

Multivariate Label-based Thematic Maps

Richard Brath¹ and Ebad Banissi²

1. London South Bank University and Uncharted Software Inc., Toronto, Canada,
richard.brath@alumni.utoronto.ca

2. London South Bank University, London, United Kingdom, banisse@lsbu.ac.uk

Abstract: The rich history of cartography and typography indicates that typographic attributes, such as bold, italic and size, can be used to represent data in labels on thematic maps. These typographic attributes are itemized and characterized for encoding literal, categorical and quantitative data. Label-based thematic maps are shown, including examples that scale to multiple data attributes, and a large number of entities. Multiple approaches to handle long labels are considered. Positional and proportional encoding apply attributes to portions of labels for encoding a large number of data attributes or quantitative values. Quantitative evaluation indicates label-based thematic maps may outperform choropleth maps for some tasks. Qualitative evaluation provides guidance for design considerations.

Keywords: Typography, cartogram, visualization, font attributes

1. Introduction

Typography has been used in maps for hundreds of years to encode data into typographic elements such as size, oblique, case, underline, spacing and font-family. Figure 1 shows an example with these many typographic variants.



Fig. 1. Portion of map from *Stieler's Atlas of Modern Geography* (1909). Labels use variation in case, weight, underlines (dash, solid, double), italics (and reverse italics), spacing and font-family to encode data. Public domain.

However, thematic maps such as cartograms and choropleths represent data on maps typically with non-textual representations, such as changing the color or size of geographic areas. Typography is typically minimal or secondary to the other elements of these thematic maps. Popular thematic types include choropleth maps and many varieties of cartograms such as contiguous shape-preserving cartograms, Dorling cartograms and so on, with some examples shown in Figure 2.

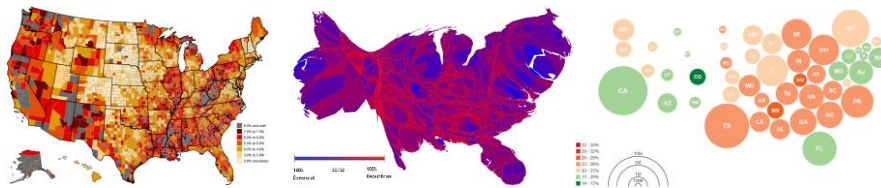


Fig. 2. Left: example choropleth map encoding data via color per county. Center: a contiguous cartogram representing data by adjusting the area of each county, plus a second data attribute by color. Right: a Dorling cartogram representing states as circles, with size and color indicating data. Sources: Bureau of Labor Statistics (www.bls.gov/web/metro/twmcort.pdf); Wikipedia (en.wikipedia.org/wiki/Cartogram), University of Washington (homes.cs.washington.edu/~j-heer/files/zoo), used with permission.

These thematic maps have some well-known problems. For example:

- Some areas can be too small to be discernable (e.g., Singapore),
- Large regions are more prominent than small regions, but some tasks require that regions are perceived equally, e.g., comparison of policies,
- The viewer may have low geographic literacy, for example, only 8.3 of 17 major countries could be identified by young Americans (GfK 2006).

Note that labels are rarely used, and do not encode data beyond the literal text. Instead, we use labels enhanced with typographic attributes that encode many dimensions of data. Our unique contributions include:

- Thematic maps where the label is the key element conveying data.
- Many choropleth and cartograms indicate only one or two data variables, e.g., color (in a choropleth map), or size and color (in a cartogram). Typographic attributes can encode many more data variables into a single label.
- Encoding data along the length of a string can enable more levels or more data. (Positional and proportional encoding).

2. Background

Given centuries of typography and cartography, plus visualization research, there is a wealth of history to inform the use of typography in thematic maps.

2.1 Cartographic Typography

Typographic variation to indicate different data in maps has existed since early maps. For example, the 13th century *Matthew of Paris' map of Great Britain* differentiates among shires in multicolor uppercase, counties in red, cities in black with some cities in boxes as shown in Figure 3.



Fig. 3. Portion of *Matthew Paris' map of Great Britain*, with different types of place names in different colors and case. Matthew Paris, 1200-59, British Library, <http://www.bl.uk/collection-items/matthew-paris-map-of-britain>. Public domain.

After the invention of the printing press, colored text was used less frequently. However, other typographic attributes, such as italics or case, were used to indicate different types of data. The differentiation occurred whether hand-engraved text, such as Mercator's map of Palestine in 1537 (Osley 1969) or movable type, as seen in Figure 4, a snippet of Munster's *Geographia Universalis* (1540).



Fig. 4. Portion of map from Munster's *Geographia Universalis* with different types of place names in italics and capitalization. Sebastian Munster, "Italia XIII> Nova Tabula" from *Geographia Universalis*, 1540. Map #83 from Otto and Marie Neurath Isotype Collection, University of Reading, used with permission.

By the mid-1800's, there were more typefaces, weights, underlines and small-caps. Maps evolved to use these typographic attributes simultaneously within a label to indicate many data attributes. Figure 5 left shows a portion of an Ordnance Survey map legend (Hodson 1926) where city labels indicate: 1) the literal name of the city with text; 2) administrative centers with italics; 3) population category with font size; 4) country via font type (serif for UK, slab-serif or serif variant for Scotland); and, 5) whether it is a town (uppercase) or village (lowercase).

Similarly, Stieler's atlas (1925) uses different typefaces, case, underlines and spacing (Figure 5 right). Stieler encodes quantitative data into labels indicating a few ordered levels, such as underlines ordered dot, dash, thin, thick, double thin, double thick; or populations indicated as a combination of case and italics ordered lowercase italic, lowercase roman, uppercase italic, uppercase roman.

