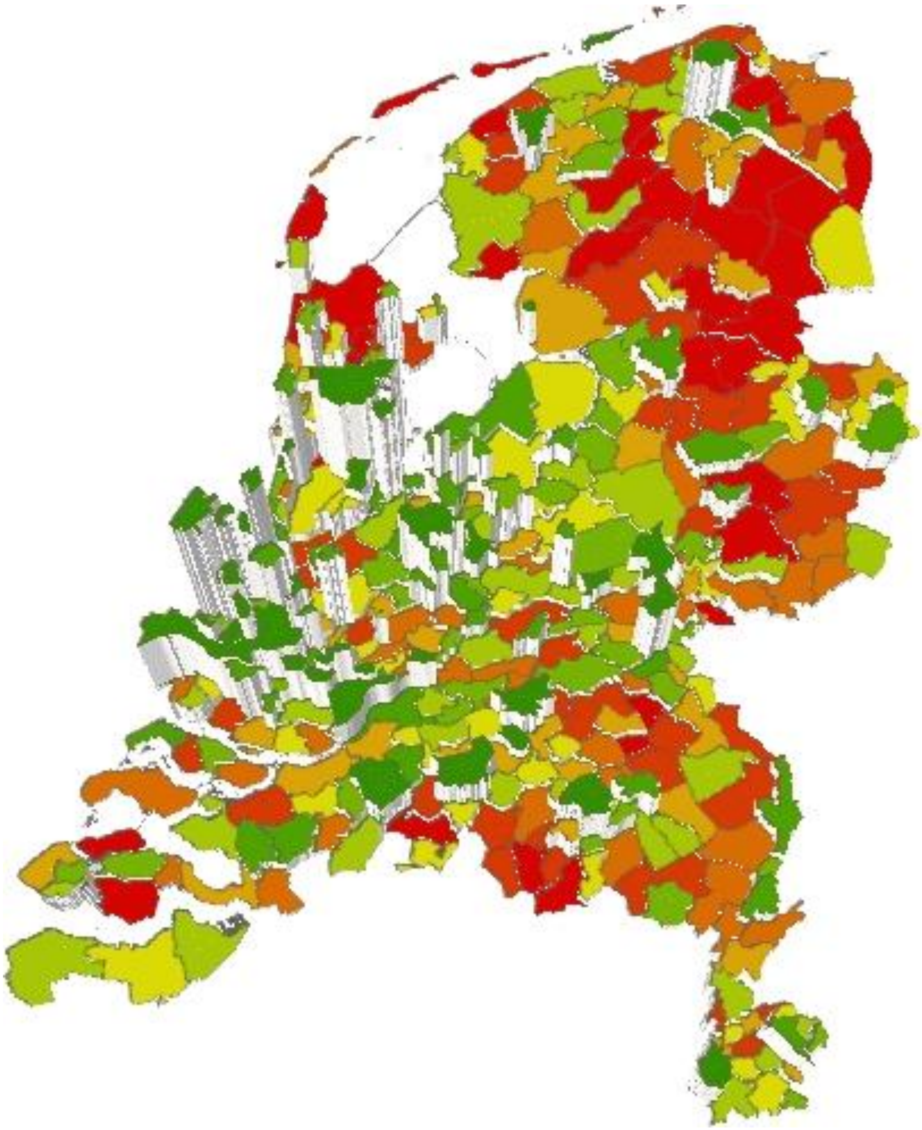
Centre for Geo-Information

Thesis Report GIRS-2013-15

The usability of 3D techniques for multivariate thematic mapping

Gert-Jan Boos

October 2013



WAGENINGEN UNIVERSITY
WAGENINGEN UR

The usability of 3D techniques for multivariate thematic mapping

Gert-Jan Boos

Registration number 87 01 28 099 100

Supervisors:

Dr.ir. Ron van Lammeren

A thesis submitted in partial fulfilment of the degree of Master of Science

at Wageningen University and Research Centre,

The Netherlands.

October 2013

Wageningen, The Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2013-15
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

PREFACE

My interest for 3D visualizations started when I worked with ESRI's CityEngine. This software allows to efficiently model a 3D urban environment using procedural rules. My goal was to learn more about this software and to write a tutorial that could be used by other MGI students. At the same time when I did this assignment as Capita Selecta I was already looking for a suitable thesis topic. When I was discussing CityEngine with Ron van Lammeren he asked me if this program could also be used for the creation of 3D thematic maps. We decided that this question was a good starting point for this thesis. Later when I went through the literature I decided that it would be more feasible to delimit the research and to focus on the 3D cartogram. This was necessary because the available knowledge on the use of 3D for multivariate thematic visualization was limited. The final research compares the 3D cartogram (as well as the cartogram and the value-by-alpha) with the choropleth technique.

I really liked working on the thesis because of my affinity with the subject of the research. In the starting phase, it was sometimes challenging to make the right decisions in order to define the topic. In most cases this was because there are so many interesting aspects to visualization. The time I needed to perform the research took longer than planned. This was partly because it took longer than expected to create the visualizations and the questionnaire. The creation of experimental visualizations such as the 3D cartogram is challenging because there are no ready to use solutions. However, it was interesting and I have learnt a lot about visualizations while working on this. I am very satisfied with the end results and I think that the questionnaire has produced interesting results.

I want to thank my supervisor Ron van Lammeren for the inspiration and valuable comments that I received during our meetings. I also want to thank the students in the thesis room for helping me by sharing their views on the topic during discussions. And I want to thank everybody who has completed the questionnaire.

Gert-Jan,

Wageningen, 18 October 2013

ABSTRACT

In general there is a lack of knowledge on the usability of new visualisation techniques. This is a missed opportunity for the visualization field because in practise the new techniques are not used (optimal). This research was primarily aimed to research the usability of 3D visualization for multivariate thematic mapping. This is done by making a comparison between the well-known choropleth and the value-by-alpha, cartogram, 3D cartogram techniques. Based on literature research a questionnaire was developed that was completed by 38 participants. The focus of this research was on the effectiveness (by measuring the percentage of correct answers and the time used to answer the questions) per visualization, in different situations (univariate, multivariate). The results give a clear picture of the usability of the different visualisation techniques in relation to each other. The results show that the choropleth is most effective for univariate thematic mapping. And that the effectiveness of the 3D cartogram was not significantly different for multivariate visualization. This shows the potential of the 3D cartogram, considering that the participants were more experienced in reading choropleth maps. The results are useful for mapmakers and researchers who are interested in multivariate thematic visualisations.

Keywords: Thematic mapping | Visualization | 3D Cartogram | Usability | Choropleth | Cartogram

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Background

A traditional thematic map represents the distribution of statistical data on a base map using colours, textures, lines or symbols to communicate the underlying values of the variables. A map can be considered thematic if the creator appeared to focus on the distribution of particular social or physical phenomena (Kessler and Slocum 2011). The majority of thematic maps can be divided into the following categories: cartogram, choropleth, dot, dasymetric, flow, isarithm, prism/fishnet, proportional symbol, land use and geologic (Kessler and Slocum 2011). Within most types of thematic maps the geometry is fixed. For cartograms the geometry is distorted to let the area of the features represent the data values. In prism maps and 3D cartograms the features are extruded in height to represent the values of the underlying variable. There are many different types and variations of thematic maps. However there is no readymade method to create the perfect visualization for a specific dataset. The type of thematic map that fits best to a dataset is not only determined by the data itself but also depends on other factors such as the story to be communicated or what patterns within the data the map has to reveal. Therefore the context and aim of a map are important factors to know before designing a thematic map. Trial and error is required for the creation of a thematic map, which is a dynamic process of propositions and verification (Morita 2011).

The most simple way to visualize thematic data is to use a two dimensional (2D) map with symbols or colours that represent the data values (choropleth) (Slocum, McMaster et al. 2009). Most of the thematic maps are choropleth or proportional symbol maps. There are guidelines that can be used to create an effective map. The well-known cartographer Bertin has provided a framework to enable people to make clear (2d) maps (Bertin 1983). Bertin's work can be divided into three parts. He recognizes cartography (and visualizations) as a dynamic graphic information system: making a map, manipulating a map, and communicating a map (Morita 2011). Many cartographers continued working on this guidelines and made improvements and additions. For example, the book 'Thematic Cartography and Visualization' from (Slocum, McMaster et al. 2009) is used by professional cartographers. For the more common 2D thematic mapping methods, like the choropleth map, there are many guidelines for example on the effects of the usage of different colour schemes.

There are also more complex approaches like the cartogram. A cartogram is a popular and effective instrument for visualization (Roth, Woodruff et al. 2010). Instead of using symbols or colours the cartogram distorts the shape of geographic regions so that the area directly represents a data variable (Heer, Bostock et al. 2010). The downside of this technique is that the changes in geometry can lead to difficulties for interpretation. Another interesting development is the 'value-by-alpha' mapping technique, where an overlaid alpha layer is used to influence the opacity of the underlying objects and thereby communicates the data. This visualization type is able to circumvent the compromise that comes with the cartogram form, perfectly equalizing the base map while preserving both shape and topology. The value-by-alpha map therefore can always be considered as a multivariate map, with a layer of interest and an equalization layer. However, the value-by-alpha map is not without its own limitations. In some maps this method can lead to lots of dead space (black or dark areas), when the higher values are concentrated in a relatively small area (Roth, Woodruff et al. 2010).

Multivariate maps show multiple data variables at the same time. Within this research a map is defined multivariate if more than one input data variable is visualized. Traditional multivariate maps often use combinations of colours, textures and symbols to map the different variables (Bertin 1983). Choropleth maps can display two variables using a bivariate colour scheme or more variables using multivariate colour schemes. However, in practise often two separate choropleth maps are used besides each other to communicate two variables. There are also more complex methods like using Chernoff Faces as symbols on the map (Flury and Riedwyl 1981). In this method multivariate data are represented graphically by faces. Different parameters of the faces such as eye size, pupil size are used.

One of the conclusions of Kessler and Slocum (2011) is that in the past the limited technology was an important limitation for the quality of the designed maps, while today they are largely a function of limited knowledge of tools and map design and the costs of using professional cartographic services. The evaluation of the maps in the research is done by comparing maps from two journals divided in different periods of time, from the year 1900 till 2000, against a list of design guidelines. This research also showed that the traditional choropleth map has become popular among geographers and according to this research it was used in more than 25% of the articles by the year 2000. The rapid developments in the field of 3D visualization offer new tools that can be used to develop new types of visualisations (Joris 2012). Especially for the newly available tools and techniques that use 3D the knowledge is limited. There are examples of 3D cartograms where the use of height is introduced to represent values. However these cartograms often are in fact 2.5D, where each object has only one height value. For multivariate mapping one could reason that 3D visualization offers opportunities to include more variables then 2D visualization considering the additional height axis (Z-axis). When 3D visualization is considered from a broader view it allows to create visualizations that can be considered as alternative or unusual. An important conclusion of (Kraak 2003) is that 'The use of alternative, if not unusual, graphics representations stimulate the visual thinking. These graphics can reveal patterns that are not necessarily visible when traditional map display methods are applied'.

1.2 Introduction of 3D thematic visualizations

In Figure 1 the process of the communication of thematic data is shown in order to provide a better understanding of the potential role of 3D within the field of thematic visualization. Thematic maps are different from general (topographic) maps because they do not directly represent objects, shapes or terrain that exist in the real world. The extra degrees of freedom 3D offers, in the form of a Z-axis, can be used to express the values of the thematic variables instead for example the height of the objects. Because thematic data essentially is 2D the final interpretation should also be 2D, regardless of the visualization method. In case of 3D visualization the thematic data will be transformed into a 3D thematic visualization. This visualization then will be published on a 2D computer screen or printed map, since 3D screens at this moment are not widely adopted. Although the viewer will still have a 3D visual perception through a variety of monocular depth cues (perspective, lightning and shading, occlusion). Finally when the user recognizes and understands the visualization, the process of interpretation starts and the user starts to understand the (2D) thematic data.

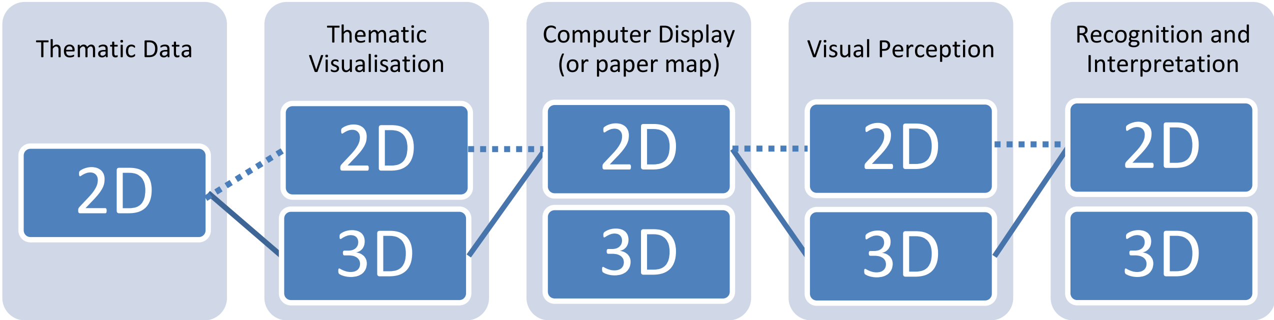


Figure 1 – Schematic visualisation of the difference between using 2D and 3D visualizations for thematic mapping. The dashed lines shows the path for 2D visualisations (such as choropleth). The solid lines show the path that is used for 3D visualisations, where different transformations between 2D and 3D occur.

1.3 Problem Definition

It is still an open question if, and under what conditions, 3D techniques are more effective than 2D techniques (Oulasvirta, Estlander et al. 2009). Within the field of geo-visualization one of the challenges is to improve the visualization of spatial thematic data. In regard of multivariate mapping the challenge is to increase the amount of incorporated variables without compromising the clarity of the map. It is plausible that 3D may offer advantages for multivariate mapping. Therefore there is considerable interest in the use of 3D techniques and to find new types of visualizations. The possibilities for 3D visualization are increasing with the technological developments. Many GIS professionals and cartographers are 'experimenting' actively with 3D and interactive thematic maps. For example they provide thematic layers for Google Earth to create 3D cartograms. There are many discussions online between professionals where different opinions are heard. Some think that 3D thematic maps are mostly 'eye candy' and that they are less effective than 3D maps. Arguments they provide often are about new problems the use of 3D may introduce. For example the occurrence of occlusion, where some the visibility of one object is blocked by another. Other cartographers are more optimistic and are claiming that this can be solved by using interactive interfaces. Despite the attention for the use of 3D in thematic mapping among professionals there is almost no information available within scientific literature. This study will explore the possibilities the use of 3D tools offers for visualization of multivariate spatial thematic data.

Existing Cartographic theories such as the guidelines from Bertin and Slocum (Bertin 1983), (Slocum, McMaster et al. 2009) are important for the creation of maps and visualizations. The guidelines are developed for 2D cartography and thus not offering a complete framework for the creation of 3D visualizations. Still many principles will be valid for the creation of 3D visualizations such as the use of colours and colour schemes. There is a challenge to complete the cartographic knowledge and guidelines with regard to the use of 3D. The choropleth map is the most commonly used visualization type for thematic statistics (Kessler and Slocum 2011). This visualization type is highly regarded among cartographers and is often used as comparison and benchmark within discussions about the effectiveness of new and experimental visualization types, including the use of 3D. However, we should not ignore the fact that the choropleth map has its own weaknesses. Choropleth maps use data or statistics that are aggregated to (administrative) areas. A result is the loss of information of the distribution of the mapped phenomena within these areas. The bigger the used units are the more information and patterns are lost. On the other hand if the used units are too small it can become really difficult to recognize them which is compromising the readability of the map. Another problem is that on the junctions between areas the colour abruptly changes which suggests that something abruptly changes in that area which in reality may not be the case. Finally the choropleth map offers limited possibilities for multivariate mapping. In practise there are many examples of choropleth maps that use bivariate colour schemes. However within maps using bivariate and especially multivariate colour schemes the readability may become compromised. In chapter two the choropleth map and other common maps will be reviewed into more detail.

