An Evaluation of Visualization Methods for Population Statistics Based on Choropleth Maps

Lonni Besançon, Matthew Cooper, Anders Ynnerman, and Frédéric Vernier

Abstract—We evaluate several augmentations to the choropleth map to convey additional information, including glyphs, 3D, cartograms, juxtaposed maps, and shading methods. While choropleth maps are a common method used to represent societal data, with multivariate data they can impede as much as improve understanding. In particular, large, low population density regions often dominate the map and can mislead the viewer as to the message conveyed. Our results highlight the potential of 3D choropleth maps as well as the low accuracy of choropleth map tasks with multivariate data. We also introduce and evaluate popcharts, four techniques designed to show the density of population at a very fine scale on top of choropleth maps. All the data, results, and scripts are available from osf.io/8rxwg/

Index Terms—Choropleth maps, bivariate maps

1 INTRODUCTION

A perennial issue with mapping statistical information onto choropleth maps is that they tend to overemphasize large, yet often sparsely populated, administrative areas because of their strong visual weight [1,44,64,88,94]. Since choropleth maps usually convey only information about the statistic of interest and no information on the spatial distribution of population, underpopulated areas tend to dominate the map space while the more densely populated are small and hard to distinguish. We hypothesize that large and sparsely populated areas could bias the average information that should be taken from a map visualization and maps could thus be inaccurately interpreted.

There are several solutions to the issue of relative areas [41]. One is juxtaposition: the two variables displayed in adjacent charts. The second is superposition: the two variables encoded simultaneously in a single chart. The last is explicit encoding: directly displaying a single variable computed from the two component variables. Juxtaposition is a viable solution: it clearly allows the user to perceive patterns in each of the two represented variables. However, recent work [22] suggests that integrating the information from two charts is more error prone. Superposition allows the user to understand the influence of both variables. However, bivariate visual representations have been said to be harder to understand [102]. It is thus still unclear which strategies should be adopted to create bivariate maps to properly convey population information in addition to the measured variable usually shown on univariate maps. We aim in this work to address this issue.

The contributions of this work are threefold. We first describe and implement several alternative bivariate map designs displaying a statistic and population distribution. We then report the results of a qualitative evaluation of these designs with 10 visualization researchers and the results of a controlled experiment with 60 participants. Finally, we introduce popcharts, designed to display statistical data over a fine-grained population distribution, and report a qualitative evaluation of this approach by the same 10 visualization researchers together with a controlled experiment with 47 participants.

2 RELATED WORK

This review considers the design and use of choropleth maps as well as previously conducted studies.

2.1 To quantify or not to quantify

The literature has primarily focused on how to design choropleth maps and the problems that each design might exhibit. We briefly sum up past work on discrete and continuous colour coding on maps.

Choropleth maps can be classified into unquantized (i.e., continuous or unclassed, e.g., [107]) and discrete (e.g., [34, 53, 71]) for colour coding. Unclassed maps, initially proposed by Tobler, create unique shades for each class as well as offer a greater fidelity to the data [17,50,99]. However, they do not account for perceptual limitations of colours [15,80,84,85,104] and recent work has shown that discrete map could...
lead to better performance [75]. We thus decided in this work to discretize the data of the statistic that we wish to show.

2.2 Bivariate maps

A common means to display two variables on a map is to use a bivariate colour coding scheme [30, 51, 82, 83, 100], but these have been shown to be hard to interpret [37, 102] and their legend hard to memorize [22]. Consequently, Bivariate colour maps are often limited to a small set of colours [22, 53, 60, 71, 83, 100]. Bivariate colour coding is an integral conjunction: the selective separation of attention between both attributes encoding is not straightforward. It is therefore not recommended for variables with different units like population on the one hand and a statistic on the other hand to both use colour [87].

Other approaches have focused on augmenting choropleth maps with additional visual elements without relying on an extra colour, starting with Charles Joseph Minard [35] and his attempt to display quantities of moves supplied to Paris by each French department in 1858 [67]. Other maps followed in the 19th century [7, 8]. More recent work has also used this approach with data-dependent glyph augmentations [105] or pixelization/glyph rotations to show uncertainty on a map [60]. Height on a 3D map has also been suggested [49] and tested to observe the evolution of populations [48]. Other approaches rely on a visual property that is closely related to colour, such as opacity [22, 85, 88]. A full taxonomy of possible approaches is described by Elmer [32] who showed that, despite common reluctance to use them, bivariate maps can be successfully interpreted. Elmer tested eight bivariate maps but did not test all possible variants in their taxonomy. We take inspiration from these approaches and incorporate 3D, opacity and glyphs in the set of techniques to display bivariate data. Other studies comparing the representation of geographical data are detailed in Sect. 3.