Large District Names	ROCKINGHAM	
Large Towns. Population, over 500,000	LIVERPOOL GLASGOW	
	(England)	(Scotland)
County Towns... d° 100,000 to 300,000	NORWICH ABERDEEN	
	(England)	(Scotland, according to population)
Other Towns... d° d°	BLACKBURN DUNDEE	
	(England)	(Scotland)
County Towns... d° 30,000 to 100,000	CHESTER	
Other Towns... d° d°	DARLINGTON PAISLEY	
	(England)	(Scotland)
County Towns... d° 10,000 to 30,000	DURHAM	
Other Towns... d° d°	MARGATE AIRDRIE	
	(England)	(Scotland)
County Towns... d° under 10,000	DORCHESTER	
Other Towns... d° d°	LICHFIELD, WEM. DALKEITH	
	(England)	(Scotland)
Parish Villages & Villages with 300 inhabitants and over	Bishopstoke	
Other Villages	Barton Barton	
	(England)	(Scotland)

The administrative status of towns is indicated by means of underlining, thus	
<u><u>BERLIN</u></u>	Capitals of Federal States (strong double underline)
<u><u>BREMEN</u></u>	Capitals of separate countries (thin double underline)
<u>MUNSTER</u>	Capitals of Provinces (strong single underline)
<u>KÖLN</u>	Capitals of large administr. distr. (thin single underl.)
<u>Stoll</u>	Capitals of small administrative districts
<u>Magde</u>	Capitals of Local Government districts
The above apply specially to Germany. In other countries the underlining is made to indicate generally the relative administrative importance of the towns. Colonial dependencies are underlined in the colour of the motherland.	
II. Classification of towns, villages, etc.	
LONDON LONDON LONDON	⊙ I. Class over 500,000 inhab.
LEEDS LEEDS LEEDS	⊙ II. " " 100,000 "
BATH BATH BATH	⊙ III. " " 50,000 "
Gotha Gotha Gotha	⊙ IV. " " 25,000 "
Clermont Clermont Clermont	⊙ V. " " 10,000 "
Irrea Irrea Irrea	⊙ VI. " " 5,000 "
Murillo Murillo Murillo	⊙ VII. " " under 5,000 "
Outside Europe, especially in countries without a regular official census, the townmarks and lettering are made to indicate in a general way the relative importance of the towns.	

Fig. 5. Small portions of map legends where multiple data attributes are encoded into labels using multiple typographic attributes. Ordnance Survey 1926, (www.davidrumsey.com/luna/servlet/s/15w94) and Stieler's 1924 Atlas (www.davidrumsey.com/luna/servlet/s/791qv7), both from the David Rumsey Map Collection, <http://www.davidrumsey.com>. Used with permission.

2.2 Visual Attribute Research

Bertin's *Semiology of Graphics* (1983) presented a systematic organization of information graphics including visual attributes (i.e., retinal variables) such as size, hue and brightness, and their representations as points, lines or areas, from which a wide variety of charts, graphs and thematic maps can be constructed. Bertin did not explicitly identify text labels as a visual attribute. Use of labels and specific typographic attributes occur tangentially, such as hand-annotated maps at the elemental level labelling proper nouns in all caps (Bertin 1983, pp. 308-309).

Since then, other researchers have expanded on Bertin's visual attributes from various fields including visualization, cartography and perception; but researchers have generally ignored text or minimized text to a single visual attribute, as shown in Table 1. The *Tyner 2010* column summarizes visual attributes from sources including Kraak & Ormeling, Krygier & Wood, MacEachren (shown as a separate column), Monmonier, Slocum and Tyner. The *Preattentive Perception* column summarizes psychology research discussed by Healey and Enns (2012).

Table 1. Table of Visual Attributes

Table of Visual Attributes		Information Visualization Researchers										Vision Research	Cartography		
Visual Group	Visual Attribute	Bertin 67	Cleveland 85	MacKinlay 86	Wilkinson 99	Ware 00	Mazza 09	Brath 11	Ilinsky 12	Chen, Floridi 13	Munzner 15	Preattentive Perception	MacEachren 95	Tyner 10	
Transform	Position	X	X	X	X	X	X	X	X	X	X		X	X	
	Length		X	X		X	X	X	X	X	X	X			
	Size (Area)	X	X	X	X	X	X	X	X	X	X	X	X	X	
	Orientation	X		X	X	X	X	X	X	X	X	X	X	X	
	Volume		X	X		X					X			X	
Shape	Shape	X	X	X	X	X	X	X	X	X	X		X	X	
	Angle		X	X				X	X	X					
	Curvature							X		X	X				
	Line Ending						X	X	X	X		X			
	Closure							X	X			X			
	Edge Type							X	X						
	Corner Type							X	X						
	Icon, glyph, etc.									X					
	Colour	Brightness	X		X	X	X	X	X	X	X	X	X	X	X
		Hue	X	X	X	X	X	X	X	X	X	X	X	X	X
Saturation				X	X	X	X	X	X	X		X	X		
Texture	Granularity	X		X	X	X	X	X	X	X		X	X		
	Pattern			X	X	X	X	X		X					
	Orientation				X	X				X					
Relation	Connection			X			X	X	X	X					
	Containment			X			X	X	X	X					
Optics	Blur				X				X			X	X		
	Transparency				X				X			X	X		
	Stereo Depth										X				
	Concave/Shade								X		X				
	Light Direction								X		X				
	Shadow								X						
	Partial occlusion								X						
Movement	Flicker					X		X			X				
	Speed					X		X			X				
	Direction							X			X				
Miscellaneous	Numerosity										X				
	Spatial Grouping										X		X		
	Arrangement											X	X		
	Resolution											X	X		
	Artistic Effects										X				
	Text Labels			X	X	X	X	X							

Past research may have overlooked text and font attributes on the assumption that other visual attributes capture this information. But this oversight leaves ambiguity and questions. For example, Bertin identifies the attribute intensity, which could be applied to a label (e.g., light grey, dark grey, black), but font weight (lightweight, bold, black) is also similar to intensity. Similarly, size can be adjusted by scaling text, or adjusting intercharacter spacing, or using variants of font width (e.g., condensed and expanded variants). What font attributes are available?

