The compositional dot map: a visualization of spatial data

Keywords: Data visualisation, data exploration, spatial data.

1. INTRODUCTION

Many statistical themes are related to geography, for instance population, migration, health, economy, and environment. Visual exploration and presentation of the underlying spatial data helps to better understand these themes and to communicate the trends and patterns to the general public.

The dot map is thematic mapping technique where single data points are shown as dots. The main purpose of the dot map is to analyse the spatial distributions of the dots, which may represent events, persons, or companies. The exploratory use of the dot map is not particularly new, since the famous map by John Snow to analyse the 1854 cholera outbreak in London [1] is also a dot map, albeit with stacked bars instead of dots. With this map, he managed to find the cause of the cholera, which was a particular water pump.

The dots in a traditional dot map are coloured uniformly. A straightforward extension is therefore to colour the dots. In this way, it is not only possible to show data distributions, but also the compositions into multiple groups. However, it is not trivial how to colour the dots without having perceptually dominant colours. Another challenge is how to deal with different zoom levels. This is important, since many maps are interactive nowadays.

The Racial Dot Map [2] shows the ethnic composition of the US population. For each person, a dot is plotted that is coloured by the ethnic origin of that person. This interactive map clearly shows the racial segregation in the US. Although this map is aesthetically pleasing, the used colour codes and alpha transparency seem to have been chosen arbitrarily.

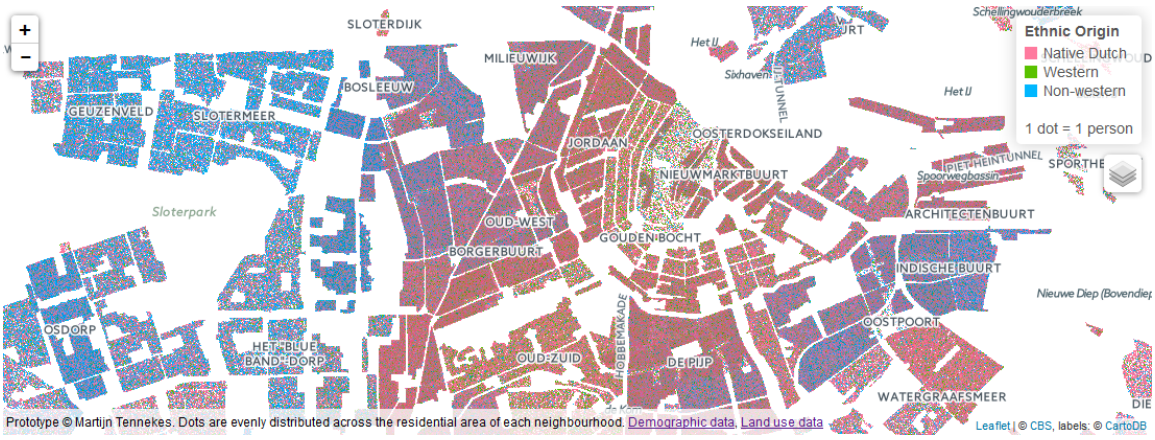

Figure 1. Interactive coloured dot map of the population of Amsterdam

We propose an algorithmic colour scheme for plotting compositional dot maps, which takes density, composition, as well as zoom level into account. We have applied this method to Dutch demographic data of ethnic origin on neighbourhood level. A screenshot of the prototype [3] is depicted in Figure 1.

2. METHODS

The main part of the interactive dot map is the creation of a tile server of bitmap images of 256 by 256 pixels. Tile servers are commonly used in interactive maps, including Google Maps, OpenStreetMap and Microsoft's Bing Maps. The bitmap images are arranged by zoom level, longitude, and latitude according to certain standards [4].

The location of the dots need to be specified. If the exact locations are not known, the dots can be generated at random per administrative area. The dots are assigned to pixels at a specific zoom level. Per pixel, the number of dots per category are counted. For lower zoom levels, these numbers can easily be summed since the zoom factor of one zoom is exactly 4, i.e., both height and width are increased by factor 2. Hence, we obtain counts per category for each pixel for various zoom levels. Notice that these counts per category contain two important values: the total, which represents the density, and the proportion, which represent the composition of the categories.

Next, the Hue Chroma Luminance (HCL) colour space is used to generate the colours. This colour space is useful for statistical graphics, since it is based on perceptual properties [5]. For instance, by varying the hue parameter, the luminance (brightness) and chroma (saturation) remain perceptually approximately constant.

The density is encoded with luminance; the darker the colour, the denser the dots. A dark pixel may therefore be interpreted as a stacking of dots, which is inevitable at lower zoom levels. The composition is encoded with hue and chroma. Each category is assigned a hue value. If multiple categories are present, the pixel hue is a mixture of these hue values. The chroma indicates the dominance of a category; if only one category is present, the chroma is maximal, whereas if all categories are equally present, the chroma is zero, which results in grey.

