

## MAPPING FLORIDA: BEYOND THE CHOROPLETH MAP

JEFF UELAND AND PETER VINCENT

*"The fundamental tool for the geographical analysis is undoubtedly the map, or perhaps more correctly the cartogram."*

### **Introduction**

With the advent of computer cartography and GIS there has been an increasing emphasis on map visualization, not only in terms of the use of color but also in terms of methods of data representation. Enough computer power is now available on the average PC to construct all but the most complex maps. Mapping is a way of visualizing the world and no map is without some sort of distortion. When mapping social phenomena however, there is ample reason to place the emphasis on data visualization with a topologically correct geography – the real purpose of mapping, rather than on the geography per se with a visually bias and often unhelpful data representation.

In this paper we want to illustrate the role of data visualization and representation by illustrating how cartograms, and particularly circular cartograms, can successfully be used to illustrate area data for Florida.

Jeff Ueland is an Assistant Professor in the Department of Geography at Ohio University. Peter Vincent is a Senior Lecturer in the Department of Geography at Lancaster University.

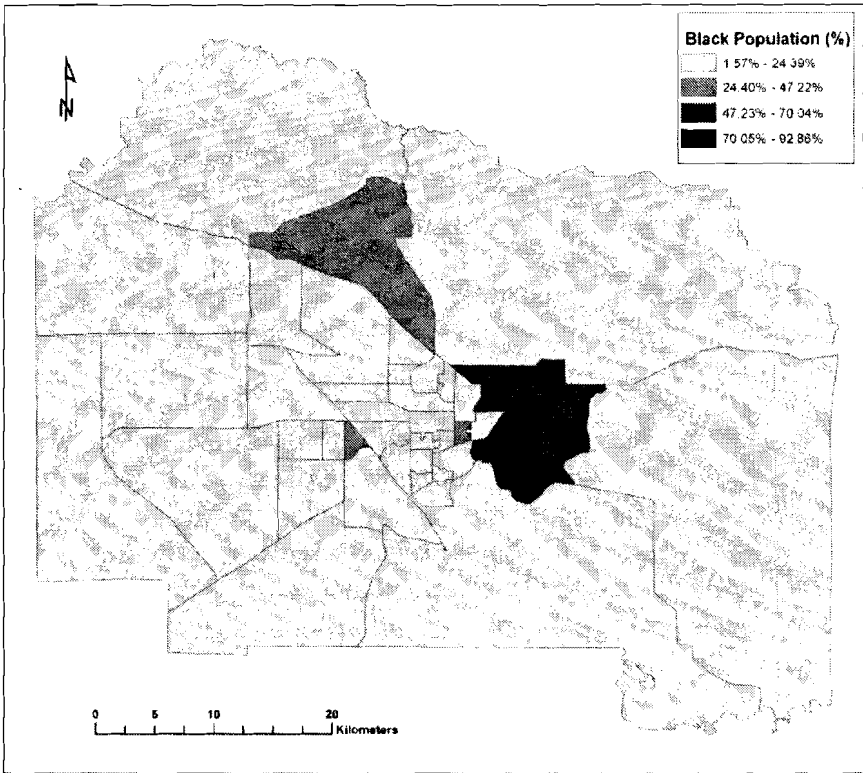


Figure 1. % Black Population in Alachua County 2000.

### Lies, damn lies and choropleth maps

Cartographically we appear to live in a taken for granted world; the world of the choropleth map. A glance at the maps in *The Atlas of Florida* (Fernald 1981) illustrates the point well. Choropleth maps are used to visualize data which have been enumerated, or summed over some defined region. Their regional boundaries create spatial bins over which there is no control in terms of shape and size. The data in each bin is then classified into categories and arbitrary shades used to represent them. The final map may be as much a result of the zones and the choice of categories and symbols as it is the underlying geography (Monmonnier, 1991).