Researchers have identified typographic conventions for categories in maps. For example, Cuff and Mattson (1982) surveyed cartographic labelling conventions including uppercase, italics and serif/sans serif as shown in Table 2.

Table 2. Table of Letter Convention of major map publishers (redrawn based on Hodges 1989)

Feature		Publisher		National Atlases		Atlases		
		Reference	Magazine	U.S.	Canada	CIA	Goode	Oxford World
Water	Salt	<i>CAPS</i>	<i>Mixed</i>	<i>CAPS</i>	<i>CAPS*</i>	<i>Mixed</i>	<i>CAPS*</i>	<i>CAPS*</i>
	Fresh	<i>Mixed</i>	<i>Mixed</i>	<i>Mixed</i>	<i>Mixed</i>	<i>Mixed</i>	<i>Mixed</i>	<i>CAPS*</i>
Land Features	Other	<i>CAPS*</i>	<i>Mixed</i>	<i>CAPS</i>	-	<i>CAPS</i>	<i>CAPS</i>	<i>CAPS</i>
	Peak	<i>CAPS*</i>	<i>Mixed</i>	<i>Mixed</i>	-	<i>Mixed</i>	<i>Mixed</i>	<i>CAPS</i>
	Range	<i>CAPS*</i>	<i>Mixed</i>	<i>CAPS*</i>	-	<i>CAPS</i>	<i>CAPS</i>	<i>CAPS*</i>
Cultural	City	<i>Mixed</i>	<i>Mixed</i>	<i>CAPS*</i>	<i>CAPS*</i>	<i>CAPS*</i>	***	<i>Mixed</i>
	Park	<i>CAPS</i>	<i>CAPS</i>	<i>CAPS</i>	-	-	<i>CAPS</i>	<i>CAPS</i>
	Political	<i>CAPS</i>	<i>CAPS</i>	<i>CAPS</i>	<i>CAPS</i>	<i>CAPS*</i>	<i>CAPS</i>	<i>CAPS</i>
Thematic	Notes	<i>Mixed</i>	<i>Mixed</i>	<i>Mixed</i>	<i>Mixed</i>	<i>Mixed</i>	<i>Mixed</i>	<i>Mixed</i>
	Legend Item	<i>Mixed</i>	<i>Mixed</i>	<i>Mixed</i>	<i>Mixed</i>	<i>Mixed</i>	<i>Mixed</i>	<i>Mixed</i>
	Legend Head	<i>Mixed</i>	<i>Mixed</i>	<i>CAPS</i>	<i>CAPS</i>	<i>Mixed</i>	<i>Mixed</i>	<i>Mixed</i>
	Title	<i>CAPS</i>	**	<i>CAPS</i>	<i>CAPS</i>	<i>Mixed</i>	<i>CAPS</i>	<i>Mixed</i>

Brewer's *Designing Better Maps* (2005) indicates use of typeface, italic/roman, hue, alignment for categorical data and size, weight, scaling (condensed/expanded), brightness, case and spacing for ordered data. Krygier (2005) indicates the use of weight, italic, size, hue, brightness, case, spacing and condensed. Krygier also recognizes but recommends against underlines. Muehlenhaus (2014) discusses qualitative attributes including font family (recommending sans serifs), oblique and type arrangement; and quantitative attributes including case, value and weight.

Visual attributes and typography occur separately in most cartography texts (e.g., MacEachren, Tyner, Muehlenhaus, Krygier), and generally typographic attributes are not considered elements for thematic maps, even though they share many similarities to visual attributes.

2.3 Typographic Attributes

Table 3 summarizes typographic attributes to encode data into labels and their relationship to well understood pre-attentive visual attributes (Wolfe & Horowitz 2004). Type attributes are listed vertically:

- *Glyphs* are the letterforms that literally encode data. Alphanumeric glyphs are orderable, but symbols are not.
- *Font Family Attributes* are type features typically made available with a typeface family, such as **bold**, *italic*, ~~typeface~~ and condensed/expanded fonts.
 - *Font weight* can encode ordered data. Up to nine weights are available in some fonts but fewer levels are perceivable at small sizes.
 - *Oblique* can have both a reverse slope, non-slope and familiar forward slope, making it feasible to represent both positive and negative values.

- *Case* can be used in Latin-based alphabets to create an ordering from lowercase, to SMALL CAPS, to UPPERCASE.
- *Typeface* can be used for qualitative data (i.e., categories). While there are 100,000+ fonts, only a few groups work at small sizes and are significantly different from each other to be noticeably different.
- *Underline* is broadly available in most fonts, and, as shown in the atlas in Figure 1, can be used to create an ordering.
- *Sequence* indicates differences which require context of a sequence of glyphs, such as a superscript or s p a c i n g. Squishing text is not recommended by typographers but a few fonts are specifically designed for scaling, e.g., *Neue Swift*. Pairs of delimiters such as (), <, ** are a potential means for differentiating some labels.
- *Font Design* indicates features manipulated by a font designer when creating a new font. While these may be feasible typographic attributes to encode information, in practice, they are not easily accessed from most applications. Parametric fonts do expose many low level features (e.g., prototypo.io).

Table 3. Table of typographic attributes, relation to visual channels and types of encoding.

