a degree of flexibility for further prototyping. Scripts can be extended to suit particular data sets or specify additional functionality to views. The provision of a series of tutorials and examples (URLs 3 & 4) supports this notion of an open visualization resource or toolkit of ideas. Cdv also includes some unsupported prototype extensions. These include procedures to create contiguity matrices and visualize local statistics based upon adjacency neighbourhoods, to create and interrogate semi-variogram clouds, to display local parallel co-ordinates plots based

upon distance inclusion of zones selected along a user-digitised route and the addition of raster images as backdrops.

5. Alternatives to choropleth maps: new interactive visualization tools

Pre-computed non-continuous cartograms can be loaded into cdv. Those distributed with the example data sets use Dorling's (1996) algorithm. Dynamic links between abstract spatial views and the familiar polygons with their geographic locations and distinctive shapes, as illustrated in Figure 16, overcome some of the issues of interpretation and retaining zone shape and geographic distances inherent in the use of static cartogram projections. They can also be used to resolve some of the problems with mapping population information in static choropleths as outlined in section 2, including the projection problem.

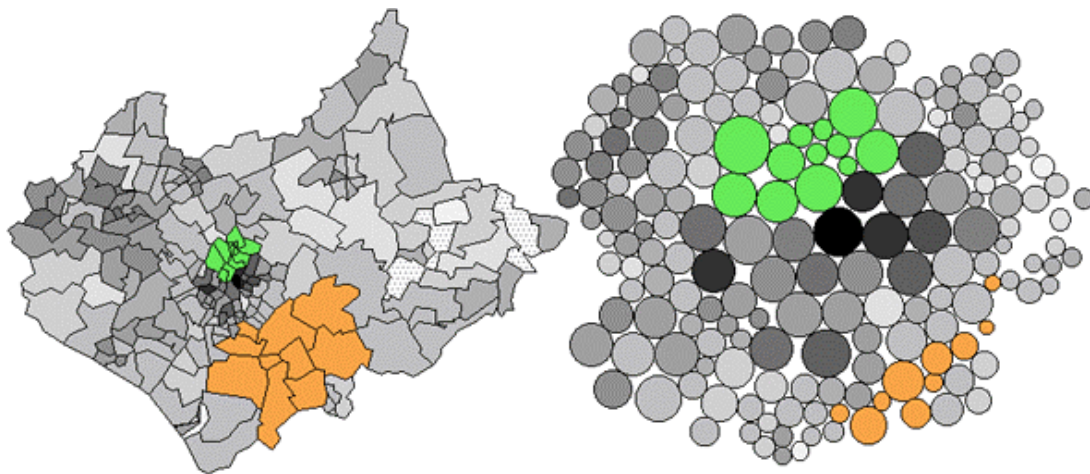


Figure 16: Linking between the views which display geographic boundaries and those containing non-continuous population cartograms aid interpretation of both views.

Such linking techniques significantly diminish the justifications for using land-area maps for symbolising data about the population. Note the astonishing over-representation of the population in rural zones in the traditional land-area based projection. As Figure 17 shows, using cartograms can produce startling results. See, for example, how a traditional two-class choropleth showing car ownership / capita in Leicestershire reveals that the majority of the county contains people with above average access to private transport. The cartogram clearly demonstrates that the majority of the population live in areas where there is below average access to private transport.