The study from Correll et al. [22] is particularly interesting for our work. They focused on uncertainty visualization with bivariate maps and compared juxtaposed univariate maps, continuous bivariate maps, discrete bivariate maps and a new version called Value-Suppressing-Uncertainty which aims to reduce the number of categories in bivariate maps. Their studies indicate that continuous colour-based bivariate maps and juxtaposed univariate maps are outperformed by discrete bivariate maps relying on a colour and transparency scheme. We take inspiration from this study but, instead of uncertainty visualization, we focus on population density and merge their results with other work [37, 102] and recommendations [32, 86] to eliminate bivariate colour-coding schemes, preferring colour and transparency [22, 85, 88].

2.3 The Benefits of Dasymetric Maps

To avoid the common pitfall of choropleth maps, cartographers have focused on dasymetric maps in which the boundaries of cartographic representation are not arbitrary/administrative but, instead, reflect the spatial distribution of the variable being mapped which may be presented at a fine scale [31]. Many projects have focused on how to derive a fine-scale distribution of the population that is more realistic than in a choroplath map [27, 55, 59, 70, 108, 113]. Researchers have sought out to confirm the benefits of such precisely mapped population densities in different domains [5, 19, 79, 91]. In the case of health outcomes, recent work has shown that choropleth maps tend to lead to either overestimation or underestimation of risk exposure and that finer-population distributions provided by dasymetric maps help [81]. In all mentioned previous approaches, however, fine-scale population distributions are often shown on classical-delimited maps, probably for two reasons. First, the classic divide into counties or other regions is usually well-known to the map readers and can help them discuss results. Second, social or health data is rarely available with such a high-precision [23, 91]. Consequently, it is important to be able to display both very fine-scale population information as well as additional information at a larger scale. This is addressed in this paper through our popchart techniques from Sect. 6 onward.

3 Visualizing Population on Choropleth Maps

We have highlighted, based on previous work, that bivariate colour coding is not ideal for simultaneous encoding of statistical and population data. Past work has argued for the use of two separable visual variables to encode bivariate maps [32, 87]. The range of tasks or possible encoding tested has, however, been limited. We propose to test several representations combining two separable visual variables as well as visual representations combining two integrated visual variables. Based on a survey of existing techniques, we present here possible representations of bivariate choropleth maps in which the population and data of interest are simultaneously represented. For each technique, we describe previously conducted evaluations and their outcomes.

We use maps of French départements to illustrate different designs of bivariate maps. French départements are well defined administrative subdivisions of France for which population and other statistics have been recorded for a long time. We chose to use unemployment rates as statistic because it is unambiguously related to population (as a percentage) and is known to be related to population density and geography. For experimental purposes most départements have comparable sizes and so perception of colour is more homogeneous across tasks and trials. Départements are also interesting because actual inhabitants have little knowledge of their relative positions or data, and it thus decreases the possible a-priori knowledge for controlled experiments.

3.1 Juxtaposed Univariate Maps

Juxtaposed Univariate Maps consist of juxtaposing (i.e., putting side by side) two choropleth maps, each showing a different variable (see Fig. 2). This technique is well established and has been used, for example, to compare crime and education [4]. While Juxtaposed Univariate Maps does not fit our constraint of mapping both variables on the same figure, they are often used as a baseline for studies on bivariate choropleth maps [22]. Furthermore, mapping population directly to a second choropleth breaks Monmonier’s rules [69] by showing magnitude in a choropleth map instead of intensity. For the sake of the experiment we chose to quantize the base choropleth and not the additional choropleth for population. Each map went with its corresponding legend and hovering over a region in one map also highlighted the same region in the other (see Fig. 2).

3.2 Absolute Maps

Absolute Maps combine the information conveyed by the two variables into a single value (Fig. 3 right). The result is likely to be very dramatic since the distribution of population is often exponential. The variation of data where, for instance, 90% of space is occupied by 10% of population is then hidden in the lowest quantile of the absolute scale. Like the population map of Juxtaposed Univariate Maps it breaks Monmonier’s rules [69] by showing magnitude instead of intensity. It does, however, make sense to consider it as a baseline to the problem we are exposing in this paper.