3. RESULTS

The method has been implemented in a prototype [3] that shows the ethnic origin of the Dutch population in three categories: native Dutch (pink), western (green) and non-western (blue). Many non-western people live in major cities, for instance Amsterdam (see Figure 1). Western people also live in major cities, but to a much lesser extent. They also live in areas near the borders, for instance in Vaals, in which the tripoint with Belgium and Germany is located. The composition of native and non-native Dutch people is clearly visible when zooming out. Figure 2 shows the Randstad, which is the urban area of the four major Dutch cities. It is clear that the dominant category in non-urban areas is native Dutch.

4. CONCLUSIONS

In our opinion, the compositional dot map is an effective and attractive mapping technique to show spatial density and composition. A couple of aspects need to be further developed. First, the assignment of the initial colours should be tested among users. Second, the method of selecting colours contains several parameters, which have to be tuned. Finally, the legend can be improved for lower zoom levels, such as depicted in Figure 2. Both composition and densities values should be reflected by the legend.

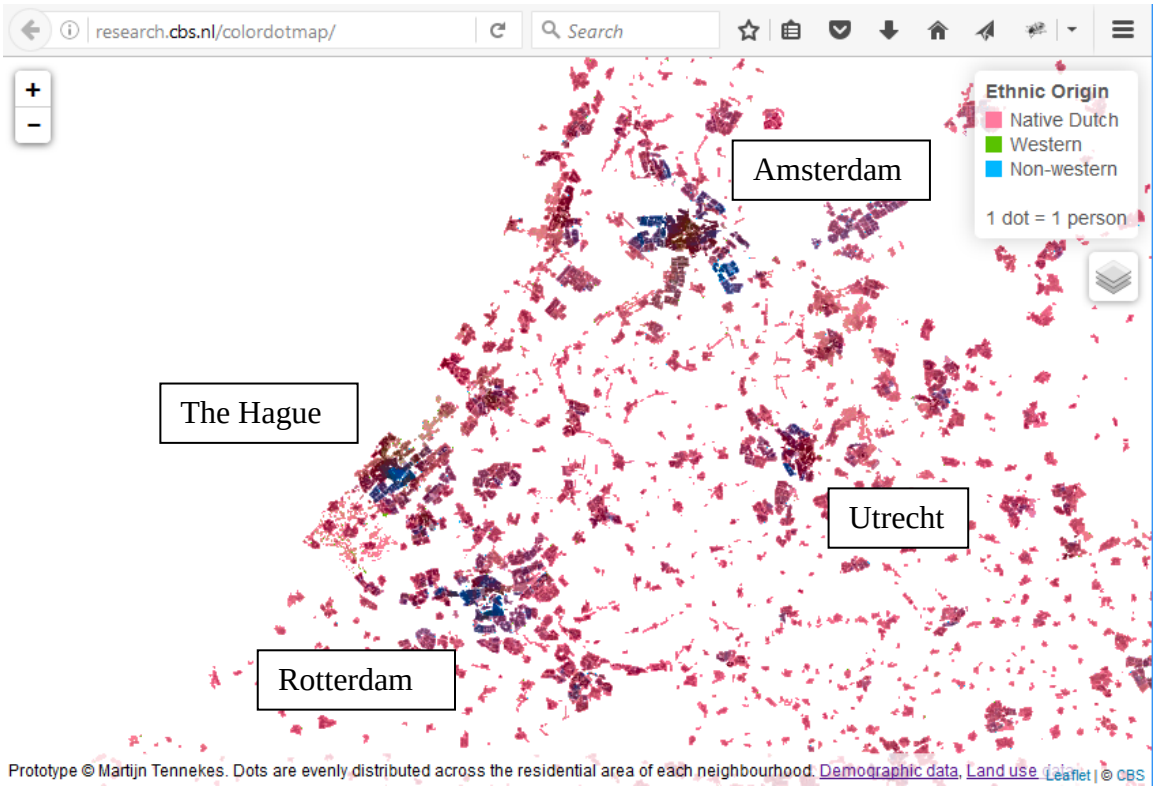

Figure 2. Interactive coloured dot map of the population of the Randstad

REFERENCES

[1] D. Cameron, I. G. Jones, John snow, the broad street pump and modern epidemiology, International journal of epidemiology (1983), 12-4, 393–396

[2] D. Cable, The Racial Dot Map (2013). URL: http://www.coopercenter.org/demographics/Racial-Dot-Map

[3] M. Tennekes, Prototype colored dot map (2016), URL: http://research.cbs.nl/colordotmap

[4] Klokan, P.P. (2008) Tiles à la Google Maps: Coordinates, Tile Bounds and Projection URL: http://www.maptiler.org/google-maps-coordinates-tile-bounds-projection/

[5] A. Zeileis, K. Hornik, P. Murrell, Escaping RGBland: Selecting colors for statistical graphics, Comput. Stat. Data Anal. (2009) 53-9, 3259–3270.