That choropleth maps have become the *de facto* mapping default can easily be determined by the number of Florida agencies which have websites using this method. Practically none of these agencies has a ‘handle with care’ warning for the unwary viewer. In particular, the viewer’s attention is often focused on the size of the area rather than the data themselves. Large areas tend to dominate the map, and for variables such as income or car ownership, these are generally associated with the least densely populated areas. Take for example the choropleth map for Alachua County shown in Figure 1. Here, the densely populated inner city census tracts are visually overwhelmed by the larger peripheral map bins. For many classes of data the absolute quantities are often functions of the areal size of the mapping bin. For example, numbers of persons, farms, automobiles are naturally related to the size of the mapping units. To overcome this problem absolute data can be transformed to a density or ratio measure but this doesn’t overcome the visual impact of different size mapping bins (Robinson *et al.*, 1984).

### **Cartograms**

Maps are called cartograms when distortions in size, and occasionally of shape and distance, are made explicit and are seen as desirable (Dorling, 1996). Cartograms are an alternative to choropleth maps and place the emphasis on the data. They are particularly helpful in visualizing the details in human geography. Two types of cartogram are usually recognized; contiguous and noncontiguous. A contiguous cartogram creates no gaps between the data bins, whereas the noncontiguous type preserves both shape and topology but not neighbor adjacency.

Simple rectangular cartograms can easily be constructed manually but are very time consuming and it was only in the 1980s that computers were first employed to do the graphical computation. Nowadays there are scripts available in some GIS packages to construct both contiguous and noncontiguous carto-

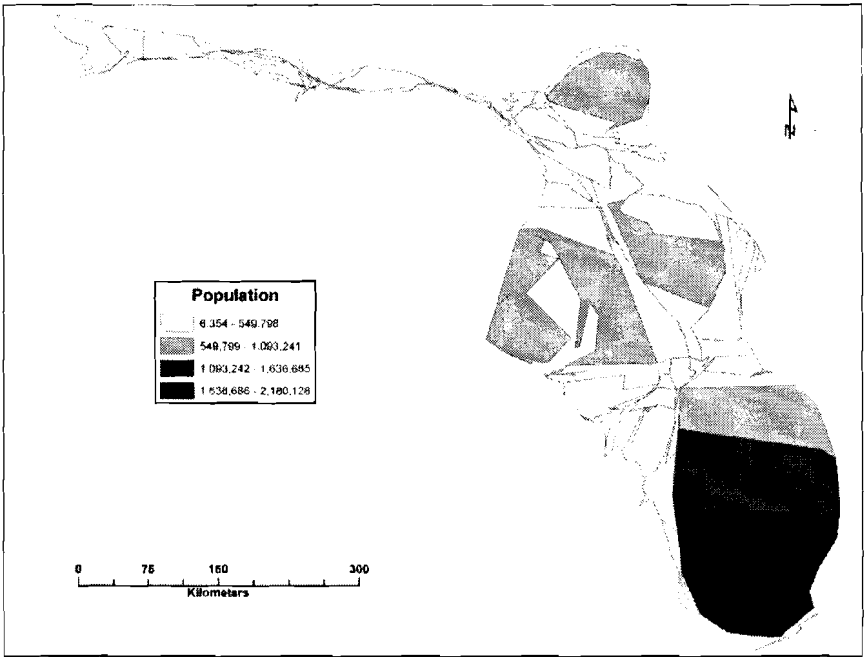


Figure 2. Contiguous Cartogram: Florida Population, 2000.

grams but the results are often not visually successful. Rubberbanding operations in the algorithms for contiguous cartograms can severely distort the overall shapes of the data bins and the final map can be visually very unsatisfactory and sometimes bizarre. Figure 2 is a contiguous cartogram created with a script call called “Contiguous Cartograms” and was downloaded from the ESRI website: <http://arcsripts.esri.com/>. The script is based on the algorithm devised by Jackel (1997) and functions iteratively. The more iterations executed the more distortion occurs across the cartogram, thereby accentuating those areas with the highest data vales. Areas with low data values are reduced so as to be almost visually insignificant. Figure 2 was created after 15 iterations of the “Contiguous Cartogram” script.