3.3 Colour and transparency maps

They have been introduced as value-by-alpha maps in the literature [88] as an alternative to cartograms to display enumeration data when looking at a specific measure linked to it, for example population information. Similar to Absolute Maps they can lead to a dramatic effect as 90% of space, occupied by 10% of population is hidden in the most translucent quantile (see Fig. 3 left). According to Roth’s work [87], value-by-alpha maps represent asymmetrical conjunctions and are useful when one variable is more important than the other. Their use could be interesting to map population to transparency values so that low-density areas are less discernable when compared to highly populated areas.
We thus conjecture that this technique has the potential to effectively work on maps has been featured in various articles showcasing the case of maps onto which additional information is encoded by rectangles with high-values mapped to their height will occlude regions behind them. We can however speculate, based on past research on the role of interaction for occlusion management [61, 109], that simple interaction techniques would avoid this issue. Interaction, on the other hand, can make it more difficult to see overall patterns since information from multiple views needs to be integrated mentally. We believe however that such interaction could be restricted to small rotations or pitch changes to avoid this problem while still preventing occlusion issues.

An early study [24] highlighted that 3D maps seemed to do as well as scaled-circle maps and that when faced with a 3D maps, readers are not likely to try to interpret volume but will focus on height even without a legend to guide them. This makes 3D Choroplethsa particularly interesting technique for our study. Niedomysl et al. [72] showed that printed 3D maps are less effective than printed 2D ones in recall tasks but possibly better when estimating the percentage of population living in a specific area (although the effect size was small). Stewart and Kennelly [95] showed that 3D prism maps using shadows could help in discriminating between population levels in different regions. We refrained from implementing shadows to limit the number of potential confounding factors in our experiments. In addition to their previously-highlighted advantages, 3D Choropleths maps rely on separable visual variables which are better for bivariate maps [32, 87]. We thus conjecture that this technique has the potential to effectively represent both population and the data of interest within the same map.

3.5 Choropleth Dot Maps
In 1865, Minard proposed to display the population in a map by adding visual elements (e.g., squares proportional to the number of inhabitants) onto each region displayed in a coloured map [68]. However, Minard never used this technique and replaced circles or squares by pie charts to depict multiple values per region. Choropleth Dot Maps also propose the display of two separable variables and is therefore better suited for a bivariate map [32, 87] and an interesting technique to evaluate.

3.6 Choropleth Bertillon Maps
Jacques Bertillon is mostly famous for his statistical classes but he also promoted graphical representations of statistical data [77]. His work on maps has been featured in various articles showcasing the history of statistical data visualization [36, 77]. Bertillon’s work shows maps onto which additional information is encoded by rectangles where the height and base length are dependent on two variables. In his map of the attractiveness of Paris compared with other French regions [7], Bertillon used the base of a rectangle to be proportional to the population, $P$, and the height to represent the attractiveness of Paris computed as the number of persons born in Paris divided by the number of inhabitants of the region $(I/P)$ which is an index dependent on the population. Consequently, the area of the rectangle is given by $P + I/P$ the number of persons born in Paris in each region. He employed a similar technique in another map: ‘Les étrangers à Paris’ [8].

While area is not ideal for quantitative comparison of data [9, 20], it remains that two of the three variables presented in the map can be encoded by length which is ideal for quantitative data [9, 20]. This idea was pushed forward with, for instance, the visualization of the evolution of space taken by pasture in Normandy onto a map (see Fig M in La Statistique Graphique [58]) or by Bertillon himself in the visualization of the Opérations du Mont de Piété [6]. We implemented this idea on top of a choropleth map leading to Choropleth Bertillon Maps (see Fig. 1(E)). Such representations relying on width and height are integral conjunctions [32] and are therefore unlikely to perform well for bivariate maps with different scales [32, 87]. However Choropleth Bertillon Maps also presents a redundant and separable encoding with hue and we postulate that it could yield interesting performances.

3.7 Deformed Cartograms
It is possible to attach the data to the regions and scale them so that their areas become proportional to their data [46]. We therefore obtain a map that uses value-by-colour for one variable and value-by-area for the population distribution (Fig. 1(D), Fig. 4 left). This idea has been exploited in statistical visualization in the 19th century [76]. An early example represented scaled-down versions of France through the years to represent the time to reach several cities with advancing transportation technology. Recent examples have used them to encode the density of population on global earthquake intensity maps [45] or residential data [43]. Cartograms, however, have inherent limitations. Their deformations impact the shape of the regions or the topology of the map, possibly hindering the readability of the map itself or the represented data [44, 88] and being able to use and interpret cartograms might require a specific learning curve [98]. However, some cartogram techniques have focused on preservation of shape or topology [78].