The problem definition behind this research is best understood when taken in consideration that it is plausible that visualizations do conditionally benefit from the use of 3D techniques. And at the same time taken into consideration that the highly regarded choropleth map still leaves room for improvement. This research will explore the usability of 3D for thematic mapping and it will explore an alternative for the traditional choropleth map, the 3D cartogram. Thereafter the usability of the 2D and 3D visualizations will be tested among a group of users by measuring the accuracy and efficiency of the visualizations regarding to information transmission. And by asking for the participants preferences Finally the usability of the visualizations will be discussed.

1.4 Research Objectives

The general objective of this research is to 'Explore the usability of 3D techniques for multivariate thematic mapping'. The question behind this objective is if multivariate thematic visualizations could benefit from the use of 3D. The hypothesis is that in some cases multivariate thematic maps are more effective if 3D is used. Theoretically the extra dimension can be used to communicate more variables.

The following research questions are derived from the general objective:

RQ 1. What are the current (3D and 2D) thematic visualization techniques?

RQ 2. How to develop multivariate thematic 3D visualizations (and their 2D equivalents)?

RQ 3. What is the usability of the developed visualizations (3D versus 2D)?

1.5 Thesis Outline

The following chapters of this report will answer the research questions from paragraph 1.4. Research question 1 will be answered within chapter 2, by reviewing a selection of visualization types. In chapter 3 the used methods for the test procedure, the development of the questionnaire and visualizations as well as for the data analysis are explained. Also the materials that are used, such as software and data sources, are described. This chapter the answer to research question 2. Chapter 4 presents the results of the questionnaire in the form of descriptive statistics and hypothesis testing as well as explanation of the results. Chapter 5 discussed the results and provides conclusions on the usability of the different visualizations and thereby answers research question 3. Finally recommendations and a reflection on the research are given.

2. REVIEW

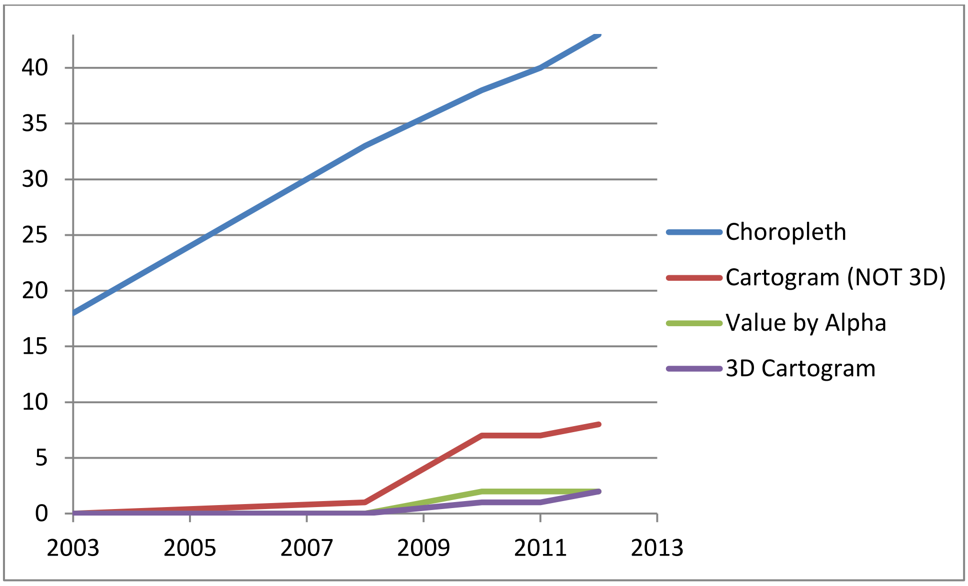


Figure 2 – The total (cumulative) amount of peer reviewed papers published in scientific journals on each mapping technique. Numbers are retrieved using the Wageningen UR online library global search function that has access to Taylor & Frances Online Journals, Social Sciences Citation Index, SciVerse ScienceDirect and others.

The aim of this chapter is to review the conventional and the experimental mapping techniques and to determine their strengths and weaknesses. Additionally, more generally applicable cartographic and visualization theory is reviewed. The fact that there is much more knowledge and experience with choropleth mapping then with other techniques is demonstrated in Figure 2. This figure shows that there are about 20 times scientific publications for choropleth mapping then for other mapping techniques. Nevertheless, it would be unfair to only make a comparison between choropleth and the 3D cartogram, because there are already additional techniques available, such as the cartogram and the value-by-alpha map, that tackle some weaknesses of the choropleth technique. Therefore the choropleth, cartogram and value-by-alpha maps are reviewed as well as the 3D cartogram. These products are all suitable for the communication of the same kind of thematic data as the choropleth. The 3D cartogram is the most used 3D technique within thematic mapping, and there are many examples available on online discussion boards and cartographer's weblogs. Still the number of scientific publications on this technique is small (Figure 2).

2.1 Choropleth map

2.1.1 Introduction

The choropleth technique, shown in Figure 3, is the most common technique for thematic maps. This map uses geographic areas (often administrative units) with shadings (colours) that communicate the values associated with the areas. Within most choropleth maps the data is classified in such way that each class represents a fair amount of data values.

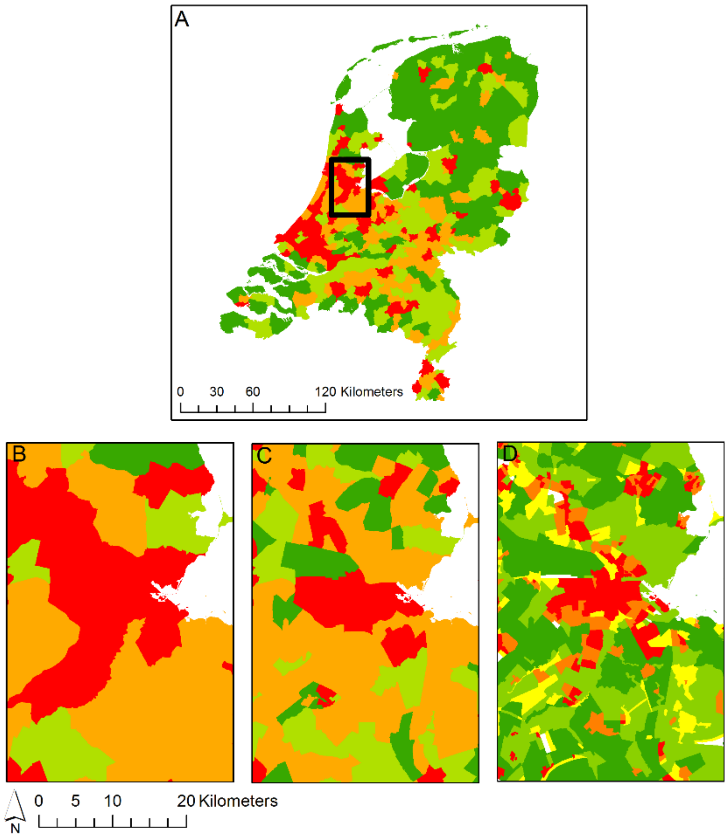


Figure 3 – This illustration shows the population density of the Netherlands on municipality level (A). The maps below are zoomed in on the city of Amsterdam. Map B still uses the municipality level as unit. Maps C and D use smaller units (districts and neighbourhoods).

2.1.2 Thematic mapping

While most choropleth maps use classification to improve the readability, there are also unclassed, i.e. continuous, choropleth maps. One of the issues with choropleth is that (continuous) shading makes the (exact) values of the areas difficult to read. Since users can only distinguish a limited amount of shades or classes. The downside of classification is that it generally leads to a loss of information. Within the article of Egbert and Slocum (1992) it is stated that 'from a mathematical standpoint, unclassed maps are ideal because they eliminate error due to data classification, but cartographers have never been able to agree on the perceptual merits of this approach'. Within Figure 4 in each map (A,B,C) the same information is visualized using different class breaks. The natural breaks method (A) tries to find natural groupings within the data and to group similar values. The quantile method (C) provides maximal contrast, because it assigns the same number of units to each class. Since the population distribution is skewed, the standard deviation method (B) assigns most values to the lower classes. If the distribution of the values is skewed it is often necessary to use one of those classification methods, because the use of an equal interval would assign too much values to the same class. Despite the fact that maps (A,B,C) are based on the same data (and are all valid) they provide completely different pictures. If an unclassed choropleth map would have been used, the users would have faced troubles because they can only distinguish a limited amount of different shading.

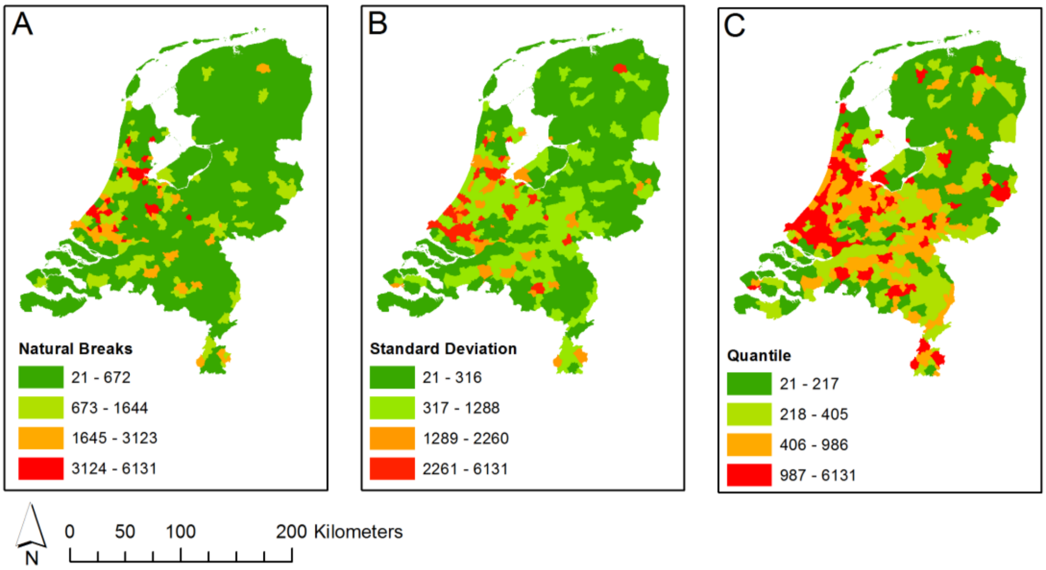


Figure 4 - Choropleth maps showing the population density of the Netherlands (inhabitants per 1000m2). In each map a different classification method is used.

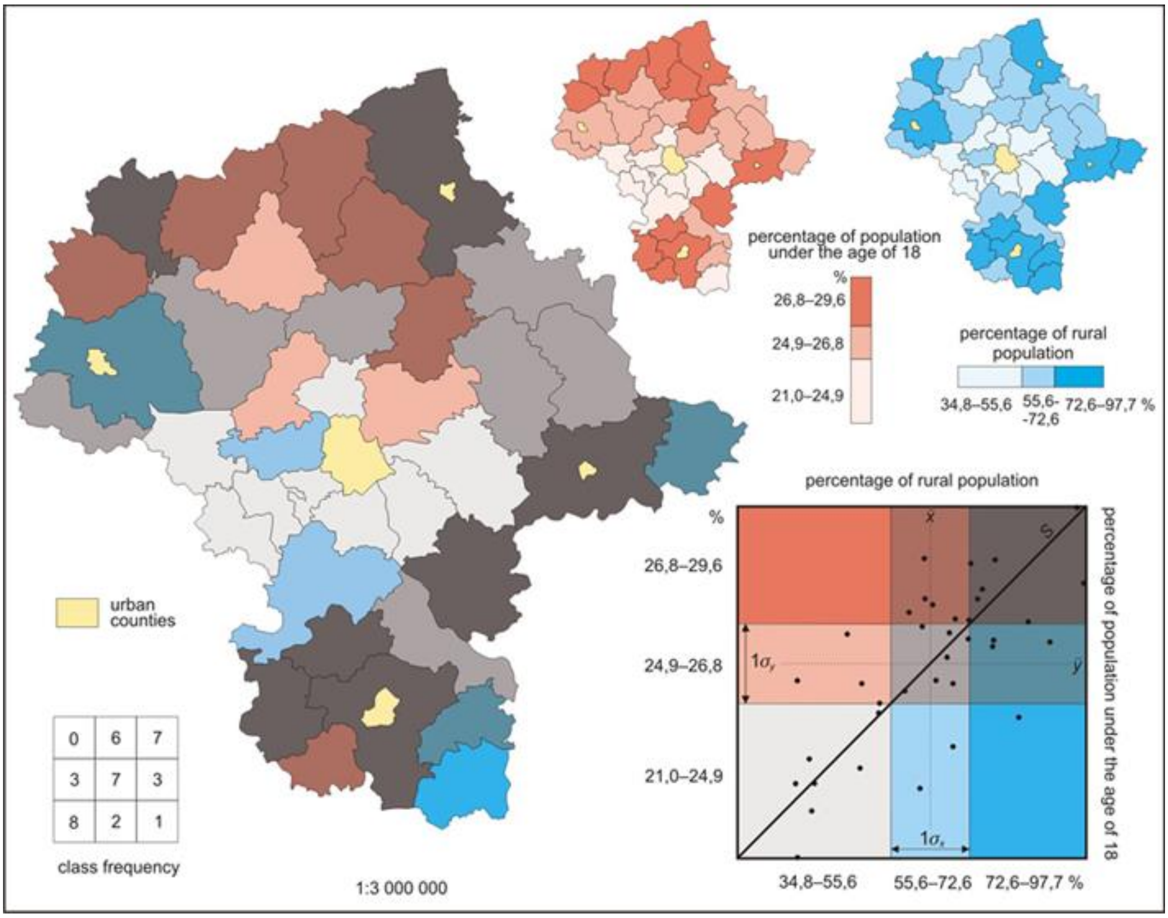


Figure 5 - (Leonowicz 2006) - Illustration of an bivariate choropleth map showing the percentage of population under the age of 18 and the percentage of rural population using an bivariate colour scheme.

2.1.3 Multivariate mapping

The choropleth allows to create bivariate or multivariate thematic maps by using a bivariate or multivariate colour scheme. This allows to visualize and to compare two separate phenomena simultaneously. In practice, multivariate maps often use the choropleth technique combined with other techniques (such as symbols) to incorporate more variables. Within Figure 5 an example is shown of a bivariate choropleth map. The research of Leonowicz (2006) showed that these maps are only effective when a limited number of classes is used and they are properly designed. This research also concludes that univariate choropleth maps are more effective for the visualization of spatial distributions within a single variable. And that bivariate choropleth maps are more effective in the communication of spatial relationships between variables. Another solution is to use the choropleth technique in combination with other techniques, such as symbols and textures. The most simple, but popular, solution for multivariate mapping with choropleth is to display multiple maps (with different variables) next to each other.

2.1.4 Geometric characteristics

The most important limitation is that choropleth uses bounded areas in order to visualize data. The spatial resolution of thematic data is often limited to the size of the smallest administrative units, for example neighbourhoods. Within large scale choropleth maps, data is often aggregated to larger and more recognizable units such as municipalities or provinces. Information about the distribution of the values within the used units is not visible on a choropleth map. This is especially a trade-off if information is available on a high resolution, for example on building, street or neighbourhood level. And with technical developments data becomes more and more detailed. Within Figure 3 A an example of a conventional choropleth with large units is provided. The use of large units makes it possible, for someone who has basic geographic knowledge of the Netherlands, to recognize specific municipalities within the map. Within Figure 3 B the same administrative units are used, but the map is focussed on a smaller part of the Netherlands (Amsterdam). When the same map is created based on smaller units, as demonstrated in Figure 3 C & D, different patterns become visible. This shows that the used (administrative) units have a strong influence on the appearance of the map. The use of administrative units for mapping is arguable because thematic phenomena often are not strongly related to these units. Also the boundaries between the areas tend to attract visual attention (Stewart and Kennelly 2010). When comparing Figure 3 A & D it becomes clear that most boundaries in map A are not representing (changes) in reality.