Group	Font Attribute	Visual Channel					Preattentive Potential	Best for encoding: Q: quantitative O: ordered C: categoric G: grouping/relationship L: literal	Example
		Position	Length/Size	Intensity	Orientation	Shape			
Glyphs	Alphanumeric Text Glyph					◆	D	L, O	ape bat cat dog 123 456
	Symbols					◆	D	C	! ? # @ #comment \$var
Font Family Attributes	Font weight		◆	◆			HP	Q (2-9 levels)	1.0 2.0 3.0 5.0 8.0
	Oblique / Italic			◆			HP	C, Q using slope angle	-2.0 -1.0 0.0 1.0 2.0
	Case inc small caps		◆				P	C , possible O (2-3 levels)	BIG Avg. Small tiny
	Typeface					◆	P	C (2-6 levels)	Swiss French German <i>Italian</i>
	Underline	◆	◆				HP	C, O, Q (using length)	plain dash single double
	Condensed		◆	◆			HP	Q, O (2-4 levels)	1000 2000 3000
Sequence	Squished		◆	◆			HP	Q	stout thin plain wide fat
	Spacing		◆	◆			HP	Q, O	tall grand venti
	Baseline shift (e.g. subscript)	◆	◆				HP	C (2 levels)	Normal ^{High} _{Low}
	Delimiters					◆	D	G	(but) *and* <or>
Font Design	X-height		◆				HP	O, Q (few levels)	ick ick ick
	Contrast / Stress angle				◆		P	O (few levels)	LOW MED. HIGH
	Serif length / Bracket size		◆				HP	O, Q (too small to see?)	see prototypo.io

◆ / * indicates primary / secondary visual channel for font attribute

† HP: Highly probable, P: probable, D: doubtful

Note: does not include traditional visualization attributes such as size, color, texture, outline, blur, shadow, etc., which are applicable to any marker including type.

The *Visual Channel* columns indicate the relation between the typographic attribute and well researched visual channels. For example, font weight at a macro-level can be understood to alter the intensity of characters while at a micro-level it can be understood as varying the size (i.e., width) of letter strokes.

Some visual channels have stronger cues for guiding attention and can visually pop-out, as summarized in the column *Preattentive Potential*. For example, both intensity and size are considered preattentive, so font weight, which uses both, is considered highly probable in this table. These preattentive capabilities imply that typographic attributes can be utilized similar to other visual attributes – that is – to create thematic maps.

3. Multi-variate Label-based Cartograms

Choropleth maps and cartograms have issues, as previously outlined in the introduction. Instead of indicating a geographic region's identity with shape, a label can be used. A label literally encodes the entity and uses visual attributes to visually pop-out data. Figure 6 shows two simple label-based cartograms of the U.S. states as squares (left) and rectangles (right). In both, data is indicated by color.

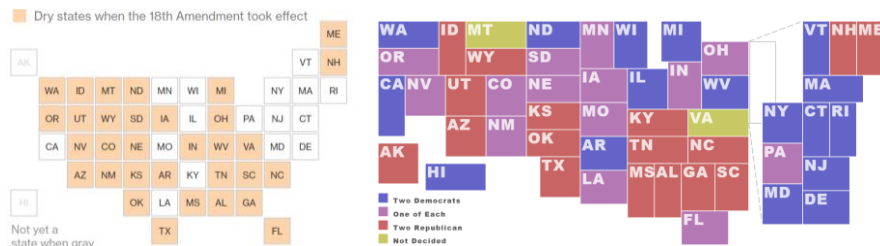


Fig. 6. Label-based cartograms. Labels indicate states and color indicates data. Left image © 2015 Bloomberg L.P., <http://bloom.bg/2k5uqf>. Used with permission. Right image by author.

3.1 Multivariate Labels

With a label-based cartogram, the label can be used as the basis to encode data using a wide variety of visual attributes including font-specific attributes. Figure 7 uses font weight to encode 2010 GDP per capita, font spacing for GDP growth, font oblique angle for inflation (reverse slope for deflation, e.g., Ireland), and color for region. This representation can be used to answer simple questions of a single variable, e.g., Are there countries with high growth rates (A: Yes, widely spaced China, Korea, Seychelles). Complex questions across multiple variables can be answered too, e.g., Are there countries with high GDP, high growth and low inflation (A. Heavyweight, widely spaced, non-italic, e.g., Macao, Sweden); or, Are there countries with low GDP, high growth rates and low inflation (A. Lightweight, widely spaced, non-italic, e.g., Kenya, Timor-Leste).

Note that only a subset font attributes are available in any given font family. While two levels of weight are fairly common (i.e., plain and bold), many fonts do not provide more. Multiple angles of italic are rarely available: here an oblique angle is created by applying a shear transformation to the text. Many fonts do not have expanded or condensed versions: font spacing (i.e., tracking) is an alternative used in this example. Many of the examples in this document use fonts *Segoe UI*, *Source Sans Pro*, or *Roboto*: high-quality, widely available, sans serif fonts with more than five weights (and with *Roboto* also offering a condensed version).