We are interested in contiguous cartograms [73] as they preserve neighboring regions information. While the limitations and advantages of cartograms have been thoroughly discussed [42, 78, 97] and while variations of cartograms are often compared (e.g., [73, 96, 106]), only a handful of studies compared them to other cartographic representations. Kaspar et al. [52] compared cartograms with choropleth maps with graduated circles. Their results seem to indicate that choropleth maps with graduated circles are easier to interpret but the tasks participants had to perform were not detailed in the paper. Sun and Li [96] also compared cartograms with other common representations but only asked users for subjective preferences.

3.8 Non Contiguous Cartogram Maps
To avoid lack of recognizability posed by standard cartograms, we tested Non Contiguous Cartogram Maps (Fig. 4 right). We shrink the size of each individual coloured region of the choropleth to represent the number of inhabitants according to the formula \(\sqrt{\text{density}}\). This method is population/area of the considered region and maxdensity the pre-computed maximum of this variable for all regions. All regions are then zoomed in all together and only the few biggest ones are translated to avoid overlapping. The regions with the highest number of inhabitants are represented bigger.
than their correct geographical size while all other regions will necessarily shrink. This trade-off between very few regions scaled up and translated and the vast majority of regions scaled down and correctly centered ensures good readability of topology and choropleth colours. The blank spaces between the regions allow the correct frontiers to be displayed between them and can help reading individual colours without influence from the neighbours. Yet, it may also degrade the view as ability to compare colours between two adjacent regions. This problem is softened by the choice of quantized choropleth maps which produce very distinct colours. We expect Non Contiguous Cartogram Maps to be easier to interpret because, unlike Deformed Cartograms, they preserves the original shape of the regions and the topology.

4 INITIAL EXPERIMENTATION AND EXPERT FEEDBACK

4.1 Experiment description

To select a subset of techniques for a controlled experiment, we ran an initial online experiment to collect feedback from visualization researchers (from several institutions) about the techniques described in Sect. 3. We showed screenshots of each visualization, displaying unemployment and population distribution in French départements. We asked them to give benefits and limitations of each representation and to rank, on a Likert-scale from 1 (Very Easy) to 5 (Very Hard), how easy it was for them to achieve specific tasks with each visual representation.

4.2 Tasks

To explore the benefits and limitations of each representation, we defined a series of specific tasks based on a review of cartography tasks. Roth [86] classified high-level user goals, objectives, operand primitives and operators for tasks in cartography. Of particular importance for our study is the categorization of operands (Space-Alone, Attributes-in-Space, or Space-in-Time) and objectives (Identify, Compare, Rank, Associate and Delineate). While comprehensive, this classification is aimed at monovariate maps and so most of the tasks and objectives described are less appropriate in the testing of bivariate maps. Furthermore, the taxonomy does not directly describe a specific objective that is often considered important in the case of cartograms: summarize. This summarization task refers to the objective when users are tempted, or even asked, to aggregate data in some parts of the map to try to form and memorize a “big picture” message. It is one of the elements of a cartograms’ taxonomy of goals [74], frequently used to compare different types of cartograms (e.g., [73]). This summarize objective is often presented in information visualization taxonomies by the names summarize (e.g., [16,112]), overview (e.g., [57,92]), or even review [93]. While not initially used to define tasks in bivariate maps, Roth’s taxonomy [86] can, nonetheless, be used to describe tasks with normalized and widely-accepted vocabulary. In Table 1 we describe the set of 5 questions that we have derived using this taxonomy.

4.3 Results

Ten visualization researchers answered this online experiment. Results of the Likert-scale ratings are presented in Fig. 5. We counted the number of times when the number of (Very) Difficult is higher or equal to the number of (Very) Easy for each technique and each question. If it happened more than twice for a technique, we looked at the qualitative feedback to determine whether we should exclude them from the user study. Non Contiguous Cartogram Maps (Q1_diff, Q3_cont, Q5_cont), Deformed Cartograms (Q2_cont, Q4_cont, Q5_cont), Transparent maps (Q1_cont, Q2_cont, Q3_cont, Q5_cont), and Juxtaposed Univariate Maps (Q3_cont, Q4_cont, Q5_cont) were candidates to be excluded.

The feedback for Transparent maps confirmed that participants had issues understanding the mapping (×8 participants), that it was difficult to know whether the colour or the transparency was changing (×4), and that colours were hard to distinguish at lower levels of transparency (for low-populated areas). We thus removed it from the pool of techniques for the controlled experiment. Similarly, while Non Contiguous Cartogram Maps were generally praised for the concept, 6 participants reported that regions were usually too small to see properly/compare. We therefore also excluded it from the pool of techniques.