2.1.5 Temporal mapping

Animated choropleth maps can be used to show temporal changes. The research of Harrower (2007) concludes that it is more difficult to create a (fixed) classification scheme that is optimal for every period of time within an animated map. Therefore they suggest that for temporal animations it is better to use unclassed choropleth maps. Another approach for temporal choropleth mapping is to create a different map for each period in time and to display them next to each other.

2.2 Cartogram

2.2.1 Introduction

The cartogram is not as widely used as the choropleth map but yet well known among cartographers. Instead of the shading the cartogram uses the object's area to communicate the values. There are different kind of cartograms, such as the continuous (shown in Figure 6) and non-contiguous cartogram and the Dorling cartogram. The contiguous cartogram is only suitable for quantitative data. For this research the continuous cartogram is used because these maps are comparable with choropleth maps and thereby is suitable for the usability test among users.

2.2.2 Thematic mapping

The cartogram is a response to the issue of the choropleth map where the values of the variables in some cases are difficult to read because of the shading. To improve the readability of choropleth maps the data is classified, which leads to loss of information. The use of area size (shape) to communicate the values conveys the information more intuitively (Sun and Li 2010). Cartograms are most effective if the intended goal is to emphasize a small number of geographic disparities (Roth, Woodruff et al. 2010).

2.2.3 Multivariate mapping

There is no tradition of using the cartogram for multivariate mapping. Most cartograms are used to visualize an isolated variable, such as population size per country. However the continuous cartogram sometimes uses multiple variables, for example the population size of each country can be used to change the countries area accordingly, while another variable is used for the shading. It is impossible to preserve the edge length ratios of the polygons and the angles.

2.2.4 Geometric characteristics

The concept of the cartogram is to communicate data by disturbing the geometry. As a result it can be hard to interpret them without additional information to help the user to locate towns and cities (Sun and Li 2010). This is especially the case when the visualized variable is distributed unequally among the area. Therefore cartogram algorithms are optimized to preserve basic properties, such as shape, orientation and contiguity as much as possible (Keim, North et al. 2004). However the user has to be familiar with the original shapes of the geometry in order to interpret the cartogram.

2.2.5 Temporal mapping

There are algorithms available that allow for the creation of animated contiguous cartograms. Such animations can reveal more information than is revealed by only a few selected cartogram snapshots (Min and Revesz 2000). The scientific knowledge on the usability of these visualizations is rather limited.

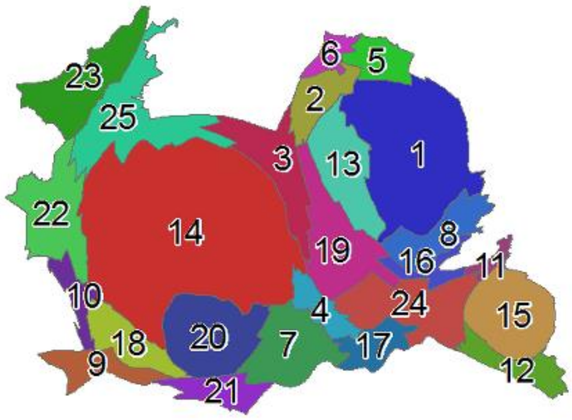


Figure 6 – Contiguous cartogram showing the population of the province of Utrecht per municipality.

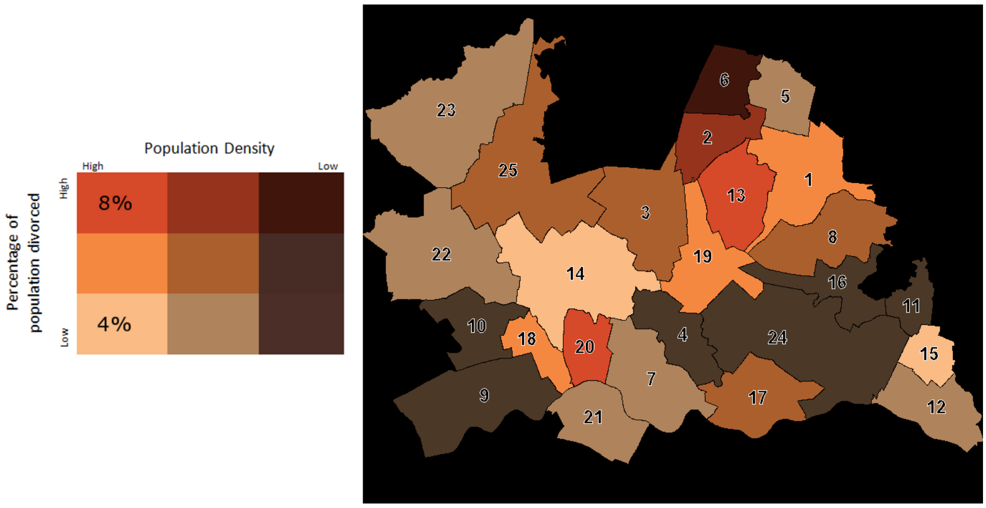


Figure 7 – Value-by-alpha example

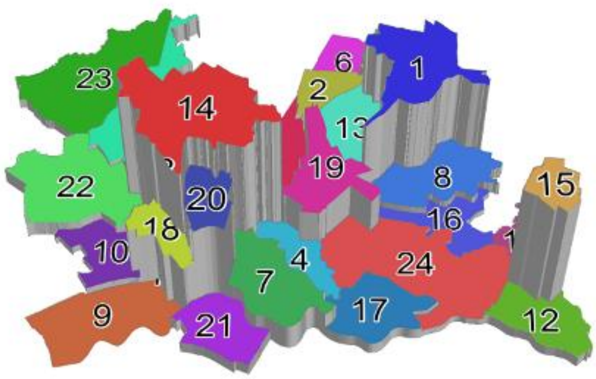


Figure 8 - 3D Cartogram

2.3 Value-by-alpha

2.3.1 Introduction

The value-by-alpha map (shown in Figure 7) is similar to the choropleth map, but in addition uses an alpha layer to communicate the values of the equalization layer. This causes enumeration units with low values in the equalizing variable to dissolve into the background. This disappearance produces a 'spotlight' effect over areas with high values in the equalizing variable, leading the map reader to focus immediately on these areas (Roth, Woodruff et al. 2010).

2.3.2 Thematic mapping

The variable visualized with the value-by-alpha layer has to be classified and therefore it is not possible to read the exact value from the final visualization. With the value-by-alpha it is not possible for users to make numerical level estimations about the equalizing variable. The value-by-alpha map is most suitable for providing a complete understanding of complex and subtle patterns.

2.3.3 Multivariate mapping

The value-by-alpha map is suitable for the communication of one variable for the colour shading (the variable of interest) and another for normalization using the alpha layer (Roth, Woodruff et al. 2010). Therefore the value-by-alpha map can never be used for univariate mapping.

2.3.4 Geometric characteristics

Within the value-by-alpha map the shape and topology is preserved. The value-by-alpha map can be used for smaller enumeration units then the cartogram. However enumeration units with small areas may be still difficult to read. On a map with small enumeration units there can be a lot of dead space (dark) because the de-emphasized units will still take up place. The value-by-alpha map is a good alternative for the cartogram when the preservation of shape and geometry is required.

2.3.5 Temporal mapping

There are no examples to be found where animation is used within the value-by-alpha map.

2.4 3D Cartogram

2.4.1 Introduction

The 3D cartogram (shown in Figure 8) and prism maps are basically the same product. The prism map uses a 3D earth globe, such as Google Earth, to project the visualization on. A 3D cartogram uses a projection to transform the map to a plane, as with the normal choropleth. The maps are comparable to choropleth maps but instead of shading they use the enumeration units' height to communicate the values. The prism map is used exclusively with interactive user interfaces, for example on desktop computers, while the 3D cartogram is also usable with a static interface. Both products are experimental, since at this moment there is not much scientific knowledge for these mapping techniques.

2.4.2 Thematic mapping

Extremely high values within the data are easier to notice on a 3D cartogram then on a normal choropleth because they will be shown as high spikes instead of a deeply coloured area. This is especially the case with small enumeration units. Other cartographers point out that users are not good at estimating the heights of the areas and comparing more subtle height differences. However this is basically the same problem as with unclassed choropleth maps, where users have difficulties with reading subtle differences in shading. The advantage of 3D cartograms is that there is no need for data classification and numerical data can be visualized in such a way that a value that is twice as big, also get an twice as big height (Slocum, McMaster et al. 2009). Shading and colour legend could support a univariate 3D cartogram because it is not possible to create a height legend. In interactive environments, especially for 3D prism maps, this is not really an problem because values could also be obtained by clicking on a feature Sandvik (2009 - 2013).

2.4.3 Multivariate mapping

The 3D cartogram is usable for multivariate communication when shading is used to communicate one variable and height for another. However this can be confusing when the user is used to univariate 3D cartogram maps. The research of Gastner and Newman (2004) shows us that there are many applications, mainly multivariate, where the cartogram provides advantages over the regular choropleth.

2.4.4 Geometric characteristics

Within the 3D cartogram it is hard to compare two areas that are far apart from each other especially when they are on the other side of the globe, this is discussed on the website of Sandvik (2009 - 2013). Also the possibility to use both height and shading to communicate the same variable is suggested (Slocum, McMaster et al. 2009). Other solutions are to label each enumeration unit with the exact value or to provide the user with two independent spinning globes with the same visualization. On the Sandvik (2009 - 2013) website it is also suggested that reading 3D prism maps is difficult because the user has to compensate for the curvature of the earth. Another problem with 3D cartograms is that in some cases occlusion can occur, where a lower area disappears behind a higher area. When using 3D the use of perspective can introduce problems, for example with some projections the scale is not constant, because large objects in the distance are rendered smaller then objects nearby. Using an orthogonal projection could be a solution, but leads to less natural looking visualizations because there is no perspective foreshortening.

2.4.5 Temporal mapping

There are examples of animated 3D cartograms. Prism maps can be shown within Google Earth together with an interactive slide bar where the visualized period can be adjusted.

2.5 Review conclusions

The choropleth is the most commonly used map and has proven to be suitable for thematic mapping. Even though the choropleth map is not difficult to create, the effectiveness is strongly depending on the used classification method and used colours, and thereby on the creator’s visualization skills. The unclassed choropleth maps are often difficult to read, because the user’s ability to distinguish colours and shades is limited. Within the value-by-alpha map the layer of interest and the equalization layer both also need classification. The cartogram and 3D cartogram both are able to visualize quantitative data without classification, by changing the shape (area or height) accordingly. However this is only the case for univariate visualizations, because when including a second variable it is still necessary to use shading. In case of univariate mapping the choropleth, cartogram and 3D cartogram techniques are usable. The value-by-alpha technique is only usable for multivariate visualizations, because it always needs two variables. The choropleth technique itself offers limited possibilities for

multivariate mapping. However in practise multivariate choropleth visualizations are often created by using multivariate colour schemes or by presenting multiple choropleth maps separately. Choropleth maps using multivariate colour schemes are difficult to read. The choropleth method can also be used in combination with for example symbols, but this is out of the scope of this research. The value-by-alpha map is only usable for bivariate mapping, where the variable of interest is communicated with shading and the normalization variable is communicated with the alpha layer. This is also possible within the cartogram. According to the literature the cartogram is most usable to emphasize a small number of disparities while the value-by-alpha is most usable for the communication of more complex and subtle patterns. For univariate mapping the choropleth is suitable for the communication of more complex and subtle patterns. Within both the choropleth and the value-by-alpha map the geometric aspects of the data are always fully maintained while the cartogram and 3D cartogram changes the geometric aspects in order to communicate the data. The cartogram is only readable for a user that is familiar with the normal geometry of the area.
All techniques are limited to the visualization of relatively large enumeration units and are not useful for the visualization of continuous data. In practice, most thematic data is already aggregated to (relatively large) administrative units before visualization. All techniques that are researched in this study are influenced by these enumeration units. Especially when large units are used, the averaging out effects of the units can lead to a significant loss of information and misinterpretation. When smaller enumeration units are used (if the data allows) the visualization can become difficult to read because recognition of the units becomes difficult, a base map or other reference data could be a useful addition. This is also the case for the value-by-alpha map, only here the alpha layer allows to use smaller (and even continuous) units then the shading layer. Temporal mapping for the choropleth, cartogram and the 3D cartogram is possible in the form of animations. In this case you need a computer of other device with a screen to view the animations. Otherwise separate snapshots can be presented and compared. For the value-by-alpha map no examples of temporal visualization can be found. There is not much scientific knowledge on the usability of the 3D cartogram for thematic mapping. The 3D cartogram has no need for classification of the (height) values, just as the area of the normal cartogram. The cartogram and 3D cartogram are both usable for univariate and bivariate visualizations when shading is used in addition. Even if in theory more variables could be incorporated. However it is questionable if this would be effective (shading with bi- or multivariate colour schemes and height). The advantage of the 3D cartogram over the normal cartogram is that the horizontal geometry is not changed, but at the same time the use of 3D introduces new problems. For the size of the enumeration units the requirements of the 3D cartogram are most likely comparable to those of the choropleth map.

3. MATERIAL AND METHODS

This chapter starts with a general explanation of how the research questions are answered. Thereafter detailed information is given on the used methods and materials. Prior to the writing of this chapter a literature research was conducted to answer research question 1. On the one hand, the strengths and weaknesses of the usual 2D thematic maps (choropleth, cartogram and value-by-alpha) needed to be studied. On the other hand, the possibilities of 3D for thematic mapping needed to be studied. Therefore, the literature research was an important part of this research. Important literature sources were cartographic textbooks (Slocum, McMaster et al. 2009), (Bertin 1983) and (Monmonier 1991) and scientific papers. The knowledge obtained from this research was used in order to develop and to test the (3D) visualizations. The results of the literature research are presented in chapter 2 of this report.

RQ 2. How to develop multivariate thematic 3D visualizations (and their 2D equivalents)?

RQ 3. What is the usability of the developed visualizations (3D versus 2D)?

(RQ 2.) The second part of this research was the development of the visualizations. Based on the literature the decision was made to test different map reading levels. First the users' ability to find (recognize) objects on the map, then the users' ability to read and compare values associated with the map units and finally the users' ability to recognize patterns and relations between variables in bivariate maps. For each map reading level equivalent visualizations are created using the different techniques. The developed visualizations are univariate and multivariate and are using the techniques described in chapter 2.

(RQ 3.) The usability of the developed visualizations was tested with the use of an online questionnaire. This was done by the determination of the accuracy of the visualizations, in other words by measuring the correctness of the participant's answers. And by the determination of the efficiency of the visualizations, by measuring the time the participants used to answer the questions. The test measures the correctness and times of the different mapping techniques in different applications. This information was used to assess the usability of the products. The accuracy was determined by the amount of correct answers per technique and application. Figure 9 provides an overview of the structure of the reports and the concepts that are used. A statistical analysis was performed to evaluate the results.

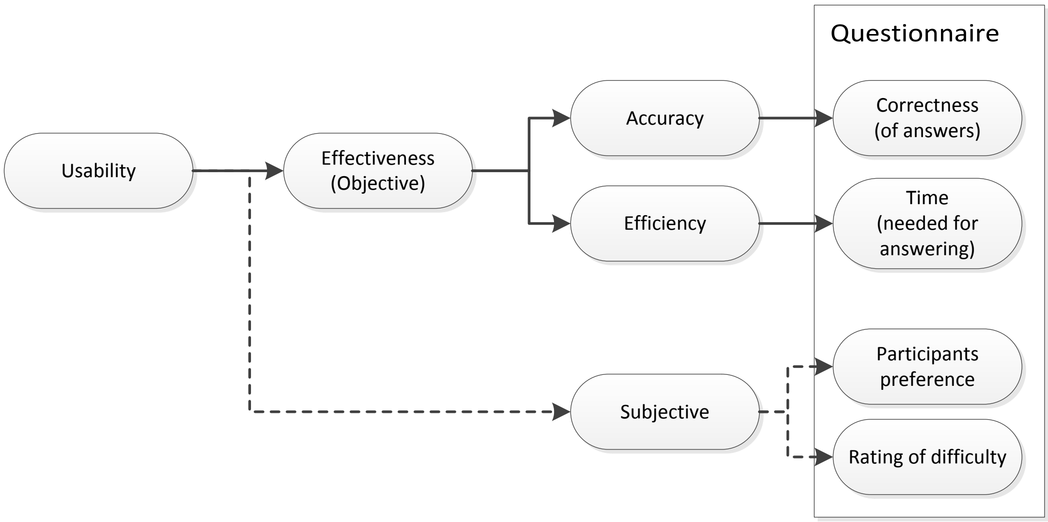


Figure 9 - Flowchart presenting an overview of the concepts used in this report.