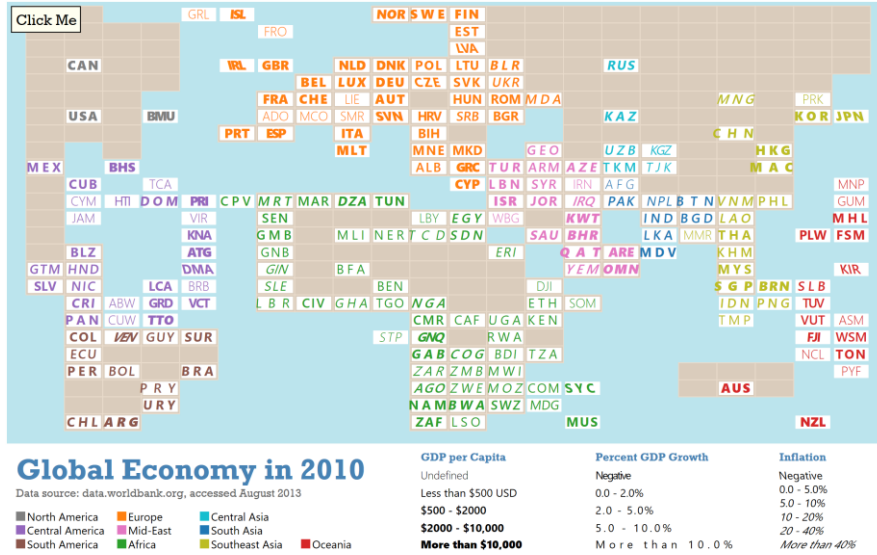


Fig. 7. A multivariate label-based thematic map indicating data via color, font weight, spacing and oblique slope angle. Image created by authors.

3.2 Scalability to Thousands of Labels

The approach can scale to different layout algorithms and higher data densities.

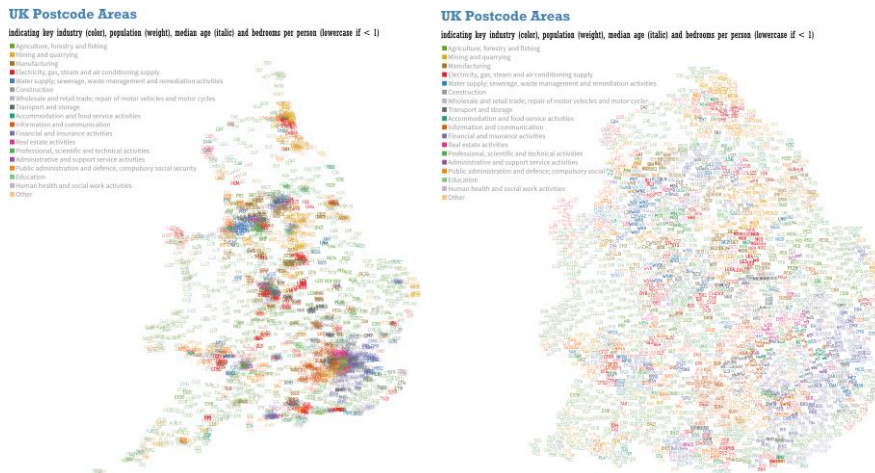


Fig. 8. Left: 2269 UK postcode districts, located geographically, showing data by font weight, oblique angle, case and color. Right: Cartogram of the same data. See Figure 9 for close-up.

Figure 8 shows 2269 UK Postcode districts plotted directly by latitude and longitude on the left; and a cartogram of the same data on the right. In the left view, most labels are occluded and the plot only illustrates postcode densities. In the right view, each label is clearly distinguished and readable on a modern monitor. Visible patterns are identifiable. For example, color indicates the occupation with the highest difference vs. national average. Green for agriculture in the west, blue for finance around east London, amber for mining in the north-east, etc.

Figure 9 shows a close up. Letters indicate postal code districts, e.g., M## for districts in Manchester. Weight indicates population. Oblique angle indicates median age: reverse slope indicates a younger age, forward slope for older age. Case indicates number of bedrooms per person: uppercase if more than one bedroom, lowercase if less. In Figure 9, patterns in oblique angle indicate regional variation in age, e.g., most of the postcodes near the top left (LL## indicating districts near Llandudno) slope right (indicating older median age) while districts within Manchester (M##) tend to slope left (indicating younger ages).

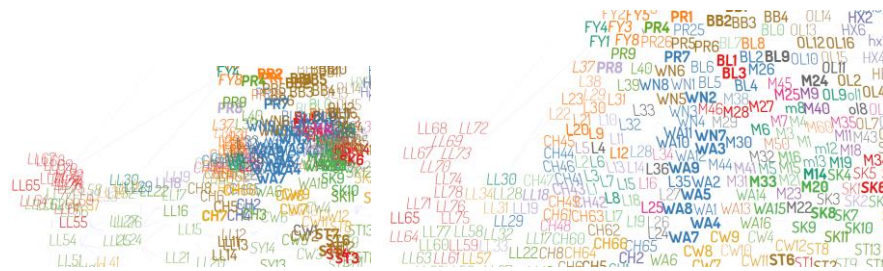


Fig. 9. Close up of Northern Wales, Liverpool and Manchester. Images by authors.

3.3 Long Labels

Both previous examples used short mnemonic codes for labels. Sometimes only labels of varying length are available. This creates problems: 1) it is difficult to fit long labels into circles or squares; and, 2) long labels have greater visual area than short labels, thereby making the long-labeled item more salient than the short-labeled item (e.g., the label *San Francisco* is more prominent than *Rome*). Solutions could include:

- *Forced Contraction*: Limit strings to an arbitrary number of characters followed by an ellipsis to indicate missing characters (e.g., San Fran...).
- *Popular Contraction*: Some place names have well known contractions, e.g., NYC for New York City.
- *Use areas with long aspect ratios*: The areas used in the cartogram can be long and narrow to better fit long labels.

- *Labels exceed areas*: While it is ideal to fit the label to the area, labels can be longer than the target area.

Figure 10 shows a portion of a cartogram of Canada census data – some cities have long labels. Labels are associated with brightly colored fixed size areas, reducing the visual prominence of label length. For example, short heavy weight labels (e.g., *Montreal*, *Toronto*) are more prominent than long lightweight labels (e.g., *Saint-Jean-sur-Richelieu*, *Kitchener-Cambridge-Waterloo*).