Juxtaposed Univariate Maps were not described very negatively in the feedback, apart from the observation that the computation of the absolute number (results of the two juxtaposed maps) could be difficult. Since it is often used and the baseline of some other studies (e.g., [22]), we decided to retain to allow comparison with other techniques. Similarly, while Deformed Cartograms were not generally well perceived (6 participants reported that they were hard to interpret, and 3 said that it was difficult to compare regions), they are popular and heavily studied (e.g., [1,64]), so we kept them in our experiment. Based on qualitative feedback only, we removed two other representations. First, Absolute Maps were, as expected, regarded as missing essential information to complete the tasks, so we removed them. Then, Choropleth Dot Maps are very similar to the idea of the Choropleth Bertillon Maps and we therefore decided to remove them from the pool of techniques. Finally, 7 participants complained about the label placement, mirroring past findings (e.g., [63,89]). In our controlled study, we thus remove labels and only display them when hovering.

We thus established a list of 4 techniques to use in a controlled experiment: Juxtaposed Univariate Maps, Choropleth Bertillon Maps, 3D Choropleths, and Deformed Cartograms.

5 CONTROLLED EXPERIMENT 1

We conducted a controlled experiment to evaluate which of the retained techniques would be helpful to show the distribution of population and the statistic of interest. The experiment is available at tiny.cc/mapstudy and its pre-registration (frozen data-analysis scripts) at osf.io/rgkpj.

5.1 Participants, Questions, and Techniques

Invitations to participate, with a link to the study, were sent by email to students and researchers working both within and outside of the fields of visualization or computer science. The email also asked them to forward the experiment to others. As such, predicting the number of respondents was difficult and we, instead, decided to stop the data collection on a specific date (after 10 days). While not common in HCI, using a time-based stopping rule to preregister a sample size is not rare and is actually in the template of preregistrations on AsPredicted.org. The techniques we used in the experiment were the ones that we had previously identified (namely, Juxtaposed Univariate Maps, Choropleth
Bertillon Maps, 3D Choropleths, and Deformed Cartograms), and we again used the data of France’s unemployment and population. The order of techniques was counterbalanced to avoid learning effects.

The set of questions was taken from our previous experiment. To avoid learning effects, 4 different but equivalent sets of questions were prepared. Each set contained 5 questions as previously described (Q1\textsuperscript{pop} to Q5\textsuperscript{sum}) and were assigned, in turn, to different representation; the same set was not always asked when using the same visualization. For each question, areas of interest were highlighted so that the completion time did not contain the search time but only the time it took for participants to answer a question with a given representation.

We did not give users the possibility to interact (rotations, dragging, zooming, or changing the pitch). While this does not reflect the real usage of such maps, it is necessary to avoid possible confounding factors in our experiments, and is not uncommon in visualization experiments [14,26,103]. While this could negatively affect 3D Choropleths, it is essential to have a fair comparison between techniques.

### 5.2 Planned Analysis Results

A total of 60 participants completed our experiment. However, 2 entered nonsensical demographics and were removed according to our preregistered exclusion criteria. We thus had valid data from 58 participants (19 females, mean age = 26.3, median = 26, SD = 4.8, range 18-41). While such data is usually analyzed with NHST and ANOVAs, recent criticism of NHST to analyze experimental data [3, 25, 29, 38, 40, 65] and recent APA recommendations [101] led us to report our results using estimation techniques with effect sizes [21]: reporting them is not always recommended [2].

5.2.1 Accuracy

Accuracy results, whether a participant had the right answer (score of 1) or a wrong answer (score of 0), are presented in Fig. 6, while Fig. 7 presents pair-wise differences (i.e., individual differences). The small overlap of confidence intervals (CIs) for Q1\textsuperscript{pop}, suggests that Juxtaposed Univariate Maps is likely to perform better than the other techniques. This is confirmed by the non-overlap with 0 in Fig. 7. Similarly, for Q2\textsuperscript{Ump}, the small overlap of CIs suggests that Juxtaposed Univariate Maps and Deformed Cartograms lead to a better accuracy (confirmed in Fig. 7). Concerning Q3\textsuperscript{Comb}, we could not find evidence of a difference between the techniques which all seemed to perform poorly (around 50% accuracy). The small overlap of CIs in Fig. 6 provides evidence that Juxtaposed Univariate Maps and Bertillon Bertillon Maps can outperform the other two techniques for Q4\textsuperscript{Comb} (confirmed in Fig. 7). Finally, for Q5\textsuperscript{sum}, the CIs in both figures provide weak evidence that 3D Choropleths and Deformed Cartograms outperform the other two techniques. In all cases where evidence of a difference are observed, the difference ranges from 13% to 22%.