3.1 Materials

For this research thematic maps using the choropleth, cartogram, value-by-alpha and 3D cartogram techniques had to be created. A schematic overview of this process is included in Appendix 3. The freely available municipality data from the CBS (Dutch Central Agency for Statistics) was used as the data source for the maps. The variables within a multivariate visualization have to be correlated (positively or negatively) to ensure that spatial patterns and relations are present. A similarity matrix was created based on the Pearson correlation coefficient. This matrix was used to select appropriate layers from within the CBS data (Appendix 1). The values of the correlation coefficient are ranging from -1 (perfect negative correlation) to 1 (perfect positive correlation), where 0 means uncorrelated. For the multivariate visualizations negative correlations with a correlation coefficient of approximately -0.6 are used. The 2D maps used in the test were developed using both ESRI's ArcMap and the Quantum GIS software. ArcMap offers more sophisticated pre-processing and visualisation options, such as labelling of the administrative units, while Quantum GIS offers a plugin that was used to create the cartogram maps. The creation of the 3D maps was a major challenge because no ready-made method or software implementation was available that met all the requirements. Eventually they were developed with ArcGIS (to export a texture image with shading and labels) and CityEngine from ESRI to extrude the data combined with the texture. The open source web based Lime Survey tool was used for the questionnaire, because it supported all the requirements and advanced functions (such as time measurements per question). The questionnaire was built with the help of the UsabilityNet (2013) website and the Dillman (2000) textbook. The results from the questionnaire were analysed using Microsoft Excel and SPSS statistical software.

3.2 Test procedure

The test was publicly accessible online, and active for eight days from the 26th of August till the 2nd of September 2013. In order to recruit participants from the MGI (Master Geo-Information Science), the private MGI Facebook group was used as well as the Wageningen University's internal mail and personal invitations. Other contacts were recruited using a general invitation on Facebook and LinkedIn, asking friends (and their friends and colleagues) and family. Both professional and inexperienced participants were needed in order to test the influence of experience on the correctness and times of the answers. The test was deliberately developed to be easy to use, also for less experienced participants. Participants could easily respond through closed answer questions and multiple choice options. These types of questions were deemed suitable for the measuring of accuracy and efficiency. And in this way a large sample group could be reached. In the final phase of the development of the questionnaire, the questions were tested on a small group of five participants. This allowed reconsidering, removing or improving of the questions that were unclear and evoked discussions. The method used for this pre-test was an in-depth interview of the five persons and a request for extensive comments as well as an exploration of their answers. This lead inter alia, to a fine tuning of some of the cartogram maps which turned out to be too difficult as well as the reconsideration of some guidance text that contained too much technical terms.

3.3 Test criteria

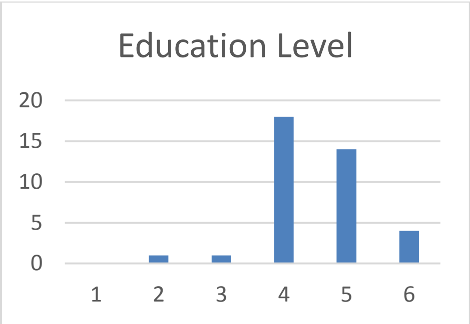
The test was focussed on the accuracy (correctness) and efficiency (time to answer) of the visualizations. In other words, the experiment tested the participants' abilities to deduce information from visualizations. Three question groups were created based on Bertin's map reading levels. First a group of questions (group 1) and corresponding visualizations was used to test the participants' ability to recognize individual mapping units (municipalities) within a map. The hypothesis behind these questions is that object recognition within a cartogram or 3D cartogram is more difficult than within a choropleth and value-by-alpha due to manipulation of the shapes. At the same time this question group was meant as a control group, since very few wrong answers were expected in this question group for the choropleth map. The second group of questions was aimed to test the users' ability to read and understand the values of the individual variables and to compare values of the different administrative units. The hypothesis behind these questions is that the effectiveness of visualization techniques regarding the communication of data values is different, because they use different ways to transmit information (shading, shapes and extrusion by height). The third question group was developed to test the participants' ability to read and understand relations and patterns between the different variables. Thus, this question group focusses on the effectiveness of the techniques for multivariate mapping. The hypothesis behind these questions is that the 3D cartogram, cartogram and the value-by-alpha techniques offer more opportunities for multivariate mapping than the choropleth technique. Within question group 1 and 2 for each visualization technique, one visualization and two questions were created. Within question group 3 for each visualization technique, one visualization and question was created. The final questions are included in Appendix 4. For each separate question the accuracy and efficiency was measured.

3.4 Data analysis

A schematic overview of the data analysis procedure is included in Appendix 2. The analysis of the final questionnaire was firstly performed in Microsoft Excel in order to obtain mostly descriptive statistics. Also an outlier removal was performed on the time measurements because some of the participants may have been distracted (or were taking a short break) while filling in the questionnaire. Therefore, the MAD (Median Absolute Deviation (Leys, Ley et al. 2013)) was calculated for each time measurement per question and any measurement with the absolute value of modified Z-score exceeding the threshold of 5 was considered as an 'outlier'. This threshold was chosen because it leads to a few outliers randomly distributed among the questions, while a stricter threshold indicated proportionally more outliers in the more difficult multivariate questions (which are most likely not actual outliers). Most questions and answers resulted in binary data (true or false) which asked for non-parametric tests. In this case the McNemar test for paired samples was used to compare the correctness of the answers of the different techniques. The Fisher's exact test for independent samples was used to compare the amount of correct answers between different subsets, such as a comparison between MGI and other participants. To confirm if the measured times were normally distributed a Shapiro-Wilk test was used. Because the times were not expected to be normally distributed, another non-parametric test was used to test if the (mean) time used to answer the questions differed between the different techniques. In this case the Wilcoxon Signed Ranks test was used because the time measurements are continuous data. After subdivision of the times into different subsets, such as MGI and other participants, the Fisher's Exact test for independent samples was used. The Fisher's Exact test is comparable to the Chi-square test, but is regarded to be more conservative and accurate. These statistics were used to test the hypotheses mentioned in the test criteria paragraph (3.3). Additionally to the measurements, the users were asked to indicate which visualization technique they preferred for the multivariate visualizations and to rate the difficulty (Likert scale). The preference was analysed by counting the number of votes per techniques. Additionally the difficulty ratings were averaged per technique and compared.

4. RESULTS

This chapter will first present an analysis of the questionnaire results using descriptive statistics and hypotheses testing. This results, and in particular the statistics, are accompanied by explanation where necessary. The results of the literature review are already reported at the end of the review chapter.



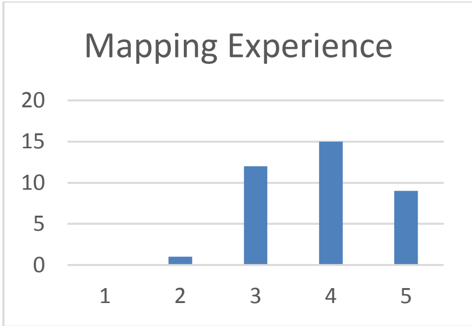


Figure 10 (left) - Education level of participants on a scale of 1 – 6, where 4 means bachelor. Figure 11 (right) - Mapping experience of participants on a scale of 1 - 5, where 3 means that the participant is reading maps occasionally.

4.1 The participants

The questionnaire is fully completed by 38 participants. The sample as a whole is relatively young (M=32, SD=13.25) and highly educated (Figure 10). Only two participants are not graduated for a bachelor diploma or higher. The majority of the participants (20 individuals) are students and staff from the MGI. They can be considered relatively experienced in reading maps. Only two participants are not reading maps occasionally while 20 participants are using maps frequently (Figure 11). There are more male participants (28) then female (10). On average the participants used 18.36 minutes to complete the questionnaire (SD=8.50). The answering of the questions related to the maps on average took 38 seconds per question (SD=20).

4.2 Simple (test) questions (group1)

All questions of the choropleth and value-by-alpha were answered correct within group 1 (M=1.00) (Figure 12). A McNemar test indicated that the 3D cartogram (M=0.92, p=0.031) and the cartogram (M=0.86, p=0.000) are less accurate than the choropleth. A Wilcoxon Signed-ranks test indicated that within question group 1 the questions on the 3D cartogram (M=21.00, Z=-1.658, p=0.097) and the cartogram (M=23.60, Z=-2.455, p=0.000) are answered slower than the choropleth (M=16.37). The time differences between the choropleth and the value-by-alpha (M=17.79) were tested equal. This was expected because the value-by-alpha works similar to a choropleth within group 1 with only simple univariate questions.

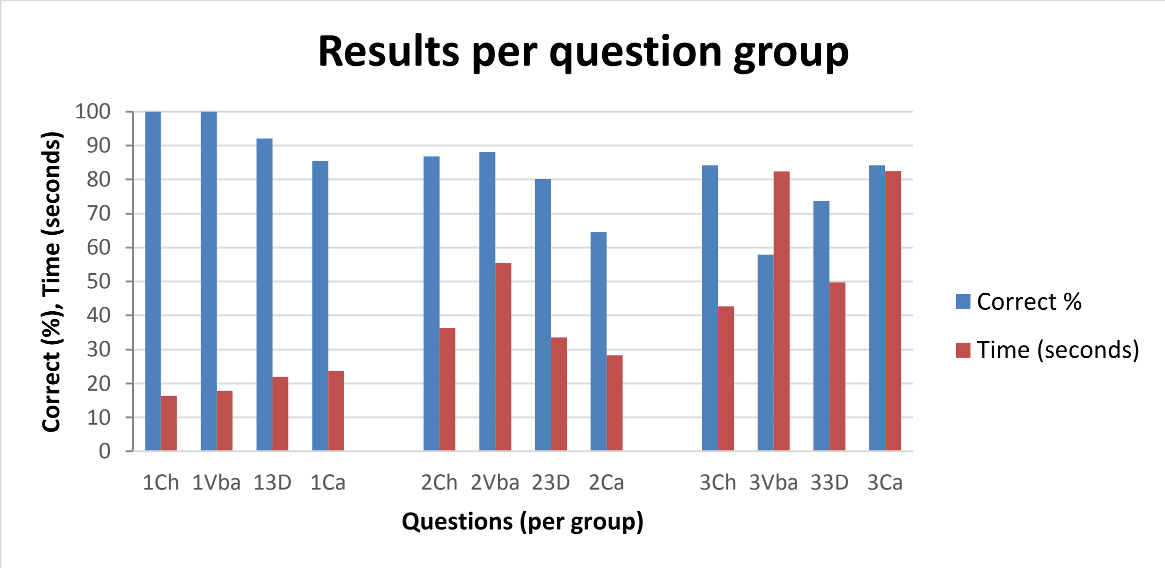


Figure 12 - Graph showing the percentage of correct answers and the average time per question for each product. The answers are sorted per question group (1=simple / test, 2=medium level, 3=multivariate).

4.3 Understanding of individual variables (group 2)

Within question group 2 the correctness of the answers of the value-by-alpha (M=0.88) and the 3D cartogram (M=0.80) tested similar to the choropleth (M=0.86). A McNemar test indicated that the answers of the cartogram (M=0.65, p=0.000) are less accurate than the choropleth within this question group. It is interesting to notice that the 3D cartogram performed better than the cartogram in the communication of univariate differences and values. The value-by-alpha was expected to perform similar to the choropleth in this question group because the questions where univariate (and the user could ignore the alpha layer). The time the participants used to answer the questions was tested equal for choropleth (M=36.35) and 3D cartogram (M=33.63). The value-by-alpha questions (M=55.51, Z=-3.816, p=0.000) were answered slower than the choropleth. While the cartogram questions were answered faster (M=0.005, Z=-2.779, p=0.005), which is notable because the questions are answered less accurate.

4.4 Understanding of multivariate relations (group 3)

Within question group 3 the correctness of the answers of the cartogram (M=0.84) and 3D cartogram (M=0.74) tested similar to the choropleth (M=0.84). In this question group the participants were expected to interpret both layers of the value-by-alpha map. A McNemar test indicated that the answers of the value-by-alpha (M=0.58, p=0.009) are less accurate than the choropleth in this question group. The time the participants used to answer the questions was not significantly different between the choropleth (M=42.72) and the 3D cartogram (M=49.75). The value-by-alpha questions (M=82.41, Z=-3.818, p=0.000) and the cartogram (M=82.48, Z=-4.210, p=0.000,) questions were answered slower than the choropleth.

4.5 Influence of participants spatial thinking skills

To discover the influence of the spatial thinking skills the participants are subdivided in two groups, MGI students / professors and other participants. Especially for the multivariate questions (question group 3) the correctness of the answers seems to be related to the participants experience and education (Figure 13). A Fischer's Exact Test showed that for the 3D cartogram the answers were tested significantly better (p=0.027) between MGI and other participants. There were no other significant differences regarding correctness and time among both groups.

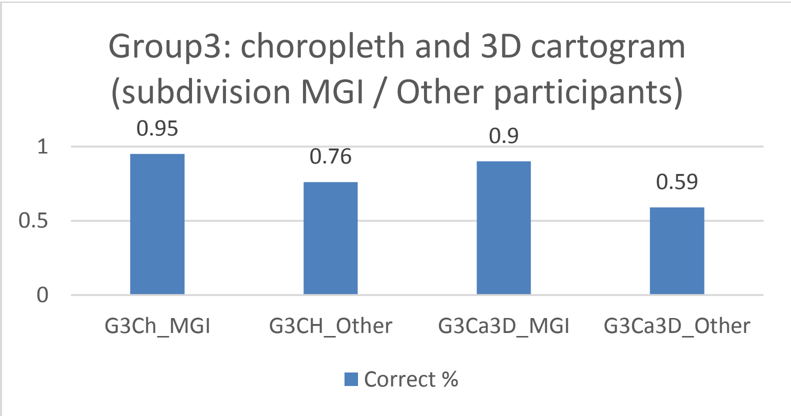


Figure 13 – Graph showing the differences in correctness of the answers for the choropleth and the 3D cartogram among the MGI and other participants.

4.6 Participants preference

As a concluding question the participants were asked which mapping technique they did prefer (Figure 14) for the questions in question group 3, in other words for the multivariate visualizations. Thereafter, they were asked to indicate how difficult the answering of the multivariate question using their preferred technique was. Most participants preferred the common choropleth technique. However the 3D cartogram still gets a considerable part of the votes. The other products were clearly less popular. The most preferred products for multivariate mapping also scored the best in this question group, with respect to the accuracy and the efficiency. There were no significant differences in the participants rating of difficulty between the preferred products. On average the difficulty was rated 2.42 (SD=0.76) on a scale of 1 to 5, which indicates that the difficulty of the questions is balanced.

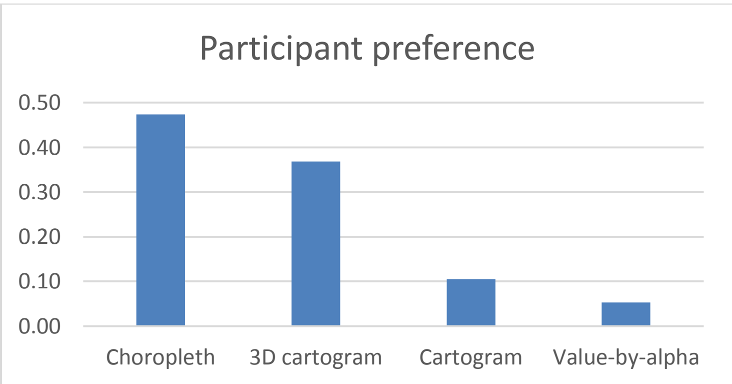


Figure 14 – The participant preference for the different products (in percentage).