Figure 3 is a noncontiguous cartogram for population in Florida counties created by a script called “Cartogram!” also

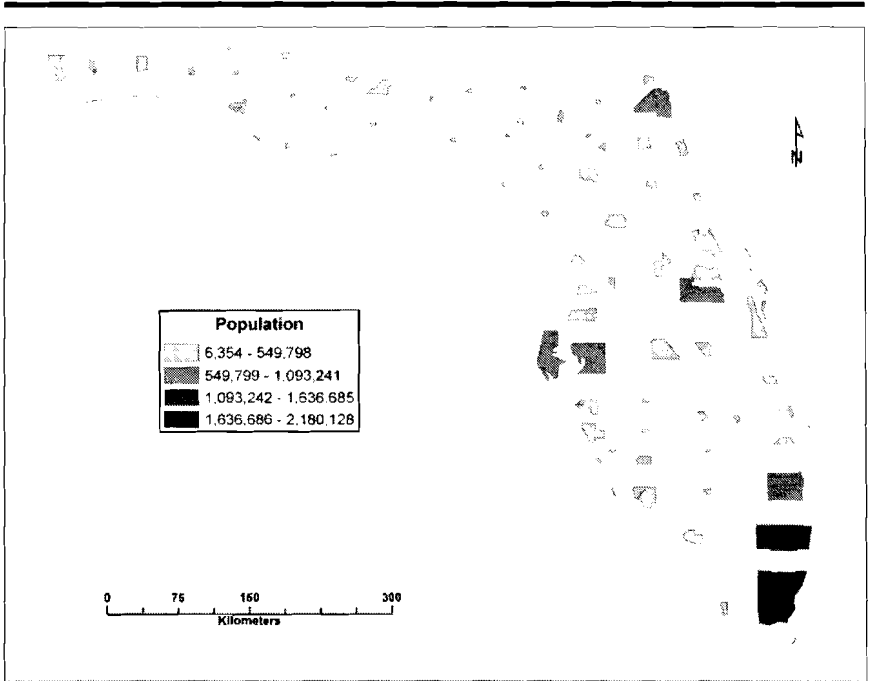


Figure 3. Noncontiguous Cartogram. Florida Population, 2000.

available at the ESRI website: <http://arcsripts.esri.com/>, and based on Olson’s algorithm (1976). This cartogram process examines the data to find the largest and smallest population values by county. The data are then scaled to produce a cartographic representation whereby the county with the largest population (Miami-Dade) is also the largest in area. During this process the actual shape of the county is preserved while not allowing any overlap between county boundaries.

### Circular Cartograms

In order to overcome the visualization problems associated with conventional cartograms Dorling (1996) developed the idea of circular cartograms. Dorling suggested that since shape-distortion is one of the central drawbacks in the use of cartograms and asked why cartograms were not constructed with simple

shape characteristics in the first place. The simplest area shape of all is the circle and Dorling developed an algorithm to develop a cartogram in which the area of the circle was proportional to the data item in question; all circles are non-overlapping and also in their correct topological relationship. At the start of the algorithm the circle centroids correspond to the original bin centroid. As the cartogram develops through iterations of the algorithm the centroids become adjusted so as to maintain topology. A simple way to imagine this is to think of each centroid joined to its neighbors by rubber bands. Each set of rubber bands can be adjusted/stretched but not broken or overlap.

Constructing the Dorling cartogram is fairly straightforward and the C code is published in Dorling (1996). Interesting web-sites which provide visualizations of the construction of Dorling cartograms can be found at: Herzog's Mapresso website: <http://www.mapresso.com/> and Dyke's CDV site: <http://www.geog.le.ac.uk/jad7/software.html>. Both sites provide Dorling applets which can be downloaded.

Circular cartograms were very successfully employed in Dorling's *A new Social Atlas of Britain (1995)* in which he noted that such circle representation was socially more just than biased visualizations produced when using conventional choropleths. Dorling used the word 'just' because of his interest in the geography of poverty and he was concerned that there should be a proper representation cartographically of the underprivileged and poor. In short, he was very aware of the ethical issues raised by visualizing such variables using conventional choropleth maps.