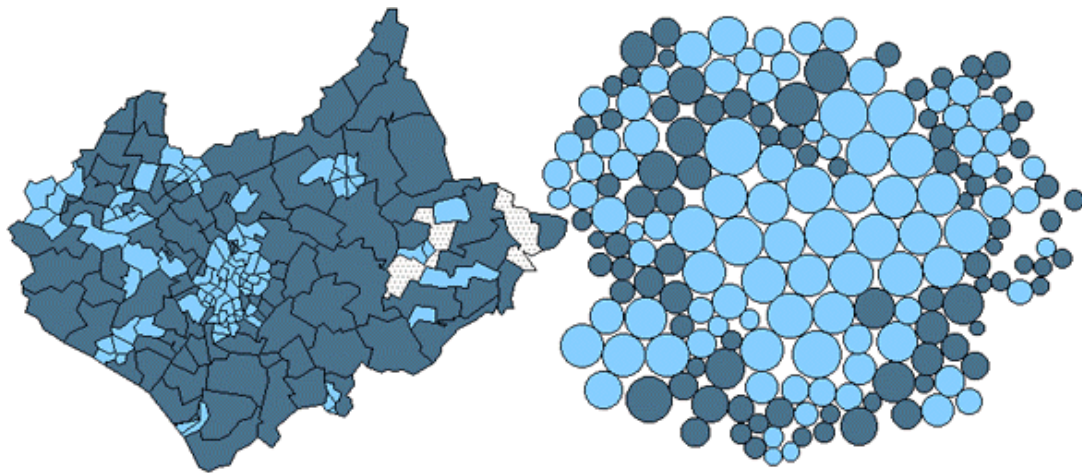


Figure 17: Leicestershire wards with above and below median values for car ownership per capita shown in dark and light colours respectively.

Multivariate distributions can also be visualized with cdv. Research has suggested that certain types of map user respond well to combinations of symbolism for displaying geographic distributions of multiple variables. CdV supports colour composites that define colours by their red, green and blue components. Different attributes can be assigned to each of the three primary hues and the range of data values attributed a strength of hue from 0% (minimum) to 100% (maximum). By combining red, green and blue percentages for each attribute, a colour is defined that is appropriate for the combination of variable values measured at each zone. The dynamic qualities of the software aid interpretation of these complex combinations of colour through a GUI that allows individual colour components to be turned on and off as required. Colour components can be used to symbolise any view. Three linked dot plots can thus be symbolised with component colours and combined with a composite map. Location is used to identify the statistical variation within variables in the dot plots, colour and linking between plots the multivariate nature of the each measurement zone and the coloured map the spatial variation of the three variables (see colour Figure 18).

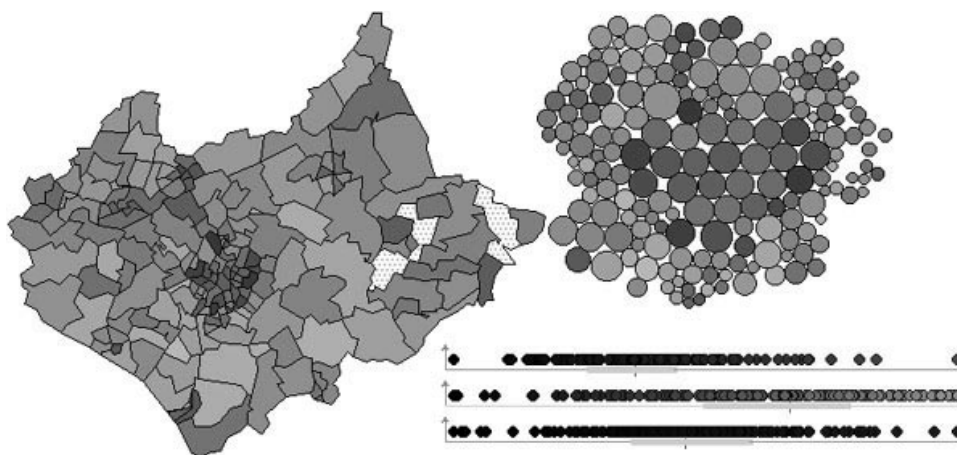


Figure 18 (Colour version only): A colour composite showing proportions of the population aged between 0 and 15, 25 and 5, and over 60 in red, green and blue respectively. A population cartogram is shown alongside the geographic projection. Dot plots display the statistical distribution of each component and the strength of hue contributed for each variable by each zone.