4.7 Participants comments

At the end of the questionnaire the participants were able to insert their comments into a feedback field. Some participants were complaining on the fact that for most question answering was mandatory and that there was no ‘no answer’ option. However, this was part of the design (the participants were asked to answer intuitively), but may have weakened the outcomes by introducing random good answers. Two people reported to have colour deficiency and were not able to distinguish all units, which could have been prevented by using different colour schemes (However, in reality this problem also often occurs with choropleth). Another participant commented that the choice for using municipal demographic data may have been in favour for the choropleth technique. However, this research was aimed to test these products in a scenario of multivariate thematic mapping, so using demographic data bounded to administrative units seems to be the right choice.

5. DISCUSSION AND CONCLUSIONS

This chapter will provide an interpretation of the results. Firstly the results of the questionnaire are discussed. This means that the effectiveness of the different visualization types is discussed within the different question groups. Thereafter the influence of the participant's spatial thinking skills is discussed. And the participant's preference regarding the visualization types is discussed. Thereafter the usability of the results is discussed considered from a broader perspective and in relation to the literature. Thereafter a short reflection on the research itself is given. And finally recommendations for further research and also for the application of the acquired knowledge are given.

5.1 The questionnaire in general

The test group of 38 participants has shown to be large enough to compare and statistically prove differences between the products. Using a larger group most likely will not have influenced the significant results. The simple (test) questions were answered quite accurate, which indicates that the participant's answers are reliable. The simple questions also confirm the conclusion of the literature review that the choropleth (and the value-by-alpha map) is more accurate and efficient for recognition of enumeration units. This was expected, because the shape or height of the choropleth map is not distorted while the other products are. From this data it can be confirmed that using the cartogram and 3D cartogram technique compromises object recognition which confirmed the findings of the Andresen, Wuschke et al. research.

5.2 The understanding of individual variables

In question group 2 the ability to read the data values of the enumeration units was tested. In more detail, to estimate which of two enumeration units relatively has the highest value and to estimate an absolute value of a specific unit. Within this test the 3D cartogram (extrusion by height) showed to be as accurate and effective as the choropleth for the communication of data values. The cartogram however performed worse, confirming the theory from the Kaspar, Fabrikant et al. (2011) research that (irregular) shapes are not suitable for the communication of (subtle) differences. It is remarkable that the questions of the (worse performing) cartogram were answered the fastest. This could be caused by the misleading irregular shapes i.e. the participants believed to know the right answer. Even if height is as accurate as colour for the communication of data values, still the 3D cartogram has other disadvantages compared to the choropleth. Mainly occlusion or distance between units will be problematic in a large map with many units. However, the 3D cartogram has no need for data classification. Generally the choropleth still seems the best option for univariate mapping.

5.3 The understanding of multivariate relations

Regarding the multivariate questions the cartogram and 3D cartogram showed to be as accurate as the choropleth. It is interesting that the accuracy of these techniques was not significantly different within this question group, especially taken into consideration that for most participants reading a cartogram, not to mention a 3D cartogram, is a new experience. This at least partially could explain the slower answering of the cartogram questions. This is an indication that these products can offer a valuable contribution to multivariate mapping and at least should be taken seriously. Moreover, it is noteworthy that the time used for answering the questions for the 3D cartogram was not significantly different from the choropleth, while in addition an animation was used to give an overview of the map. However, for this specific question it is not so important to be able to read the entire map in full detail, or to compare (subtle) differences between individual units. Because it is focussed on the recognition of patterns between the variables that are present throughout the entire map. It is also not unthinkable that the choropleth is benefited by the fact that most participants are already experienced with this technique.

5.4 Influence of the participants' spatial thinking skills

Because the answers of the participants related to the MGI scored significantly more accurate using the 3D cartogram for multivariate questions than the other participants, it may be that reading 3D maps depends (stronger) on the participant's spatial thinking skills. Considering these results it can be concluded that 3D cartogram maps in particular are suited for experienced users. In those scenarios the fact that the 3D cartogram offers possibilities for the inclusion of additional variables could also be beneficial. For example, a bivariate colour scheme combined with height.

5.5 The visualizations

Despite the fact that the majority of the participants preferred the conventional choropleth technique for multivariate thematic mapping, also a no-negligible part preferred the 3D cartogram. This is an interesting result because the 3D cartogram is less researched and applied then the cartogram, but still more preferred. It is also interesting to notice that a reasonable proportion of the participants preferred to use the 3D cartogram technique, instead of the familiar choropleth. This can be linked to the conclusion from the van Lammeren, Houtkamp et al. (2010) research, which learns that the visualization technique influences the users cognitive responses. Although the influence of the visual appeal of different visualizations is not included in this research, it probably plays a significant part in the participant's preference.

5.6 Usability of the results

The results of this research are essentially usable to get an idea of the usability of the researched mapping techniques. By the time of writing this is the first scientific research that compares the usability of the 3D cartogram with other products. The results offer an overview and perspective of the accuracy and efficiency of these products. This is important because this is a valuable support for the on-going discussions on experiments and examples with these techniques. This research will not give a final and fully complete answer on the question which product is most suitable for multivariate visualizations, nor to present you the perfect mapping technique. However, the results of this study regarding to the usability of the 3D cartogram and cartogram for multivariate mapping are promising. Because the accuracy and efficiency of the 3D cartogram was not significantly different from the choropleth within the multivariate questions, which also applies to the accuracy of the cartogram.

5.7 Reflection on the research

The general objective 'to explore the usability of 3D for multivariate thematic mapping' has largely been achieved. The results provide insight in the usability of the 3D cartogram for multivariate thematic mapping, and in addition for univariate mapping and object recognition. The 3D cartogram is selected because (in literature) it seemed the most suitable 3D technique for multivariate thematic mapping. The results of this research truly contribute to the knowledge of the usability of 3D, or at least the 3D cartogram, for multivariate mapping. In addition, this research also provides a wider view of the differences between 3D and various 2D techniques. The online questionnaire among mainly professionals turned out to be a useful instrument for the gathering of information. The literature review offered useful guidance in the creation of the visualizations and the questionnaire.

5.8 Recommendations

Based on the results further research into the usability of 3D cartograms for multivariate mapping (with three or more variables) is recommended among experienced map readers. This research may reveal if the 3D cartogram in this situation actually is more effective than the common cartogram. Also, it can be recommended to experiment and to work in practice with 3D cartograms for multivariate thematic mapping, so that the map readers can gain more experience with this product, as the results show that experienced map readers performed significantly better reading the 3D cartogram. This influence of map reading experience on the user's performance is also reported in the Ozimec, Natter et al. (2010) research, but may be even stronger for 3D visualizations. It would be interesting to perform a similar research among professionals who are experienced in working with 3D. Based on the conclusions from this research the expectation is that under these specific conditions the usability of the 3D cartogram will be higher. Another interesting question based on the literature is how the 3D cartogram would perform in maps with more enumeration units, considering that occlusion in this case will have a larger influence. Regarding the 2D cartogram based on the Sun and Li (2010) research it is recommended to include the pseudo-cartogram in further research because it is reported to be more accurate in most situations.

6. Bibliography

Andresen, M. A., K. Wuschke, et al. CARTOGRAMS, CRIME, AND LOCATION QUOTIENTS.

Bertin, J. (1983). Semiology of Graphics, The University of Wisconsin Press.

Dillman, D. A. (2000). Mail and Internet surveys: The tailored design method. 2nd ed. New York, John Wiley & Sons, Inc.

Egbert, S. L. and T. A. Slocum (1992). "Exploremap: An Exploration System for Choropleth Maps." Annals of the Association of American Geographers **82**(2): 275-288.

Flury, B. and H. Riedwyl (1981). "Graphical Representation of Multivariate Data by Means of Asymmetrical Faces." Journal of the American Statistical Association **76**(376): 757-765.

Gastner, M. T. and M. E. J. Newman (2004). "Diffusion-based method for producing density-equalizing maps." Proceedings of the National Academy of Sciences of the United States of America **101**(20): 7499-7504.

Harrower, M. (2007). "Unclassed Animated Choropleth Maps." Cartographic Journal, The **44**(4): 313-320.

Heer, J., M. Bostock, et al. (2010). "A tour through the visualization zoo." Commun. ACM **53**(6): 59-67.

Joris, B. (2012). "ArcGIS Actueel Magazine." (012012).

Kaspar, S., S. Fabrikant, et al. (2011). "EMPIRICAL STUDY OF CARTOGRAMS."

Keim, D. A., S. C. North, et al. (2004). "CartoDraw: a fast algorithm for generating contiguous cartograms." IEEE Trans Vis Comput Graph **10**(1): 95-110.

Kessler, F. C. and T. A. Slocum (2011). "Analysis of Thematic Maps Published in Two Geographical Journals in the Twentieth Century." Annals of the Association of American Geographers **101**(2): 292-317.

Kraak, M.-J. (2003). "Geovisualization illustrated." ISPRS Journal of Photogrammetry and Remote Sensing **57**(5–6): 390-399.

Leonowicz, A. M. (2006). "Two-variable choropleth maps as a useful tool for visualization of geographical relationship." Geografija 42/1: 33–37.

Leys, C., C. Ley, et al. (2013). "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median." Journal of Experimental Social Psychology **49**(4): 764-766.

Min, O. and P. Revesz (2000). Algorithms for cartogram animation. Database Engineering and Applications Symposium, 2000 International.

Monmonier, M. (1991). How to lie with maps. Chicago, The University of Chicago Press.

Morita, T. (2011). "Reflections on the Works of Jacques Bertin: From Sign Theory to Cartographic Discourse." Cartographic Journal, The **48**(2): 86-91.

Oulasvirta, A., S. Estlander, et al. (2009). "Embodied interaction with a 3D versus 2D mobile map." **13**(4): 303-320.

Ozimec, A.-M., M. Natter, et al. (2010). "Geographical Information Systems–Based Marketing Decisions: Effects of Alternative Visualizations on Decision Quality." Journal of Marketing **74**(6): 94-110.

Roth, R. E., A. W. Woodruff, et al. (2010). "Value-by-alpha maps: An alternative to the cartogram." The Cartographic Journal **47**.

Roth, R. E., A. W. Woodruff, et al. (2010). "Value-by-alpha maps: An alternative to the cartogram." Cartographic Journal, The **47**(2): 130-140.

Sandvik, B. (2009 - 2013). "thematicmapping.org." 2013.

Slocum, T. A., R. B. McMaster, et al. (2009). Thematic Cartography and Geovisualization.

Stewart, J. and P. J. Kennelly (2010). "Illuminated Choropleth Maps." Annals of the Association of American Geographers **100**(3): 513-534.
Sun, H. and Z. Li (2010). "Effectiveness of Cartogram for the Representation of Spatial Data." Cartographic Journal, The **47**(1): 12-21.
UsabilityNet. (2013). "UsabilityNet." from http://www.usabilitynet.org/tools/surveys.htm.
van Lammeren, R., J. Houtkamp, et al. (2010). "Affective appraisal of 3D land use visualization." Computers, Environment and Urban Systems **34**(6): 465-475.

Appendix 1) Index of CD

In addition to this report a CD is provided that contains additional files and documents. In this Appendix the content of the CD is given.

- Digital copy of thesis report

- Literature
  All digital available sources listed in the bibliography are included on the CD.

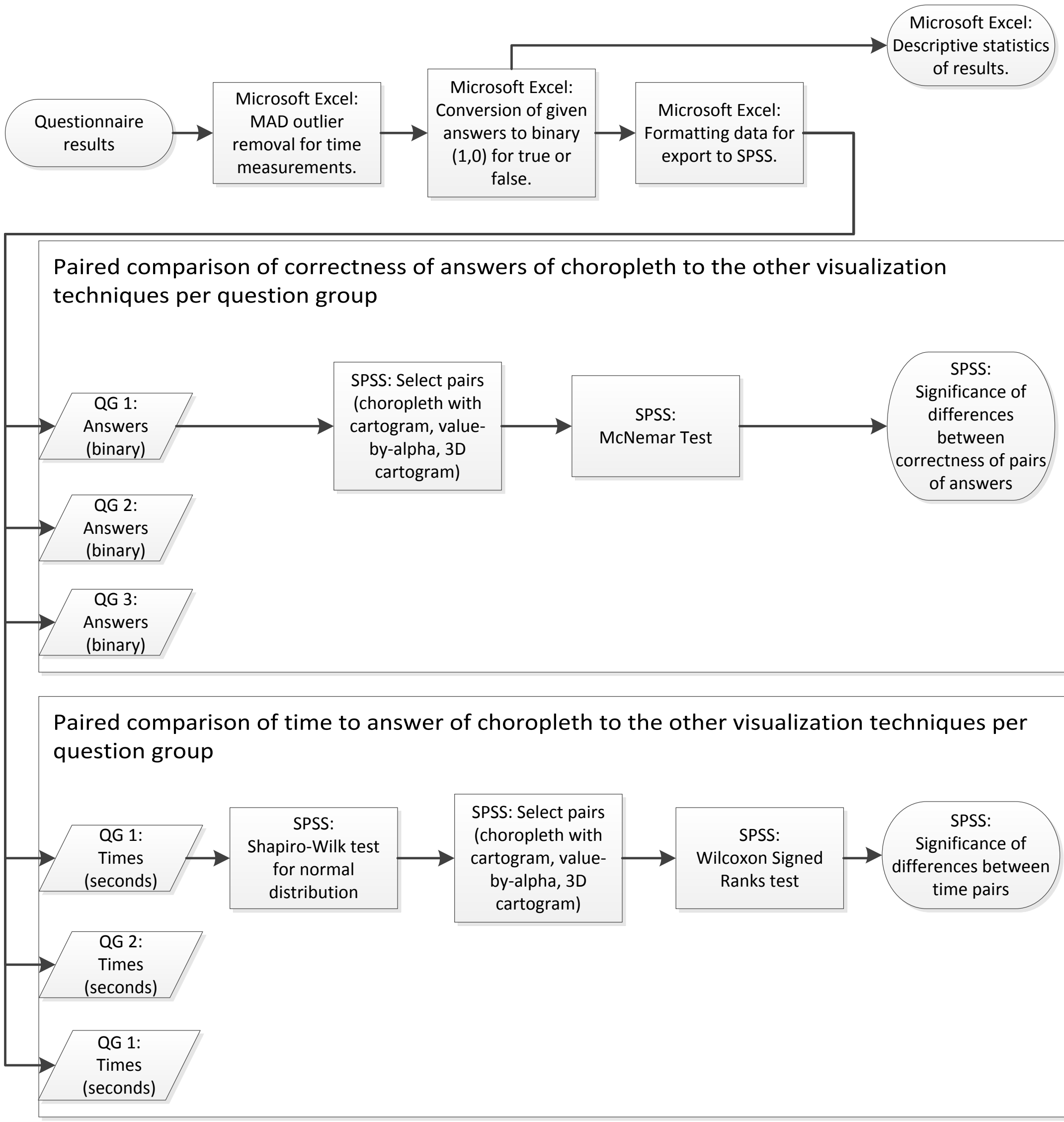
- Similarity matrix
  The similarity matrix that has been used to select the variables for the multivariate visualizations from the CBS dataset is included on the CD. This matrix shows the correlation between the variables based on the Pearson correlation coefficient.