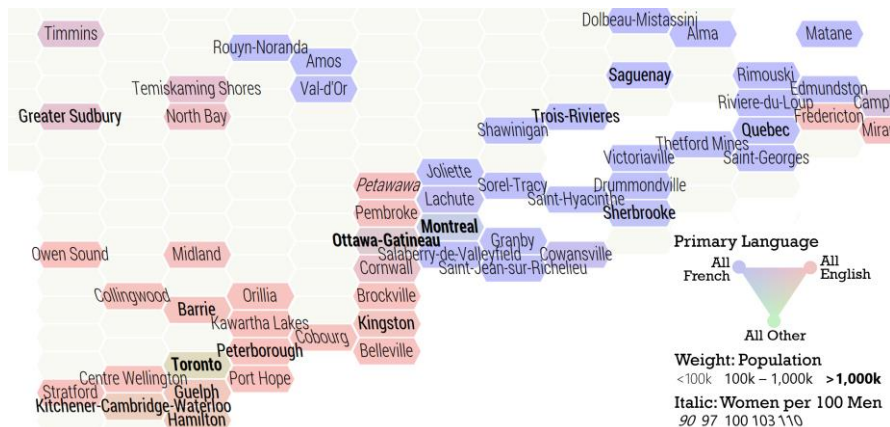


Fig. 10. Close up of cartogram of Canadian metropolitan areas. Image by authors.

3.4 Positional and Proportional Encodings

When labels use a sequence of characters, these characters can be manipulated individually. Instead of encoding a value by applying it to all characters, it can be encoded to a subset of characters:

Positional Encoding: By treating each character independently, each position can encode different data. For example, the cartogram in Figure 11 indicates nine different risk factors: bold, italic and underline are applied independently to each letter of the three-letter ISO code. At a macro-level, country codes with few additional formats are less risky than country codes with many formats.

Proportional Encoding: Any typographic format can be applied to a length of characters to indicate quantities. For example, in Figure 12, the proportion of uppercase is used to indicate life expectancy: JPN has longer lives than USA, which in turn is longer than Mda, and in turn longer than zaf.

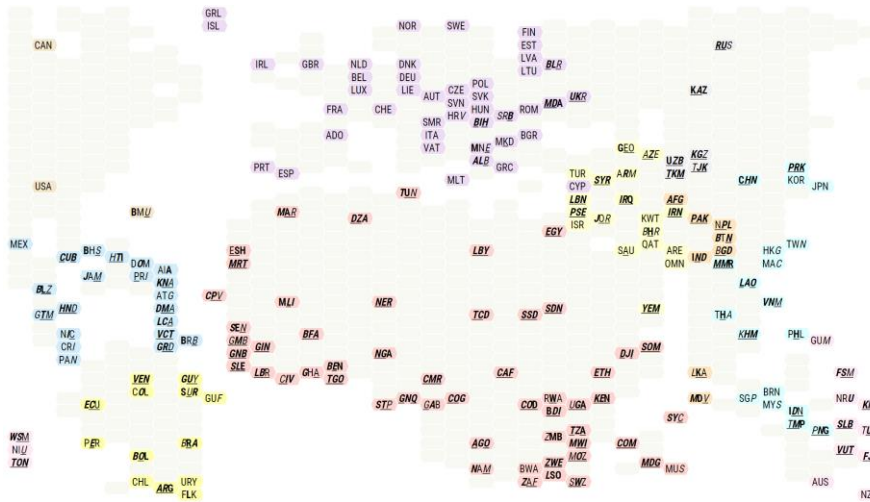


Fig. 11. Map indicating nine different risk factors indicated by bold, italic and underline applied independently to each letter of three-letter ISO code. Image by authors.

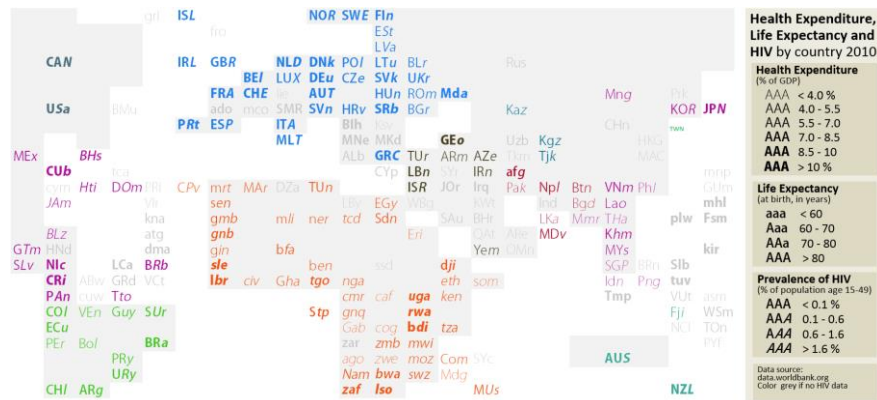


Fig. 12. Map indicating life span using proportions of uppercase, and prevalence of HIV by proportion of italic to indicate data. Image by authors.

3.5 Label-based Layouts

While the focus of this article is on the use of font attributes in cartograms, different algorithms have been used to produce the above cartograms:

Grid-based Cartogram: The cartograms in figures 7, 10 and 11 use a regular mesh-based cell deletion approach (e.g., rectangular grid or hexagonal tiling). It

starts with a large, regularly subdivided mesh such that each geographic item (e.g., country, district, city, etc) has a unique cell location. Then, successive rows (or columns) are merged such that merging two adjacent rows (or columns) retain all items and no two items occupy the same cell.