A tool that allows more than three variables to be assessed in combination, by using location rather than colour, is the parallel co-ordinates plot (Wegman, 1990). Here the premise is that by using locations to show data values on orthogonal axes, as is the case with scatter plots, the number of variables that can be mapped in a plot is restricted to two, with the possibility of adding a third if a pseudo-3D projection is employed. By plotting data along parallel axes this number is increased dramatically and the variation of cases between variables can be assessed more precisely than in the 3D alternative. Effectively such plots are parallel dot plots with links between cases formed by lines rather than dynamic highlighting (Figure 19) and they allow more than three variables to be assessed concurrently. The distribution of lines, the patterns that they form, and envelopes that result, highlight outliers, multi-modal distributions, clusters, and the nature of relationships in multidimensional data space. In cdv, parallel coordinate plots link to the map and other views. There are $n!$ (where n is the number of axes) possible views through this multidimensional data space. The initial ordering of the axes chosen by cdv minimises Pearson's product moment coefficient between adjacent axes but dragging axes in real time will re-order them. Figure 19 shows how pairs of axes can be interrogated to display both the correlation coefficient and a scatter plot showing the relationship between any pair of variables.

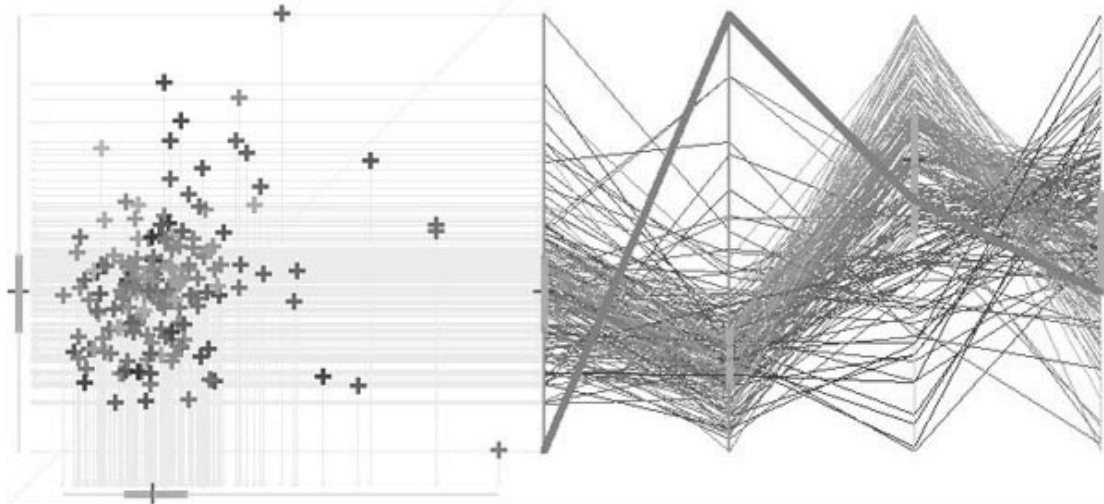


Figure 19: A parallel coordinates plot showing the three age groups used in Figure 18, along with those aged between 16 and 24. The axes are ordered by age from left to right, but can be moved interactively. The scatter plot shows the relationship between selected axes on more familiar orthogonal axes. The coloured version retains the colour components used in Figure 18.

6. Conclusion

Although laid out as a 'how to do it', *Rough Guide* to mapping census data followed by an introduction to a specific software system, this *Case Study* has in fact attempted to develop an important argument about traditional cartography and visualization. Our basic argument is that dynamic mapping provides the potential for engagement with data and exploring information in a way that pre digital mapmakers and users could only have dreamed about. It has been said that computer software and the availability of digital spatial data have given everyone the equivalent of a *cartographer on the desktop*. If this were all that it had done, then the issues raised in Section 2 of this *Case Study* would be most important. Mapping software finds use as a means of reducing the labour involved in drawing maps of census variables, but the

traditional attitude would have been to use it to search for some optimum, single map representation of the data. Such maps clearly have a role in *communicating* to others a research result, and producing and 'reading' such maps still requires careful consideration of the issues we outlined in that section.

However, this isn't all that the *cartographer on the desktop* has allowed. As Sections 3 and 4 have demonstrated, some of the traditional cartographic problems can be overcome by using dynamism, linking, and the careful use of colour. The techniques can be used to search for the 'best' way of displaying a distribution to demonstrate a known research result, but, as we have illustrated in Section 4, they have even greater utility in a visual approach to spatial data exploration.