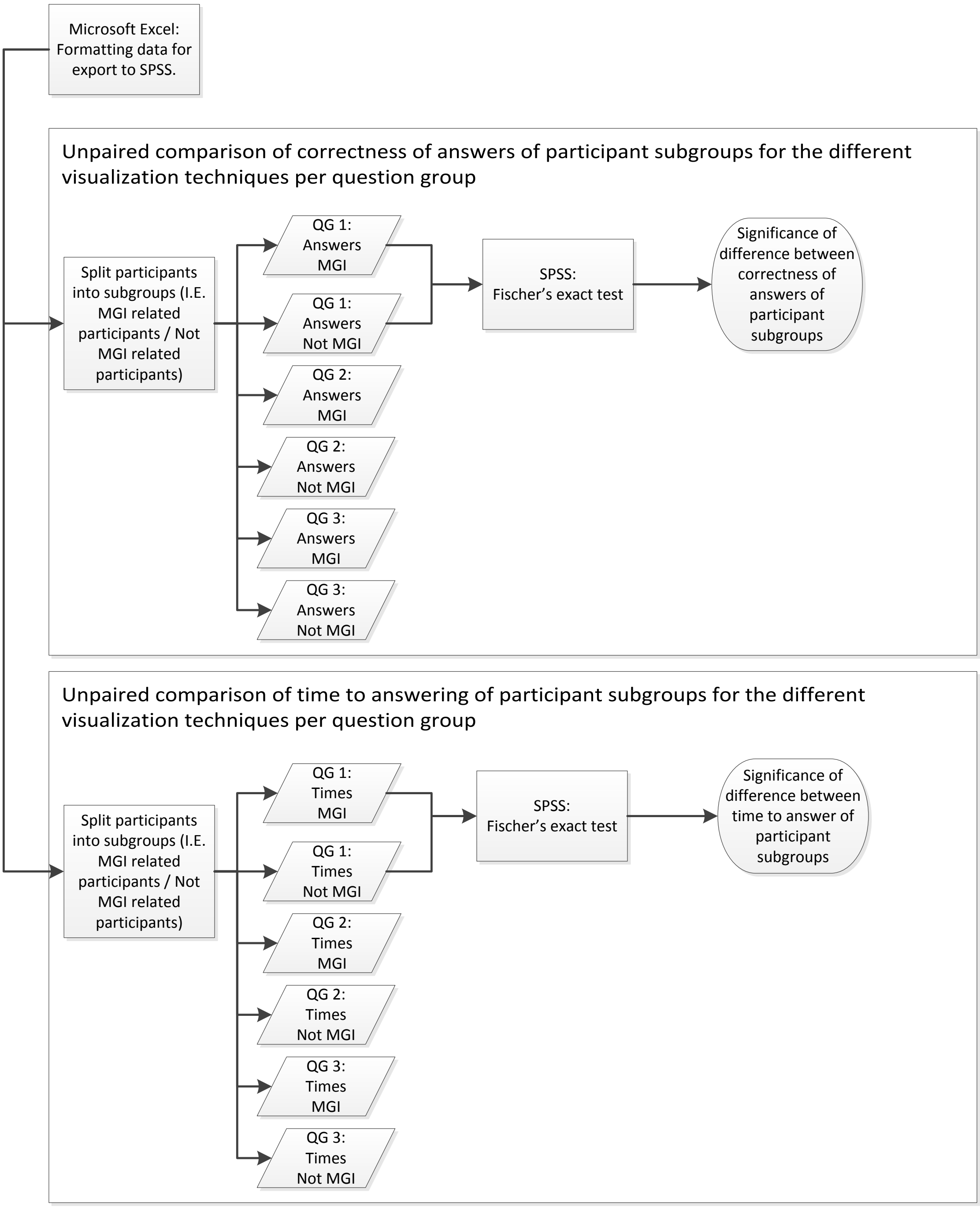
- Questionnaire and visualizations
  A backup of the LimeSurvey website and database that has been used to host the questionnaire online is included on the CD. This can be used to repeat the questionnaire, or for further research and development based on this questionnaire.

- Full questionnaire results
  The 'raw' questionnaire results as well as the Excel and SPSS datasheets that have been used for the statistical analysis of the questionnaire are included on the CD.

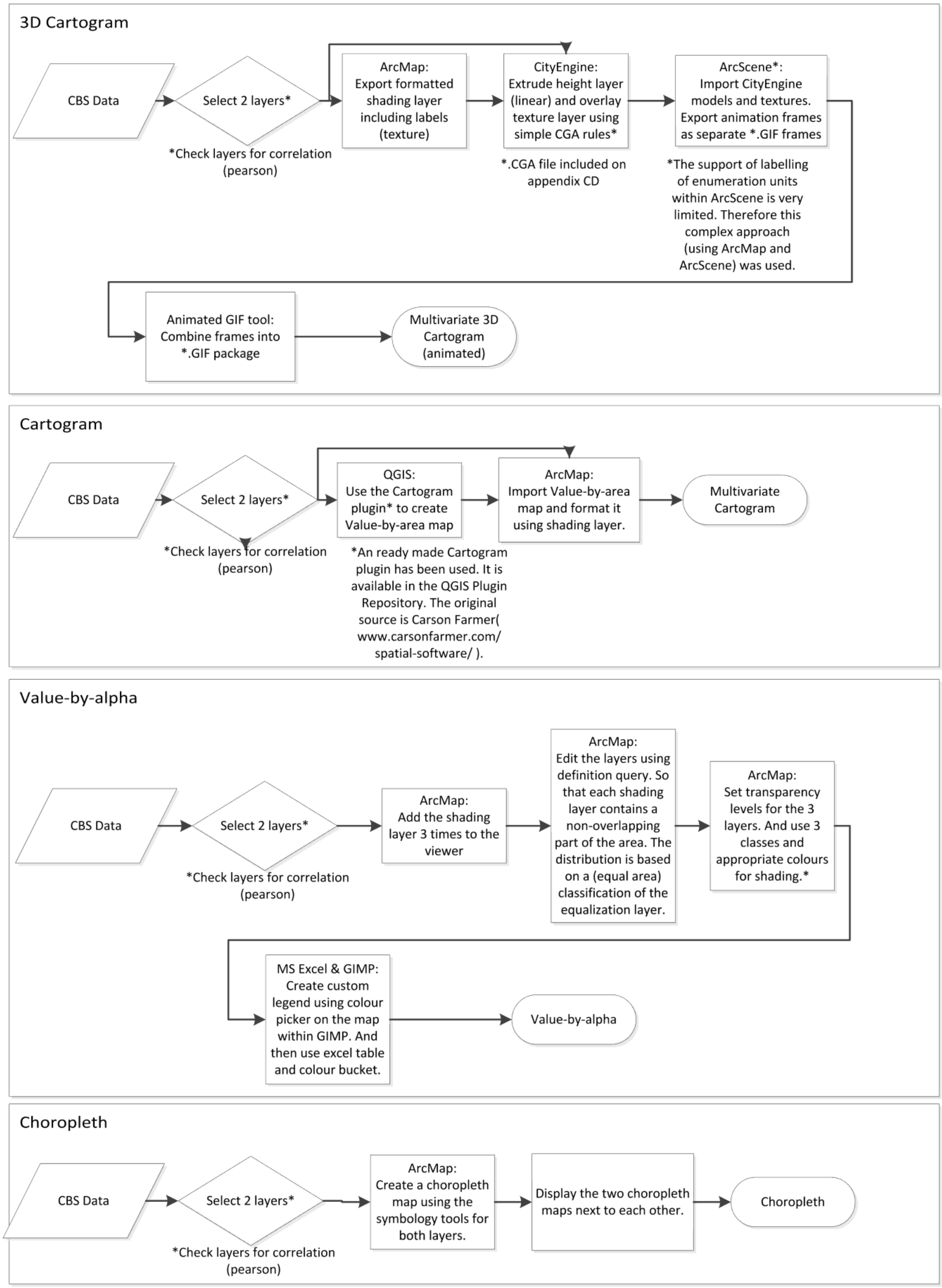
Appendix 2) Flowchart of the statistical analysis of the questionnaire results

This flowchart shows the process that was used to derive descriptive statistics and to perform significance tests from the questionnaire data. To keep the table readable duplicate branches are removed from the graph. For example, the McNemar test for paired comparison of correctness of answers is not only performed for question group 1, but also for 2 and 3 etc. This is also the case for the subgroups. In the research also age, mapping experience, education level and gender subgroups are tested in the same way as MGI. However, only the MGI subgroups showed a significant difference.




Microsoft Excel: Formatting data for export to SPSS.
Unpaired comparison of correctness of answers of participant subgroups for the different visualization techniques per question group
Split participants into subgroups (I.E. MGI related participants / Not MGI related participants)
QG 1: Answers MGI
QG 1: Answers Not MGI
QG 2: Answers MGI
QG 2: Answers Not MGI
QG 3: Answers MGI
QG 3: Answers Not MGI
SPSS: Fischer's exact test
Significance of difference between correctness of answers of participant subgroups
Unpaired comparison of time to answering of participant subgroups for the different visualization techniques per question group
Split participants into subgroups (I.E. MGI related participants / Not MGI related participants)
QG 1: Times MGI
QG 1: Times Not MGI
QG 2: Times MGI
QG 2: Times Not MGI
QG 3: Times MGI
QG 3: Times Not MGI
SPSS: Fischer's exact test
Significance of difference between time to answer of participant subgroups

Appendix 3) Flowchart of the creation of the visualizations



Appendix 4) The questionnaire

Thesis Multivariate 3D visualizations

Welcome to this questionnaire which is part of my MSc research on multivariate thematic mapping and the use of 3D techniques. Completion of the questionnaire will take about ten minutes. This questionnaire is aimed to measure the effectiveness of different visualization techniques for thematic mapping. Try to answer the questions intuitively and within a reasonable time. You cannot return to a question that you have already answered. I want to thank you for helping me with my research.

Gert-Jan Boos

There are 33 questions in this survey

Personals

The first questions are personals

1 What is your gender?

Please choose **only one** of the following:

- ◯Female
- ◯Male

2 What is your age?

Each answer must be between 1 and 120

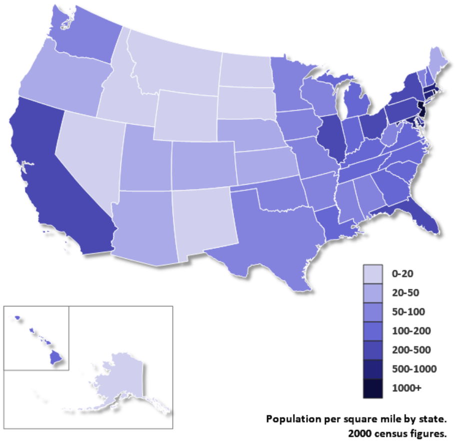
Please write your answer here:

- years

3 How experienced are you with the use of thematic maps?

Please choose **only one** of the following:

- ◯1 - I have never read or created thematic maps
- ◯2
- ◯3 - I read thematic maps occasionally but I never create them
- ◯4
- ◯5 - I read and create thematic maps frequently



The image above shows an example of thematic map that used the choropleth technique. The map in this example shows the population density of a country. Choropleths are maps that show demographic and / or other statistical data. There are different techniques that can be used for the creation of thematic maps, such as the cartogram, value-by-alpha and 3D cartogram.

4 What is the highest level of education you achieved?

Please choose **only one** of the following:

- ◯Primary education
- ◯Secondary education
- ◯Diploma / Professional qualification
- ◯Bachelor or equivalent
- ◯Master or equivalent
- ◯Doctoral or equivalent

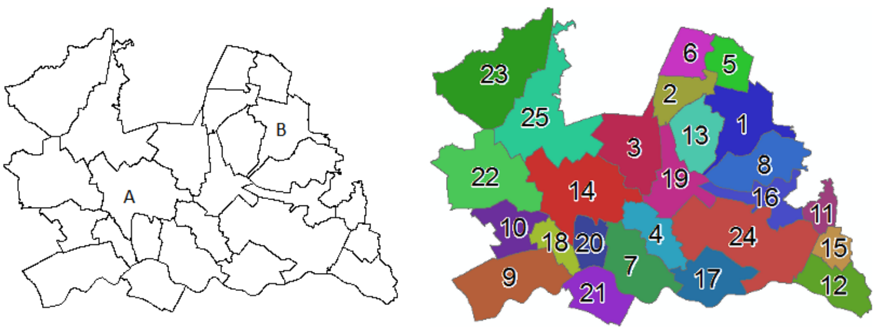
5 Are you a MGI (Master Geo-Information Science) student or professor?

Please choose **only one** of the following:

- ◯Yes
- ◯No

G1-Simple

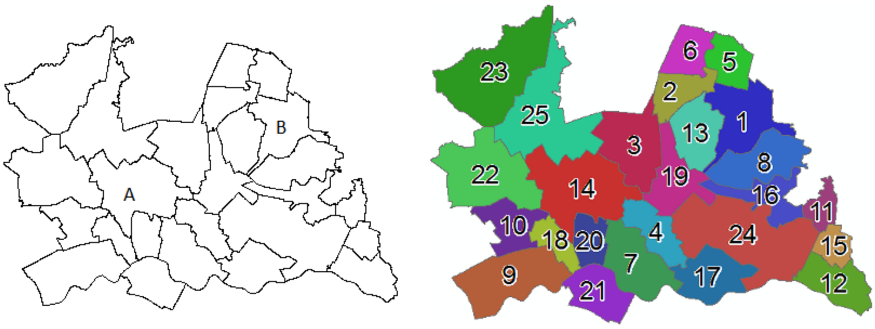
1Ch1 The choropleth map with numbered units shows the municipalities of a province of the Netherlands. The inset map shows two locations (A and B).



Question: Which number on the choropleth corresponds with area A on the inset map?

Please write your answer here: (14)

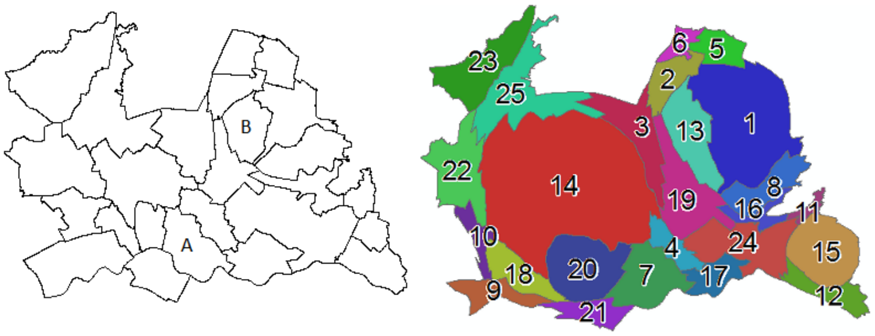
1Ch2 The choropleth map with numbered units shows the municipalities of a province of the Netherlands. The inset map shows two locations (A and B).



Question: Which number on the choropleth corresponds with area B on the inset map?

Please write your answer here: (1)

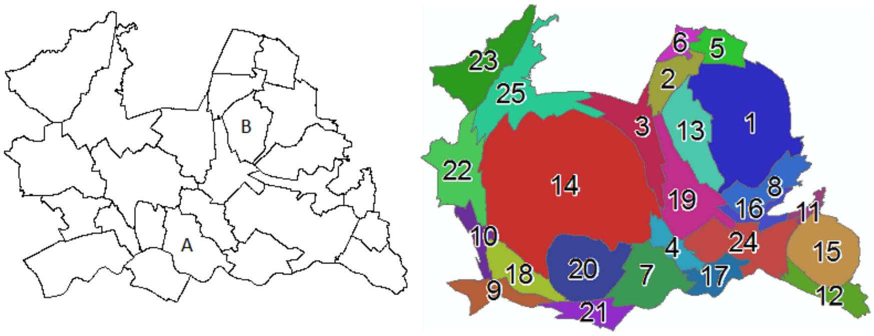
1Ca1 The cartogram shows the municipalities of a province of the Netherlands, where the municipality area is distorted reflecting their population size. The inset map on the left shows the undistorted geometry.



Question: Which number on the Cartogram corresponds with area A on the inset map?

Please write your answer here: (7)

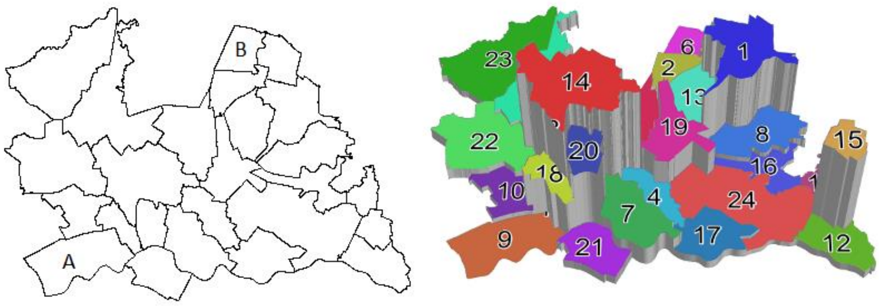
1Ca2 The cartogram shows the municipalities of a province of the Netherlands, where the municipality area is distorted reflecting their population size. The inset map on the left shows the undistorted geometry.