The creation of circular cartograms using the CDV applet was utilized to create the Dorling cartograms in this paper. CDV requires the user to construct two files. The first is an ArcInfo un-generate file that represents map bins. A un-generate file is a list of x and y coordinates for each vertex of a give data bin. The second file is a tab delimited text file which contains the variable(s) that are to be represented as a cartogram. Layouts for these two files are

(a)

END	X	Y
18	-36265.51000000	3394062.00000000
	-52185.64000000	3443775.00000000
	-44769.02000000	3443380.00000000
	-14384.82000000	3441131.00000000
	-14384.82000000	3441131.00000000
	-17596.77000000	3382360.00000000
	-17596.77000000	3382360.00000000
	-23617.07000000	3386999.00000000
	-39203.53000000	3379126.00000000
	-56575.06000000	3379370.00000000
	-56575.06000000	3379370.00000000
	-52185.64000000	3443775.00000000
END		

(b)

ID	County	Population
1	Escambia	284547
2	Santa Rosa	121708
3	Jackson	45796
4	Holmes	18893
5	Okaloosa	171174
6	Walton	38486
7	Washington	20502
8	Nassau	56859
9	Gadsden	44084
10	Leon	219081
11	Jefferson	13096
12	Madison	17761
13	Hamilton	12808

Figure 4. Examples of data inputs for CDV: (a) is an un-generate file of a county (b) is the tab delimited data file.

ID	New X	New Y	Population	Radius
1	8.755638	236.7741	15.40470187	
2	32.92931	228.722	10.07478658	
3	97.81853	221.9487	6.180022633	
4	78.80788	230.1055	3.969417267	
5	53.08924	219.8574	11.9479994	
6	72.29932	222.6651	5.665362193	
7	85.38681	225.3727	4.13498946	
8	223.4616	224.1335	6.88614084	
9	116.8775	218.8508	6.063407954	
10	141.0929	217.3746	13.5169482	
11	156.4827	206.1851	3.304803375	
12	171.1041	208.4032	3.848664017	
13	188.4753	214.5749	3.268262657	

Figure 5. Example of a CDV output file.

illustrated in Figures 4a and 4b respectively. The construction of the second file is most easily generated in Microsoft Excel and is then converted by CDV to an internal format. Both files are loaded into CDV and the iterative process started. When finished, CDV creates a new file that contains the new x coordinate, y coordinate, radius for each new circular data bin, as seen in Figure 5. These data can then be brought into a cartographic or GIS package for further enhancement. Figure 6 illustrates a Dorling cartogram of Florida population by county. The data output from CDV were imported to ARCGIS for map construction.

In Figure 7 we have taken the Dorling approach one step further to demonstrate how it may be used to visualize multivariate data. Here, the size of the circle represents the total population of each county, the height of the column represents number indi-