7. Future Directions & Technical Considerations

Collaboration and dissemination were important aspects of Project Argus. Working relationships were developed with a variety of organisations to share ideas and data. Links with the KINDS Project at Manchester Computing (URL 5) led to the notion of combining interactive graphics for data exploration with the distribution of census data from a National data server. The potential for an Internet based visual interface to the UK Census, or Visualization Gateway, is currently under investigation through Project JANUS at Manchester Computing (URL 6). The project has three main objectives towards which work is progressing:

1. Transferring the key deliverables of the ARGUS and KINDS projects into an integrated service and hosting them on MIDAS.
2. Educating the community by promoting, using and supporting these products.
3. Developing software that will enable users to perform exploratory spatial data analysis of census data effectively forming a visualization gateway to the UK Census of Population over the Internet.

The ARGUS resources have been successfully housed on the MIDAS machine. The *cdv* software has been used in workshops demonstrating exploratory visualization of census data to academics, researchers, statisticians, and civil servants in the UK and Europe. It has also been used with locally appropriate data in undergraduate and postgraduate geography programs at the Universities of Leeds, Keele & Leicester. A series of generalised polygons and population cartogram bases is being developed for enumerated zones, specifically for visualization with *cdv* or similar products.

Technological considerations relating to the final stage are currently being assessed. An attractive option would be a Java application combining the 2D and 3D census mapping. Java provides much of the functionality of Tcl/Tk, and is quicker, but at the expense of the high-level flexibility and hands-on nature of the scripting language. It is, however, a language for Internet applications and couples neatly with the VRML script node meaning that interactive querying and linking can be applied to the three dimensional representations in a similar way to the 2D interaction provided by *cdv* (Moore *et al.*, 1997). The release of the Java 3D API in the spring of 1998, with its 3D graphics, will extend possibilities further. Whilst *cdv* was a demonstrator, for which a hands on, flexible, scripted approach was ideal, the visualization gateway will be an application, for which the speed, security and robustness of a compiled application are more appropriate. These considerations are influenced by developments initiated by Sun Microsystems to integrate Java and Tcl applications. Jacl and Tcl-Blend enable Tcl commands to be written in Java and Java programs to

call Tcl procedures. These capabilities are extremely recent, and at the time of writing their utility for adding processing power to spatial Tcl applications remain untested. They might provide the means for combining the kinds of 3D graphics required to visualize a population surface with the kinds of dynamic graphics provided by *cdv* in two dimensions. These developments, along with the resolution of some of the Internet security issues that both languages currently address by omitting functionality, would enable an integrated visualization gateway to be developed for census data using a combination of the two languages. Both authors are currently engaged in a project that uses Tcl/Tk to add a map interface to panoramic photographs and VRML/Java to display three-dimensional data sets. Combination of the approaches is under investigation.

8. Acknowledgement

The figures used to illustrate this Case Study are based on Census data and zone boundary data for the 1981 UK Census of Population for the County of Leicestershire. These data are Crown copyright. Permission to display and distribute these, given to the JISC/NTI Project ARGUS based at Leicester University and Birkbeck College London, from the Office of National Statistics and the ED-LINE Consortium is gratefully acknowledged.

9. Select Bibliography

- Anselin, L. (1995) Local indicators of spatial association - LISA. *Geographical Analysis*, 27: 93-115.
A key paper for anyone interested in local indicators.
- Bailey, T.C. and A.C. Gatrell (1995) *Interactive Spatial Data Analysis*. Harlow: Longman.
Not primarily about visualisation, but an excellent introduction to the statistical methods that are nowadays applied to spatial data.
- Baxter, R.S. (1974) Some methodological issues in computer drawn maps. *The Cartographic Journal*, 13: 145-155.
An early paper that covers much of the ground covered in this Case Study.
- Bertin, J. (1983) *Semiology of Graphics*. Madison: University of Wisconsin Press.
- Borden D. Dent (1985) *Cartography: Thematic Map Design*. Dubuque IA: Wm.C.Brown:
Of the very large number of basic cartography texts, this is highly recommended. Chapter 6 provides a detailed account of the problems in choropleth mapping.
- Brewer C.A., 1994, Colour use and guidelines for mapping and visualization. In MacEachren, A.M., & Taylor, D. R. F. (eds) *Visualization in Modern Cartography*, Oxford: Pergamon, p. 123-147.
A very careful analysis of colour use on maps and graphics, highly recommended.
- Census Research Unit (1980) *People in Britain: a census Atlas*. London: HMSO
This was a very innovative product for its time, making use of the signed chi-square to show area value relative to national expectations and what at the time was very new high resolution laser colour printing technology. Such technology is now commonplace.
- Crawford, P.V. (1973) The perception of graduated squares as cartographic symbols. *The Cartographic Journal*, 10: 85-88.