Question: Which number on the Cartogram corresponds with area B on the inset map?

Please write your answer here: (13)

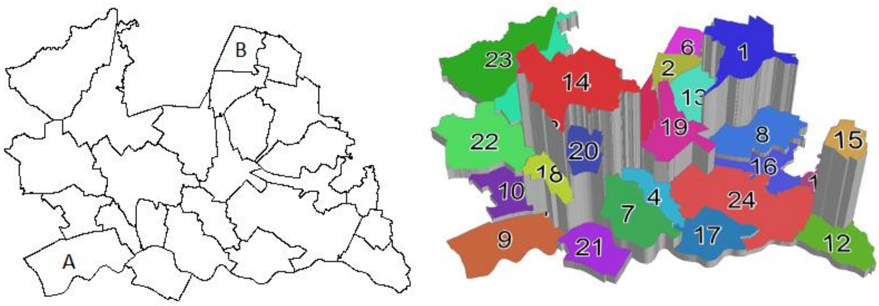
13D1 The 3D cartogram shows the municipalities of a province of the Netherlands. This map uses the height of the units to communicate the values. The inset map shows two locations (A and B).



Question: Which number on the 3D cartogram corresponds with area A on the inset map?

Please write your answer here: (9)

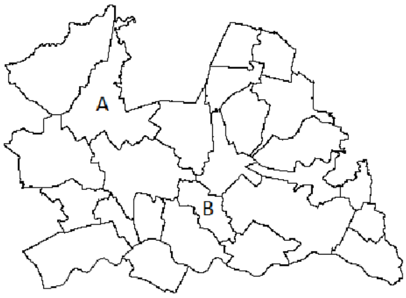
13D2 The 3D cartogram shows the municipalities of a province of the Netherlands. This map uses the height of the units to communicate the values. The inset map shows two locations (A and B).

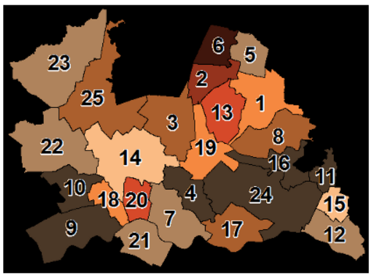


Question: Which number on the 3D cartogram corresponds with area B on the inset map?

Please write your answer here: (6)

1Vba1 The Value-by-alpha map shows the municipalities of a province of the Netherlands. The inset map shows two locations (A and B).

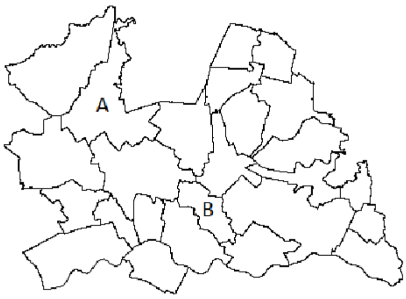


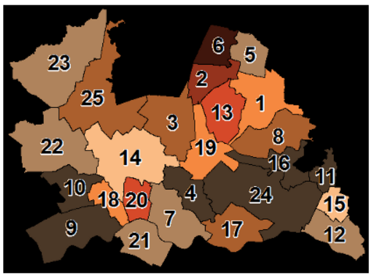


Question: Which number on the Value-by-alpha corresponds with area A on the inset map?

Please write your answer here: (25)

1Vba2 The Value-by-alpha map shows the municipalities of a province of the Netherlands. The inset map shows two locations (A and B).



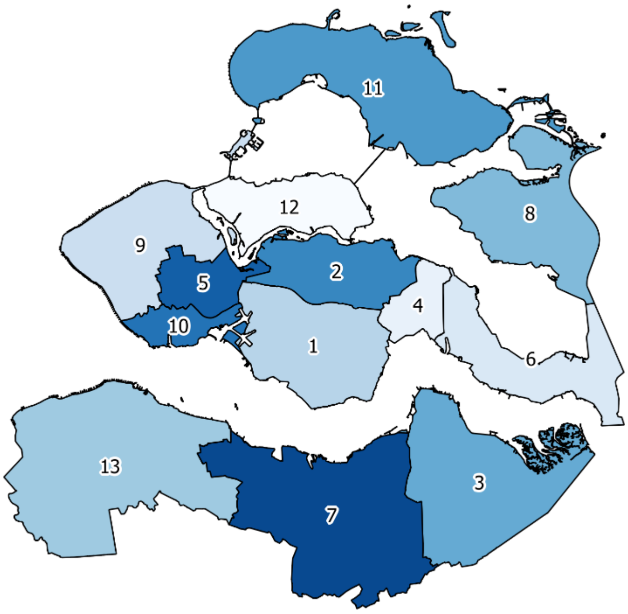


Question: Which number on the Value-by-alpha corresponds with area B on the inset map?

Please write your answer here: (4)

G2-E

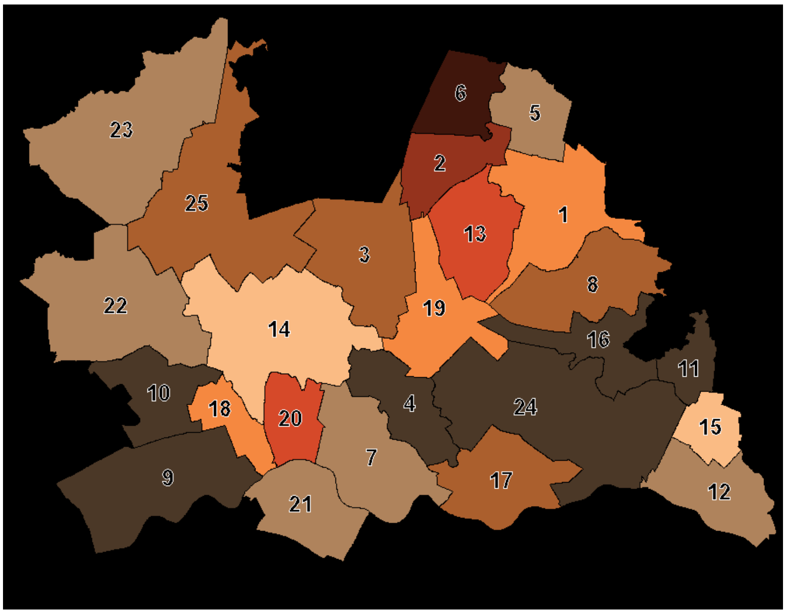
G2E1 The choropleth map shows the population size of the municipalities in a part of the Netherlands. The colour of the areas is used to communicate the values.



Please choose **only one** of the following:

- ○The population of 12 is larger than 2
- ●The population of 2 is larger than 12

G2E2 The Value-by-alpha map communicates by different colour one variable of interest together with a variable for equalization (colour brightness e.g. light red into dark red). The variable of interest in this map is the percentage of the divorced population within a municipality. The equalizing variable is the population density.

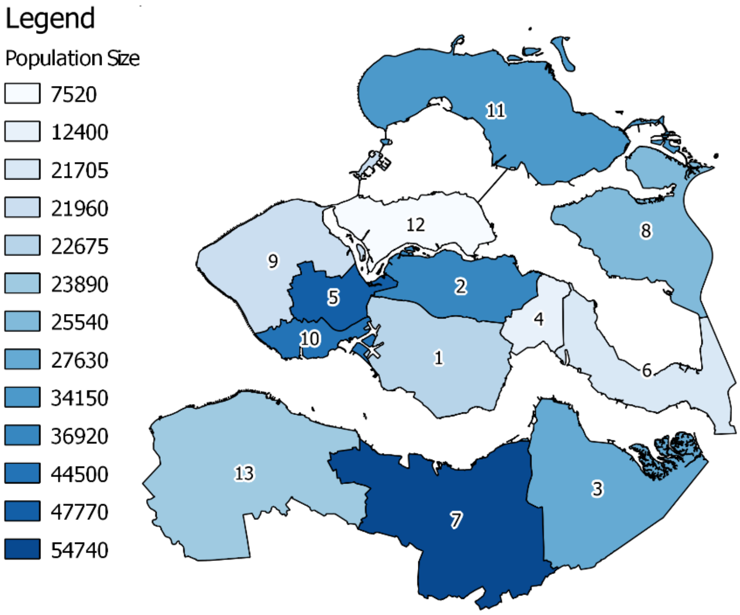


Please choose **only one** of the following:

- ●The percentage of divorced is higher in area 6 than 2
- ○The percentage of divorced is higher in area 2 than 6

G2-Medium

2Ch1 The choropleth map shows the population size of the municipalities in a part of the Netherlands. The colour of the areas is used to communicate the values.

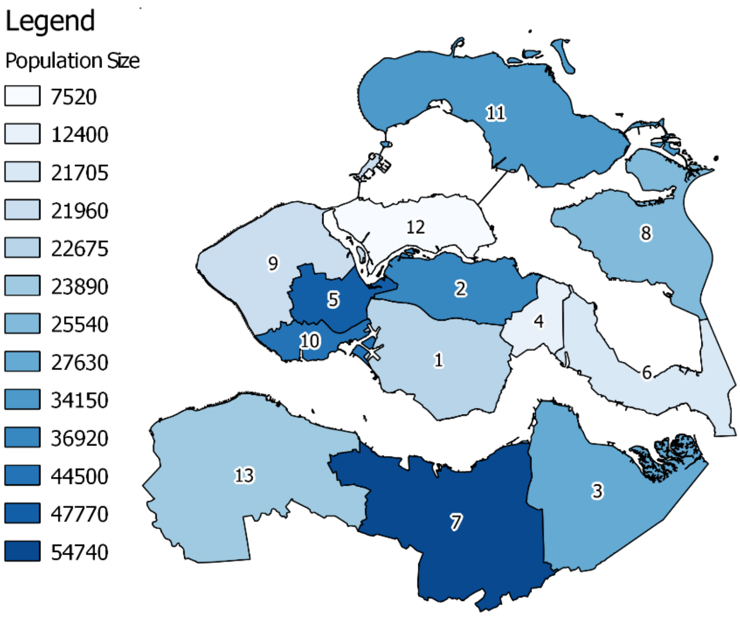


Question: Is the population size larger in area 9 or in area 3?

Please choose **only one** of the following:

- ● The population size is larger in area 3 than 9
- ○ The population size is larger in area 9 than 3

2Ch2 The choropleth map shows the population size of the municipalities in a part of the Netherlands. The colour of the areas is used to communicate the values.

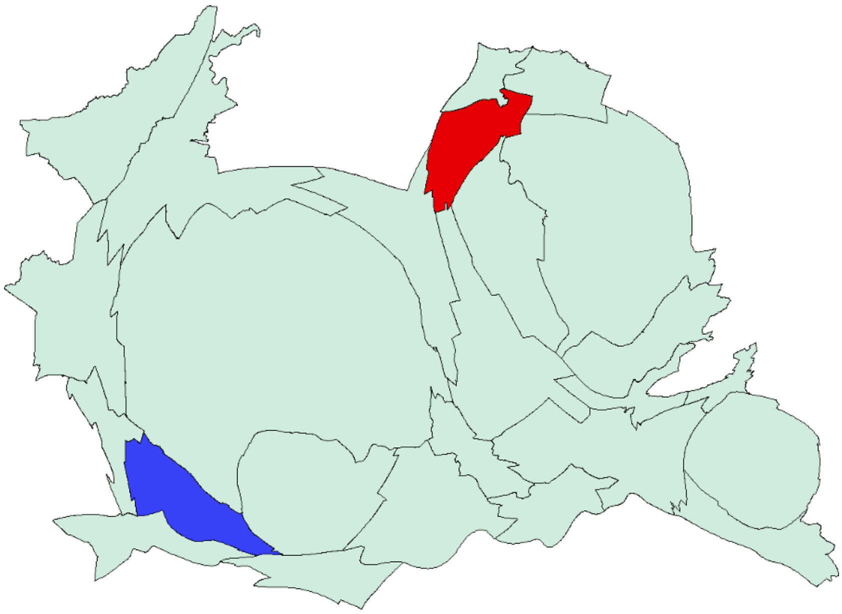


Question: How large is the population size of area 8?

Please choose **only one** of the following:

- ● The population size of area 8 is 25540
- ○ The population size of area 8 is 27630
- ○ The population size of area 8 is 34150

2Ca1 The cartogram shows the municipalities of a province of the Netherlands. The area of the municipalities is used to visualize the population size.

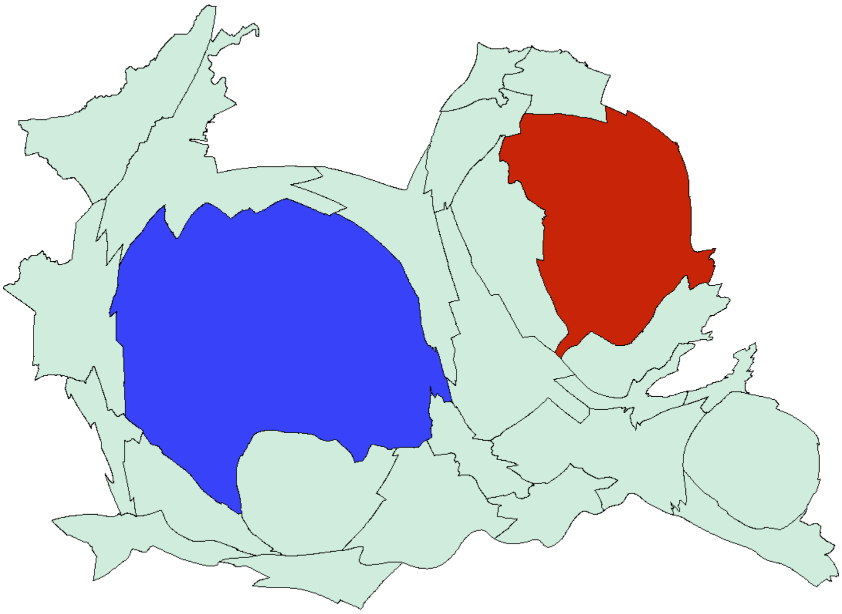


Question: Is the population size larger in the blue area or in the red area?

Please choose **only one** of the following:

- ○ Red is larger than blue
- ● Blue is larger than red

2Ca2 The cartogram shows the municipalities of a province of the Netherlands. The area of the municipalities is used to visualize the population size.

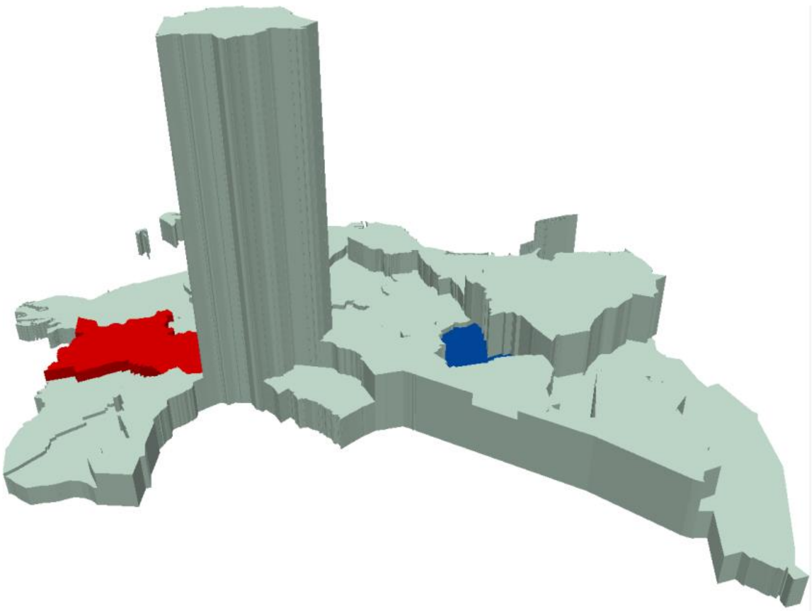


Question: The population size of the red area is 148.250. How large do you estimate the population size of the blue area?

Please choose **only one** of the following:

- ◯The population size of the blue area is 222.375
- ◯The population size of the blue area is 259.437
- ●The population size of the blue area is 316.275

23D1 The 3D cartogram shows the municipalities of a province of the Netherlands. The height of the municipalities is used to visualize the population size.