1Effect size refers to the means we measured. We do not use standardized effect sizes [21], reporting them is not always recommended [2].

2A p-value-approach reading of our results can still be inferred [54].

5.2.2 Completion Time

We analyzed log-transformed data to correct for positive skewness and present antilogged results as is standard for such data analysis processes [90] and common in HCI (e.g., [10, 13, 47, 56, 110]). Consequently, we arrive at geometric means.\footnote{An arithmetic mean uses the sum of values, a geometric mean uses the product of values.} They dampen the effect of extreme trial completion times which tend to bias an arithmetic mean. Results and pair-wise ratios are plotted respectively in Fig. 8 and Fig. 9.

For Q1\textsuperscript{pop}, the CIs in Fig. 8 give evidence that 3D Choropleths can be faster than the other techniques. Fig. 9 confirm this finding and suggest that Choropleth Bertillon Maps could be slower, even though the effect is smaller. For Q2\textsuperscript{Ump}, our results seem to suggest that Choropleth Bertillon Maps are slower than the other equally fast techniques. This is confirmed by the CIs in Fig. 9, providing strong evidence that Juxtaposed Univariate Maps and Bertillon Bertillon Maps can outperform the other two techniques for Q4\textsuperscript{Comb} (confirmed in Fig. 7). Finally, for Q5\textsuperscript{sum}, the CIs in both figures provide weak evidence that Juxtaposed Univariate Maps and Deformed Cartograms outperform the other two techniques. In all cases where evidence of a difference are observed, the difference ranges from 13% to 22%.

Fig. 6. Mean accuracy. Error bars: 95% Bootstrap Confidence Intervals.

Fig. 7. Pairwise differences in accuracy. Error bars: 95% Bootstrap CIs.
approximately 1.3 times slower than Deformed Cartograms. For both $Q^4_{\text{comb}}$ and $Q^5_{\text{comb}}$, the small overlap between Juxtaposed Univariate Maps and the other techniques provides evidence of a difference which is confirmed in Fig. 9. Juxtaposed Univariate Maps is approximately 1.3 times slower than the other three techniques for $Q^4_{\text{comb}}$, and approximately 1.4 times slower than the other three techniques for $Q^5_{\text{comb}}$.

5.2.3 Ranking
Participants’ ranking is shown in Table 2, as planned in the pre-registration. We see that 3D Choropleths was ranked as the favorite technique by the most participants with Choropleth Bertillon Maps not far behind. Deformed Cartograms was ranked as the least favorite by the highest number of participants (23) with the worst mean and median, indicating that this technique is the least favorite overall. The best median scores are obtained by Juxtaposed Univariate Maps, Choropleth Bertillon Maps, and 3D Choropleths while the best mean is obtained by 3D Choropleths. This highlights that there is not a clear favorite technique, but that there might be a slight preference for 3D Choropleths.

5.3 Additional analysis: qualitative feedback
Participants could also comment on the limitations and benefits of each technique. The fully categorized and raw data is available on osf.io/8rxwg/. We summarize here the main insights. Juxtaposed Univariate Maps were reported to be a standard and simple visualization (× 5 participants) and to provide readable information thanks to the two separated maps (× 12). On the other hand, they were said to make the combination of both variables hard (× 15). Having to switch between two maps was also reported as a drawback (× 6). For 3D Choropleths, the main reported issues were: occlusion (× 23), and the difficulty to compare widely separated regions (× 4). Concerning their benefits, it was reported that interaction would make them better—and solve the occlusion problem—(× 5), that comparisons are easy to make (× 5), that the maps are visually appealing (× 5), and that it was easy to combine both pieces of information provided by the visual variables (× 12), though the opposite was also reported (× 2). Deformed Cartograms were reported to make it hard to compare population information (difficulty to compare differently-shaped regions, × 19); to destroy the geography of the country (× 11), to be ugly (× 2), and to make sparsely populated regions hard to see (× 2). Nonetheless, they were reported to be clear (× 2) and visually appealing or interesting (× 4). Finally, Choropleth Bertillon Maps were reported to make all the information easily accessible without having to mentally combine them (× 15), but also to create a lot of overlap (× 16), to be not visually pleasing, or even to be ugly (× 5), and to make region comparisons difficult (× 5).