This is typical of the literature dealing with how map users 'read' symbol types.

- Dobson, M.W. (1973) Choropleth maps without class intervals?: a comment. *Geographical Analysis*, 5: 358-360.
Presents the view that choropleth maps should classify their data into bins.
- Dobson, M.W. (1974) Refining legend values for proportional circle maps. *Canadian Cartographer*, 11: 45-53.
- Dorling, D. (1995) *A New Social Atlas of Britain*. London: John Wiley and Sons.
An extended essay in cartographic visualisation making very effective use of a population area cartogram base to map various social indicators. The philosophy and some code for this are both detailed in:
- Dorling, D. (1996) Area Cartograms: Their Use and Creation. *Concepts and Techniques in Modern Geography*, 59, 69 pages. Norwich: Geobooks.
- Dykes, J.A. (1996) Dynamic maps for spatial science: A unified approach to cartographic visualization. In Parker, D. (ed) *Innovations in GIS 3*, London: Taylor & Francis, London, 177-187.
An introduction to Tcl/Tk for dynamic cartography.
- Dykes, J.A. (1997a) Exploring spatial data representation with dynamic graphics, *Computers and Geosciences*, 23: 345-370
A WWW based paper addressing dynamic means of representation, with interactive examples of dynamic cartography.
- Dykes, J.A. (1997b) cdv: a flexible approach to ESDA with free demonstration software. Proceedings, British Cartographic Society, 34th Annual Symposium, Leicester, 100-107
Includes a WWW-based interactive example demonstrating Bertin's visual variables in Tcl/Tk.
- Evans, I.S. (1976) The selection of class intervals. *Transactions of the Institute of British Geographers*, 2: 98-124.
An absolutely classic paper that should be compulsory reading for all would-be choropleth mappers. Evans draw attention to the need to make the classification scheme used suit the statistical, possibly also the spatial, distribution of the data being mapped.
- Ferreira, J., Jr., and Wiggins, L.L., 1990, The density dial: A visualization tool for thematic mapping. *Geo Info Systems*, 10: 69-71.
An example of the use of interaction better to design choropleth class intervals.
- Flannery, J.J. (1977) The effectiveness of some common graduated point symbols in the presentation of quantitative data. *Canadian Cartographer*, 8: 96-109.
In order to correct for the perceptual underestimation of the size of large circles, 'Flannery's law' suggests an adjustment to the relationship between the area of a circle symbol and the value it represents.
- Fotheringham, A.S., Charlton, M.E. and Brunson, C. (1996) The geography of parameter space: an investigation into spatial non-stationarity. *International Journal of Geographical Information Systems*, 10: 605 - 627.
- Getis, A. and Ord J.K. (1992) The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24: 189 - 206.
Both the above deal with extensions to the LISA idea.
- Kennedy, S. (1994) Unclassed choropleth maps revisited: some guidelines for the construction of unclassified and classified choropleth map. *Cartographica*, 31: 15-25.