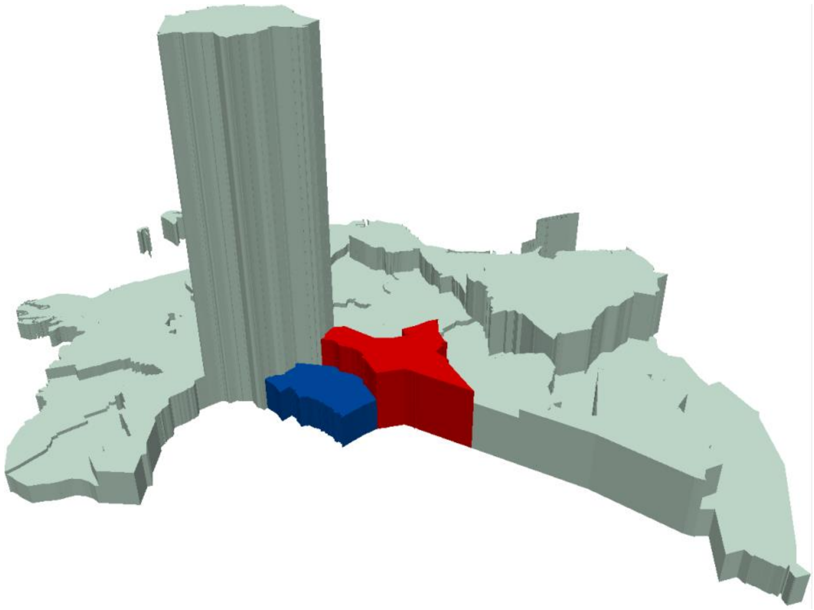


Question: Is the population size lager in the blue area or in the red area?

Please choose **only one** of the following:

- ◯Blue is larger than red
- ⬤Red is larger than blue

23D2 The 3D cartogram shows the municipalities of a province of the Netherlands. The height of the municipalities is used to visualize the population size (linear scale, 2x the population = 2x the height).

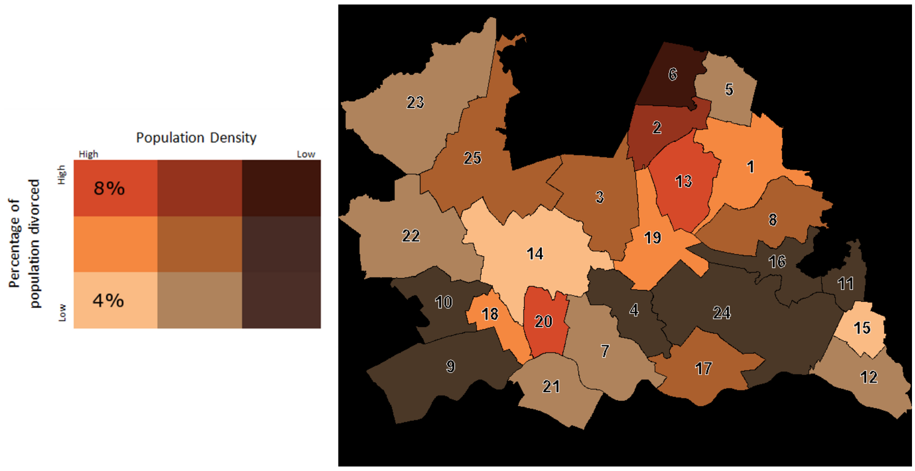


Question: The population size of the blue area is 18.455. How large do you estimate the population size of the red area?

Please choose **only one** of the following:

- ◯The population size of the red area is 26.305
- ⬤The population size of the red area is 34.780
- ◯The population size of the red area is 43.578

2Vba1 The Value-by-alpha map communicates by different colour one variable of interest together with a variable for equalization (colour brightness e.g. light red into dark red). The variable of interest in this map is the percentage of the divorced population within a municipality. The equalizing variable is the population density.

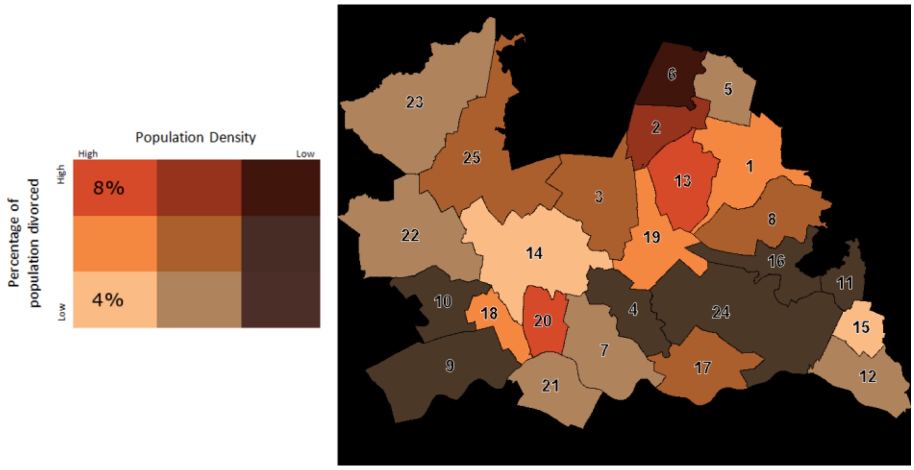


Question: Is the percentage of divorced persons within the population higher in area 2 or in area 3?

Please choose **only one** of the following:

- ● The percentage divorced is higher in area 2 than in 3
- ○ The percentage divorced is higher in area 3 than in 2

2Vba2 The Value-by-alpha map communicates by different colour one variable of interest together with a variable for equalization (colour brightness e.g. light red into dark red). The variable of interest in this map is the percentage of the divorced population within a municipality. The equalizing variable is the population density.



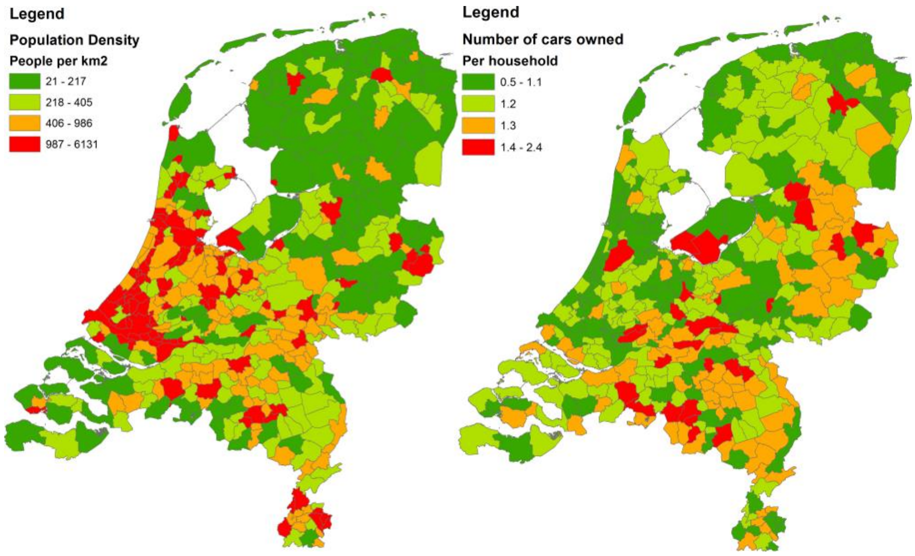
Question: How large is the percentage divorced persons within area 22?

Please choose **only one** of the following:

- ○ 2%
- ● 4%
- ○ 8%

G3-Multivariate

3Ch1 The left choropleth shows the population density of the municipalities. The right map shows number of cars owned per household.

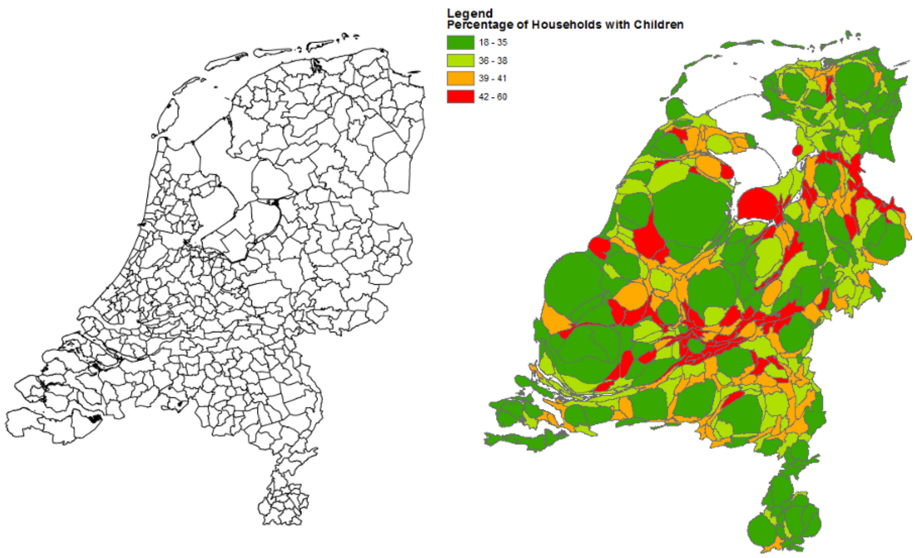


Question: Based on the two maps, what can you say about the relation between the population density and number of cars owned per household?

Please choose **only one** of the following:

- When the population density increases, the amount of cars decreases
- When the population density increases, the amount of cars increases

3Ca1 The area of the multivariate cartogram reflects the number of inhabitants per municipality. The colour of the units reflect the number of households with children.



Question: What can you say about the relation between the number of inhabitants and the percentage of households with children?

Please choose **only one** of the following:

- When the number of inhabitants increases, the percentage of households with children decreases
- When the number of inhabitants increases, the percentage of households with children increases

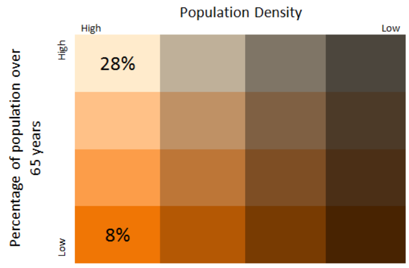
33D1 The height of the areas in the 3D cartogram reflect the population density. The colours of the areas reflect the amount of motorcycles per inhabitant.

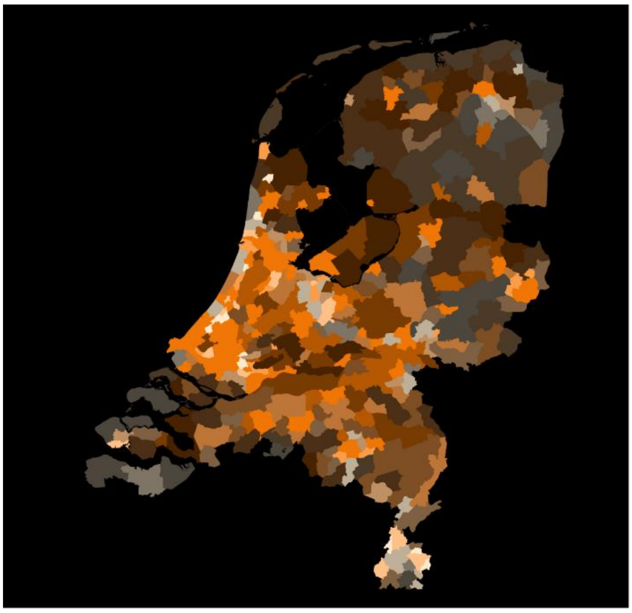
Question: What can you say about the relation between the population density and the ownership of motorcycles?

Please choose **only one** of the following:

- When the population density increases, the ownership of motorcycles decreases
- When the population density increases, the ownership of motorcycles increases

3Vba1 The Value-by-alpha map communicates by different colour one variable of interest together with a variable for equalization (colour brightness eg light red into dark red). The variable of interest in this map is the percentage of the people over 65 years per municipality. The equalizing variable is the population density.





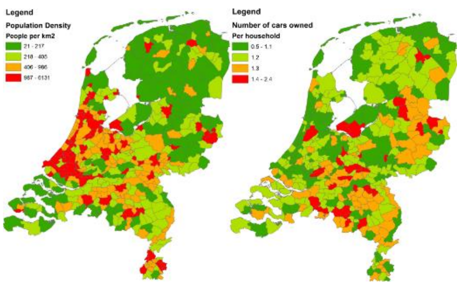
Question: What can you say about the relation between the population density and the percentage of the population above 65 years?

Please choose **only one** of the following:

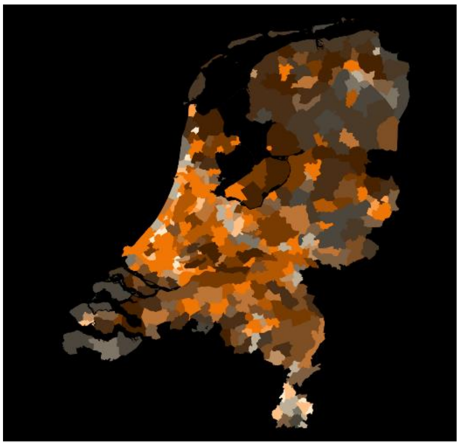
- ⬤ When the population density increases, the percentage of population over 65 years decreases
- ◯ When the population density increases, the percentage of population over 65 years increases

Appraisal

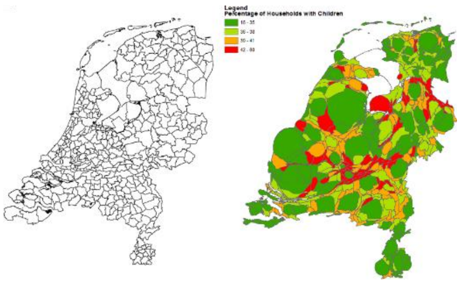
28 Which mapping technique did you prefer for the multivariate questions, where you had to find relations between population density and other variables?



(A)



(B)

(C) 

(D)

Please choose **only one** of the following:

- A - Choropleth
- B - Value-by-alpha
- C - Cartogram
- D - 3D Cartogram

29 You have indicated that you prefer the Choropleth mapping technique. How difficult do you think it is to read Choropleth maps compared to the other maps?

Only answer this question if the following conditions are met:
Answer was 'A - Choropleth' at question '28 [App1]' (Which mapping technique did you prefer for the multivariate questions, where you had to find relations between population density and other variables? (A) (B) (C) (D))

Please choose **only one** of the following:

- Very Simple
- Simple
- Average
- Difficult
- Very Difficult

30 You have indicated that you prefer the Value-by-alpha mapping technique. How difficult do you think it is to read Value-by-alpha maps compared to the other maps?

Only answer this question if the following conditions are met:
Answer was 'B - Value-by-alpha' at question '28 [App1]' (Which mapping technique did you prefer for the multivariate questions, where you had to find relations between population density and other variables? (A) (B) (C) (D))

Please choose **only one** of the following:

- Very Simple
- Simple
- Average
- Difficult
- Very Difficult

31 You have indicated that you prefer the Cartogram mapping technique. How difficult do you think it is to read Cartogram maps compared to the other maps?

Only answer this question if the following conditions are met:
Answer was 'C - Cartogram' at question '28 [App1]' (Which mapping technique did you prefer for the multivariate questions, where you had to find relations between population density and other variables? (A) (B) (C) (D))

Please choose **only one** of the following:

- Very Simple
- Simple
- Average
- Difficult
- Very Difficult

32 You have indicated that you prefer the 3D Cartogram mapping technique. How difficult do you think it is to read 3D Cartogram maps compared to the other maps?

Only answer this question if the following conditions are met:
Answer was 'D - 3D Cartogram' at question '28 [App1]' (Which mapping technique did you prefer for the multivariate questions, where you had to find relations between population density and other variables? (A) (B) (C) (D))

Please choose **only one** of the following:

- Very Simple
- Simple
- Average
- Difficult
- Very Difficult

Thank you

This is the end of the survey. Thank you for filling in this survey.

Gert-Jan Boos

33 If you have any comments or questions on this questionnaire you can fill in the textbox below.

Please write your answer here:

Thank you for filling in this survey. I want to thank you helping me with my research. The results will be presented during my colloquium and in my thesis report.

Gert-Jan Boos

Submit your survey.
Thank you for completing this survey