Force-Directed Cartogram: The cartograms in figures 8 and 9 use a force-directed layout based on an extracted mesh. First, all items are placed using latitude and longitude. Then, a Delaunay triangulation connects adjacent items into a triangular mesh. Next, a force-directed layout is used to push apart items which are close together (repulsive charge), while keeping adjacent items adjacent (edge forces). The result of a force-directed layout may have a few items that remain partially overlapped. A final step walks through all the pairwise overlapping bounding boxes and pushes labels apart.

Non-Geographic: When a label-based approach is used, labels can be rearranged into other visualizations – such as scatterplots, Venn diagrams, distributions, tag clouds, heatmaps, graphs and so on. Figure 13 shows a scatterplot of countries plotting birth rate vs. death rate, with label size indicating country population. With a label-based representation, consistent markers (i.e., the label) can be utilized across different visualization techniques maintaining constancy thereby reducing cognitive load and reducing effort to learn multiple representations.

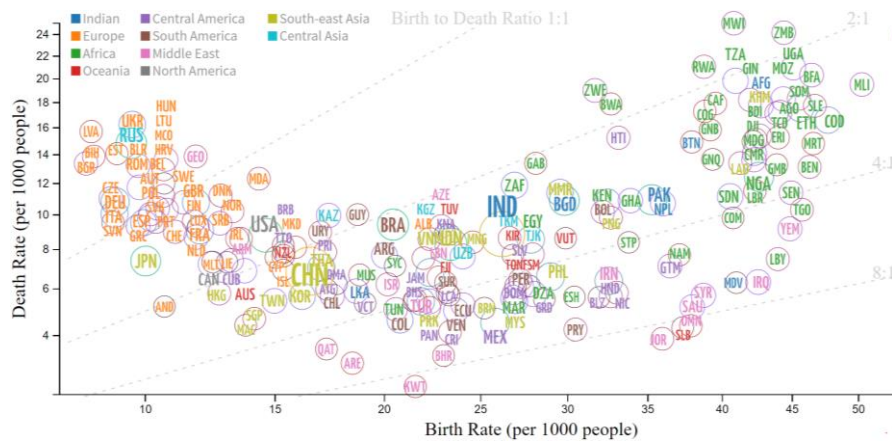


Fig. 13. Country birth rate vs. death rate. Image by authors.

4. Discussion

There are many different aspects to the effectiveness of typographic attributes that could be evaluated: speed of perception vs. ease of decoding, interference between attributes when using multiple attributes in the same label, potential infor-

mation loss, legibility and number of levels per attribute which can be perceived, impact of typographic attributes on readability, aesthetics of typographic-attribute based maps, intuitive encodings of typographic attributes, and so on. Evaluations are on-going: some have been done and many more should be considered.

Information lossiness is an evaluation of the fidelity per visual attribute in an encoding; and an estimate of permutations across multiple attributes to compare relative lossiness (Brath and Banissi 2015). A simple example is shown in Figure 14: both maps use 5 levels of hue. However, the colors of countries can be clearly identified in only 143 out of 187 countries in a choropleth map (at 520 x 310 resolution), while the corresponding label-based cartogram has all 187 countries with values readable (at the same resolution). In this example the choropleth has a fidelity of 76% compared to the cartogram. This approach can be extended to representations with many more attributes to consider the effect of tradeoffs between different encodings. In general, label-based cartograms outperform choropleths and cartograms as many areas otherwise too small to be distinguished can be made visible and explicitly labelled, e.g., compare left and right images in Figure 8.

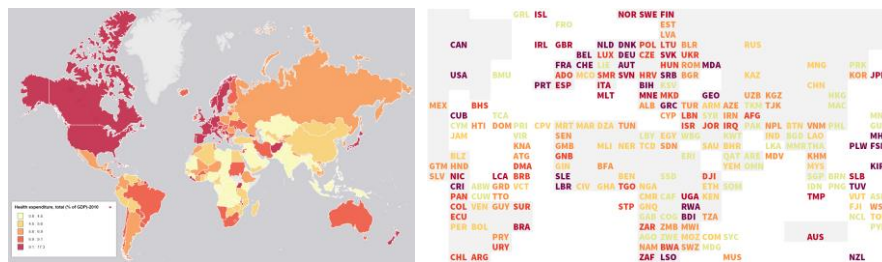


Fig. 14. Comparison of a choropleth map and the equivalent label-based cartogram. Left: The World Bank: Health expenditure, total (% of GDP) databank.worldbank.org, used with permission. Right image by authors.

Identification and location tasks measure responses. Seventeen participants used the two maps in Figure 14 in a similar approach to the National Geographic map literacy survey (GfK 2006) and summarized in Table 4. The identification task required the viewer to identify two circled countries in each map type. The location task required the viewer to indicate the color of two named countries for each map type. Viewers were either graduate students from computer science or employed in the domain of computer software development. Given that the subjects were older than the 18-24 year olds in the National Geographic study, countries used for the tasks were less commonly referenced in popular media, such as Niger, Albania, Ivory Coast and Slovenia.

Overall, the ISO-labeled cartogram outperformed the choropleth map 2.2 times. This implies the textual mnemonic with relative positioning acted as a better aid for recall than the geographical outline with accurate positioning. The performance difference was significantly greater for the identification task. Possibly this

was due to the ISO-label acting as a mnemonic cue for identification, or perhaps it is a result of the small study size. A larger study would be desirable.

Table 4. Percent correct responses to map identification and location tasks.