- Langford, I (1994) Using empirical Bayes estimates in the geographical analysis of disease risk. *Area*, 26: 142-190.
A clear, how to do it, account of using Bayes' theorem to modify mapped values according to how well they are known.
- Langford, M and Unwin, D.J. (1992) Generating and mapping population density surfaces within a geographical information system. *The Cartographic Journal*, 31: 21-26.
Shows how careful use of ancillary information can enhance choropleth mapping by allowing dasymetric mapping.
- MacEachren, A. M., 1995, *How maps work: representation, visualization and design*. New York: The Guildford Press.
- McCormick, B.H., DeFanti, T.A., and Brown, M.D. (eds) (1987) *Visualization in Scientific Computing*, Computer Graphics, 21(6), special issue.
- Marshall, R.J. (1991) Mapping disease and mortality rates using empirical Bayes estimators. *Applied Statistics*, 40: 283-294.
- Martin, D. (1989) Mapping population data from zone centroid locations. *Transactions, Institute of British Geographers*, 14: 90-97
- Mersey, J.E., 1990, Colour and thematic map design: the role of colour scheme and map complexity in choropleth map communication. *Cartographica*, 27, Monograph 41, 157 pages.
- Monmonier, M. (1989) Geographic brushing: enhancing exploratory analysis of the scatterplot matrix. *Geographical Analysis*, 21: 81-84.
Early examples of geographic brushing where polygons and scatter plots are linked.
- Monmonier, M. (1991) *How to Lie with Maps*. Chicago & London: Chicago University Press
Based on Huff's well-known book How to Lie with Statistics, this shows how cartography can be used to create false impressions. Chapter 10, Color: Attraction and Distraction should be compulsory reading for those new to the field.
- Moore K.M, Wood, J.D. & Dykes, J.A., 1997, Using java to interact with geo-referenced VRML within a virtual field course, ICA Visualization Commission, Gavle, Sweden, June 1997,
<http://www.geog.le.ac.uk/mek/usingjava.html>
- Muehrcke, P. (1990) Cartography and geographic information systems, *Cartography and Geographic Information Systems*, 17: 81-84.
- Muller, J.C. (1979) Perceptions of continuously shaded maps. *Annals, Association of American Geographers*, 69: 240-249
Examines how users perceive continuous choropleth maps and concludes they are able to understand them.
- Ord, J.K. and Getis, A. (1995) Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis*, 27: 286-306.
- Ousterhout, J.K. (1994) *Tcl and the Tk Toolkit*, Addison-Wesley, Reading, Ma., 480p
Tcl/Tk creator and guru introduces the Tcl scripting language and Tk graphical toolkit.
- Petersen, M.P (1979) An evaluation of unclassed cross-line choropleth mapping. *American Cartographer*, 6: 22-23.
- Stuetzle, W., 1988, Plot windows. In Cleveland, W.S. and McGill, M.E. (eds), *Dynamic graphics for statistics*. Belmont CA: Wadsworth Inc., p. 225-245.

- Tobler, W.R. (1973) Choropleth maps without class intervals. *Geographical Analysis*, 5: 26-28.
- Tobler's paper first introduced classless choropleth maps of the sort that are nowadays often the default option in software. Petersen's paper evaluates their use in practice.
- Unwin, D.J. (1981) *Introductory Spatial Analysis*. London: Methuen
A dated account of basic spatial statistical analysis that deals with visualisation issues.
- Wegman, E.J. (1990) Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association*, 85: 664-675.
Excellent introduction to and account of using parallel co-ordinates to analyse and interpret multidimensional data distributions.
- Wright, J.K (1933) A method of mapping densities of population with Cape Cod as an example. *Geographical Review*, 26: 103-110.

10. Other Related Resources.

There are significant additional resources available from the JISC/NTI Project ARGUS, notably a set of WWW pages hosted on the MIDAS service and some other public domain map/Visc packages such as MANET (from University of Augsburg), X-Gobi (from USA) and X-LISP-STAT (see Brunson, 1997). A further *Case Study* in this volume by Wise, Haining and Signoretta reviews and evaluates some of these.

List of WWW Locations

1. The Martin/Bracken estimates:
gopher://midas.ac.uk/11/midas/datasets/census.dir/surpop.dir/
2. Project ARGUS: www.midas.ac.uk/argus/
3. Tcl/Tk for Mapping Tutorial www.geog.le.ac.uk/jad7/Tcl/ICA/index.html
4. Tcl/Tk for Mapping Examples
www.elsevier.nl/inca/homepage/miss/cageo/dykes/index.htm
5. Project KINDS www.midas.ac.uk/KINDS/
6. Project JANUS www.midas.ac.uk/JANUS/