5.4 Discussions
Unsurprisingly, when focusing on a single variable ($Q^1_{\text{avg}}$ and $Q^2_{\text{avg}}$), Juxtaposed Univariate Maps provide a higher accuracy than other techniques and are relatively fast. When combining the two variables ($Q^3_{\text{avg}}$) it seems that most techniques would perform equally well in completion time and accuracy. This reinforces previous findings comparing some of the designs that we have tested [32]. For $Q^4_{\text{comb}}$, asking participants to combine both variables and to consider geographical information, Juxtaposed Univariate Maps and Choropleth Bertillon Maps seem to give a better accuracy although Juxtaposed Univariate Maps maps are much slower. Finally, 3D Choropleths and Deformed Cartograms seem to lead to particularly accurate answers when participants are asked to combine both variables and aggregate this over several regions ($Q^5_{\text{comb}}$) and are also much faster than Juxtaposed Univariate Maps. These results align with previous work suggesting that Deformed Cartograms can be useful for summarizing tasks. However, 3D Choropleths have not been suggested as interesting summarizing techniques, and our findings are thus more surprising.

Overall, Deformed Cartograms were not really appreciated by participants and only proved to be better for summarizing. Our qualitative results also mirror past results: the deformations hinder the understanding of the map [43, 88]. 3D Choropleths were very appreciated and could be easily improved with interaction to avoid occlusion issues. They provide the information very quickly (Fig. 8) but are not the most accurate technique. The accuracy they have for summarizing tasks makes them an excellent candidate to replace Deformed Cartograms.

Juxtaposed Univariate Maps, often considered as a baseline, are among the most accurate techniques for tasks relying on locating or comparing (mirroring previous findings e.g., [22]), and could thus still be used as a baseline in future studies. Their accuracy, however, seems to come at the cost of time. Finally, despite their popularity and the positive feedback, Choropleth Bertillon Maps did not outperform Juxtaposed Univariate Maps but always had a good overall accuracy, showing that glyphs on choropleth maps can exhibit good performances.

Our data suggest that the accuracy was poor when participants combined information from the two maps or visual variables (~50% for $Q^3_{\text{comb}}$, ~65% for $Q^4_{\text{comb}}$). Currently most choropleth maps displaying social data do not include the separation of the population (e.g., [53, 107]) which can be dramatically misleading [45, 78, 88]. However, even if they showed the distribution of population, our results indicate that such maps could still be misunderstood. While past work [32,86] suggests that the techniques we tested should perform well (because they are made of two separable visual variables) our results indicate that they do not give high accuracy. This suggests that new techniques should be developed and investigated for this type of task. Since the task in $Q^3_{\text{avg}}$ was quite similar to that in $Q^4_{\text{comb}}$, we can hypothesize that combining information from two visual variables on a map requires some training before being accurate enough.

Finally, despite our quite large sample size, we observe in the figures

Table 2. Ranking between most (1) and least (4) favorite technique.

| Technique          | Mean | Med | SD  | 1st | 2nd | 3rd | 4th |
|--------------------|------|-----|-----|-----|-----|-----|-----|
| Juxtaposed Univariate | 2.4  | 2   | 0.9 | 10  | 19  | 19  | 1   |
| 3D Choropleth Maps  | 2.2  | 2   | 1.1 | 19  | 12  | 14  | 9   |
| Deformed Cartogram  | 2.8  | 3   | 1.1 | 8   | 14  | 9   | 23  |
| Choropleth Bertillon| 2.3  | 2   | 1.2 | 17  | 15  | 9   | 13  |
We use the default package of our Geo-Information System to produce (also called dasymetric map) on top of the standard choropleth to convey To exploit the potential benefits of dasymetric maps, as described earlier, we have developed popchart which aims at overlaying data on a choropleth map to augment the fine-scaled population distribution it shows. We have explored four representations described in the following subsections.

6 Popchart: Fine-Scale Population on Choropleths

To exploit the potential benefits of dasymetric maps, as described earlier, we have developed popchart which aims at overlaying data on a choropleth map to augment the fine-scaled population distribution it shows. We have explored four representations described in the following subsections.

6.1 Popchart dasymetric overlay maps

Popchart dasymetric overlay maps overlay a transparent choropleth (also called dasymetric map) on top of the standard choropleth to convey population. With two levels of granularity we use another transparent colour on top of the coloured region to display population. In all the maps we used black to encode local population data. (Fig. 10a) shows large cities and suburbs but the larger population density along valleys and rivers are not clearly visible. Here the total population is given by the sum of the perceived opacities. This technique can suffer from the same perception bias as choropleth maps in that a city presenting a larger area can be perceived as having a greater total population.