Task and percent Correct Responses	Choropleth Map	ISO-labeled Cartogram	ISO Cartogram Outperformance
Identify circled country	15%	65%	4.4x
Locate country and indicate color	53%	85%	1.6x
Total	34%	75%	2.2x

Critiques from typographers and visualization experts raise different concerns. *Legibility* relates to the clear decoding of individual alphanumeric glyphs and is an issue related to type size. For example, fewer levels of weight will be discernable at small type sizes. *Readability* relates to the ease of reading type: encoding data proportionally along the length of a string (e.g., USA) may be more difficult to read than an attribute applied across full words. Related to readability is familiarity with common formats and typefaces: e.g., a switch between serif and sans serif may be unnoticeable given our familiarity with both whereas a switch to **Blackletter** may be more noticeable. *Intuitive encoding* helps the viewer decode a visualization when multiple variables are encoded. For example, using font weight to indicate population (Figure 8) or GDP (Figure 7) is an intuitive relationship between the magnitude of the font weight and a measurement of magnitude. Using an arbitrary combination of weight, slope and underline (Figure 11) is non-intuitive. *Long label bias* is raised by many critics – a few approaches have been presented here but not evaluated. Finally, one researcher noted that the use of type as a thematic attribute may require re-evaluation of attributes as outlined by Bertin.

5. Conclusion

At a high-level, we have provided various evidence regarding the potential for a new way of thinking about thematic maps based on textual labels. Over the last few decades, researchers have codified the mechanics of colorful choropleths and contorted cartograms. However, the historic record of the last 700 years has hinted at the hidden potential of typographic possibilities. We have organized these typographic attributes, applied them to thematic maps, and shown how they can lay out, scale up (to many attributes and/or many labels), handle long labels and even encode small ranges of quantitative values.

How far can thematic labels go? The many images herein are only the start. One obvious extension is interactivity (implied by the *Click Me* button in Figure 7) which suggests animation, zooming, filtering, hyperlinks and analytics at the click of a mouse. But labels aren't constrained to geographic entities: they can be

anything. Imagine a thematic map of local coffee shops labeled with the barista on duty and their ratings; or a version of Facebook with friends arranged by proximity and thematic attributes for current activities. There are many possibilities.

References

- Bertin, J. (1983). *Semiology of Graphics*, trans. William Berg, Madison, WI: University of Wisconsin Press.
- Brath, R., & Banissi, E. (2015). Evaluating lossiness and fidelity in information visualization. *SPIE/IS&T Electronic Imaging* (pp. 93970H-93970H). International Society for Optics and Photonics.
- Brath, R., & Banissi, E. (2016). Using Typography to Expand the Design Space of Data Visualization. *She Ji: The Journal of Design, Economics, and Innovation*, 2(1), 59-87.
- Brewer, C. (2005). *Designing Better Maps: A Guide for GIS Users*. ESRI Press.
- Chen, M., & Floridi, L. (2013). An analysis of information visualization. *Synthese*, 1-18.
- Cleveland, W. (1985). *The Elements of Graphing Data*, Summit, NJ: Hobart Press.
- Cuff, D. & Mattson, M. (1982). *Thematic Maps: Their Design and Production*. NY: Methuen.
- GfK (2006). Final Report National Geographic-Roper Public Affairs 2006 Geographic Literacy Study. <http://on.natgeo.com/QRp3aj>. Accessed: 29 Jan 2017.
- Healey, C. & Enns, J. (2012). Attention and Visual Memory in Visualization and Computer Graphics. *IEEE Transactions on Visualization and Computer Graphics*, 18(7), 1170-1188.
- Hodges, E.R.S. (1989). Cartography for the Scientific Illustrator. In D.G. Cole, (Ed.), *The Guild Handbook of Scientific Illustration*. NY: Van Nostrand Reinhold.
- Hodson, Y. (1999). *Popular maps: The Ordnance Survey Popular Edition One-Inch Map Of England And Wales, 1919-1926*. Charles Close Society.
- Iliinsky, N. (2012). Properties and Best Uses of Visual Encodings, in Complex Diagrams (blog), <http://complexdiagrams.com/properties>. Accessed: 29 Jan 2017.
- Krygier, J. (2005) *Making Maps: A Visual Guide to Map Design for GIS*. NY: Guildford Press.
- MacEachren, A. (1995). *How Maps Work: Representation, Visualization, and Design*, NY: Guildford Press.
- MacKinlay, J. (1986). Automating the Design of Graphical Presentations of Relational Information. *ACM Transactions On Graphics (TOG)*, 5(2): 110–141.
- Mazza, R. (2009). *Introduction to Information Visualization*. London: Springer.
- Muehlenhaus, I. (2014). *Web Cartography: Map Design for Interactive and Mobile Devices*. Boca Raton, FL: CRC Press.
- Munster, S. (1540). Italia XIII> Nova Tabula in *Geographia Universalis*,
- Munzner, T. (2015). *Visualization Analysis and Design*. Boca Raton, FL: CRC Press.
- Osley, A.S. (1969). *Meractor: A monograph on the lettering of maps, etc in the 16th century Netherlands, etc*. NY: Watson-Guption Publications.
- Stieler, A., & Haack H. (1909). *Atlas of Modern Geography*, 9th Ed. Gotha, Germany: Justus Perthes.
- Stieler, A., & Haack H. (1924). *Atlas of Modern Geography*, 10th Ed. Gotha, Germany: Justus Perthes.
- Tyner, J. (2010). *Principles of Map Design*. NY: Guildford Press.
- Ware, C. (1999) *Information Visualization: Perception for Design*, San Francisco: Morgan Kaufmann.
- Wilkinson, L. (1999). *The Grammar of Graphics*. London: Springer.
- Wolfe, J.M. & Horowitz, T.S. (2004). What Attributes Guide the Deployment of Visual Attention and How Do They Do It? *Nature Reviews Neuroscience* 5(6) 495–501.