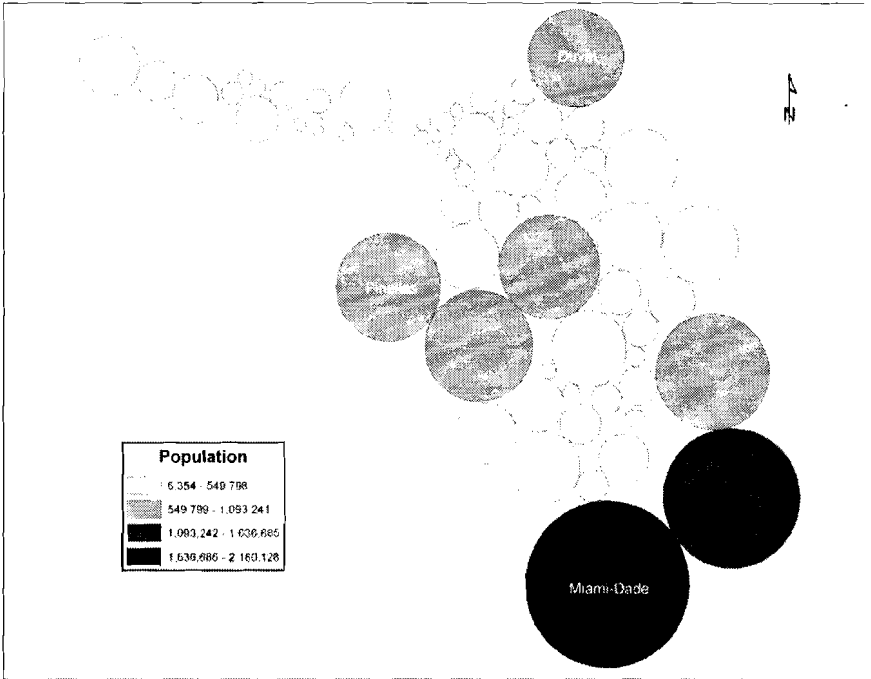


Figure 6 Dorling Circular Cartogram: Population in Florida, 2000

viduals below the poverty level, and the shading of each circle represents median household income. Figure 7 was created in ArcScene, a module available in ARCGIS’s 3-D analyst.

**Conclusions**

Cartograms can now be computed relatively easily and overcome many of the problems associated with conventional choropleth maps. Cartograms are excellent for visualizing data which is surely the central role of maps. In many situations, particularly where mapping social data, there is a strong case to be made for using methods that are not only cartographical unbiased but also ethically sound. Of course, map users might have to get used to seeing data displayed in relatively unconventional ways but there is nothing wrong in getting map users to ponder over a

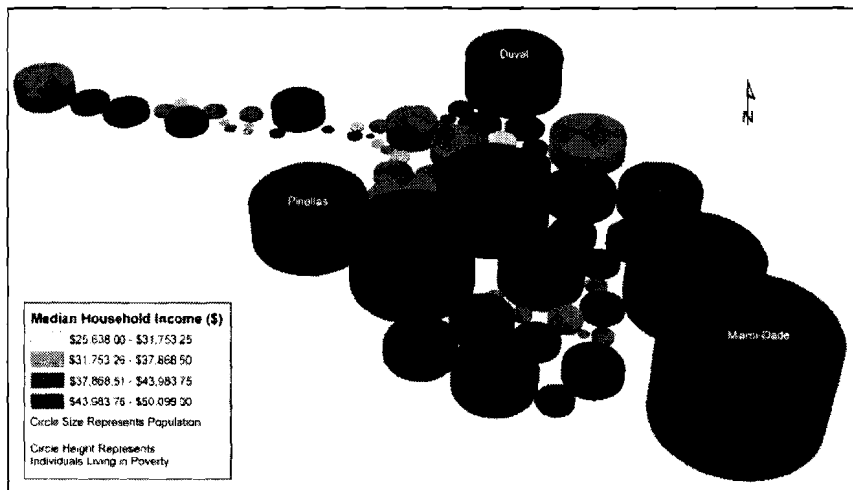


Figure 7. 3-Dimensional Darling Circle Cartogram: Population, Income, and Poverty in Florida, 2000

cartogram. This is surely much better than a cursory look at a conventional choropleth map with all its inherent visualization problems.

## REFERENCES

- Dorling, D. (1995) *A New Social Atlas of Britain*. Chichester: Wiley.
- Dorling, D. (1996) *Area Cartograms: Their Use and Creation*. Concepts and Techniques in Modern Geography, 59.
- Fernald, E. A. (1981) *Atlas of Florida*. Gainesville, FL: University of Florida Press.

Jackel, C. B. (1997) "Using ArcView to Create Contiguous and Noncontiguous Area Cartograms." *Cartography and Geographic Information Systems* 24: 101-109.

Monmonnier, M. (1991) *How to Lie with Maps*. Chicago, University of Chicago Press.

Olson, J.M. (1976) "Noncontiguous Area Cartograms." *The Professional Geographer* 28: 371-380.

Robinson, A. H., Sale, R. D., Morrison, J. L. and Muehrcke, P.C. (1984) *Elements of Cartography*. New York, NY : Wiley.

Stamp, D. (1962) President of the Royal Geographical Society, 1962: 135.