6.2 Popchart Dot Maps

Popchart Dot Maps use dots to represent population per city. Large dots can overlap suburbs as seen in Fig. 10b for Lyon and St Etienne. Still, it clearly shows population distribution along rivers, coasts and valleys (e.g., south of Lyon). Here the total population is given by the sum of the sizes of the circles (when they are visible).

6.3 Popchart Heatmap Maps

Popchart Heatmap Maps use heatmap overlays. Heatmap is a popular technique to aggregate a large number of points to display on a map. We use the default package of our Geo-Information System to produce a heatmap from transparent (low population density) to black (high population density). Large cities are visible but their centre can be harder to identify: the ‘fuzzy’ nature of heatmaps makes it difficult to distinguish the distribution of population between a main city and its suburbs. For instance, in Fig. 10c, it is difficult to see that there are more people in the East of Lyon than in the West. Here the total population is relative to the total amount of ‘ink’ present in a region.

6.4 Popchart 3D PrismMaps

Popchart 3D PrismMaps, relies on height to display the number of inhabitants at the city level. The 3D extrusion for each city is based on the limits of the city and the height is defined by the number of inhabitants. The colour of the top and border of the 3D volumes encodes the statistical value associated with the enclosing region. As people tend to be concentrated in a few big cities, most of the volumes are very flat making the resulting map appear similar to a tilted standard choropleth. Among these flat coloured volumes one can distinguish big cities as very high ‘skyscrapers’ (Fig. 10d) sometimes surrounded by lower ‘towers’ in the suburbs. Only a few cities, behind the biggest ones, tend to be occluded by the 3D perspective. The total population of a given region is given by the sum of all the heights of the volumes.

Fig. 10. Cities of St Etienne(S), Lyon(L) and Grenoble(G) at higher level of detail using Popchart dasymetric overlay maps, Popchart Dot Maps, Popchart Heatmap Maps and Popchart 3D PrismMaps.

fig. 11. Answers to Likert-scale ratings from visualization researchers.

7 Initial Experiment with Visualization Researchers

To find a set of techniques for a controlled experiment, we conducted an investigation with visualization researchers to get feedback on the techniques described in Sect. 6. We follow the same procedure described in Sect. 4 and describe our tasks (Table 3) using Roth’s taxonomy [86].

The same 10 visualization researchers answered. Results are presented in Fig. 11. We counted how often the number of (Very) Difficult is greater or equal to the number of (Very) Easy scores for each technique. If it happened more than twice, we looked at the qualitative feedback to determine whether we should exclude a technique. Popchart dasymetric overlay maps (Q5, Q6, Q4,) and Popchart 3D PrismMaps (Q2, Q3, Q4, Q5, Q6) were candidates to be excluded. Popchart dasymetric overlay maps was reported as difficult to compare (×3) and interpret (×4). We therefore excluded it from our experiment. Popchart 3D PrismMaps was, despite its low ratings, generally well perceived (4 participants reported that all information could be computed with this technique) so we kept it in the pool of techniques.

Based on this experiment, we thus decided to keep popchart 3D prismMaps, popchart Heatmap maps, and Popchart Dot Maps.

8 Controlled Experiment 2

We conducted a second experiment to evaluate which of the previously defined techniques would be more helpful to convey the spatial distribution and the statistic of interest with a fine-scale distribution of population. The code used for the experiment is available at tiny.cc/mapstudylong with a frozen preregistration at osf.io/svbb8a.

8.1 Participants, Questions, and Techniques

Invitations to participate were sent by email, including a link to the study, to students and researchers working in computer science and in other fields. The email also invited them to forward the experiment.

We again used a time-based stopping rule (10 days) to preregister a sample size. The previously identified techniques used in the experiment are popchart 3D prismMaps, popchart Heatmap maps, Popchart Dot Maps in a counterbalanced order to avoid learning effects. We also used the data on France’s unemployment and population.

The tasks were also taken from the initial experiment with researchers. To avoid learning effects, 3 different, but equivalent, sets of questions were prepared. Each set contained 4 questions (Q1 (comb), Q2 (comb), Q3 (comb), Q4 (comb)) and were assigned, in turn, to different visualization techniques. For each question, areas of interest were highlighted so that the completion time does not contain the search time but only the time it took participants to answer the question. To limit possible confounding factors we did not give participants the ability to interact with the map.
Accuracy results are shown in Fig. 12, pairwise differences in Fig. 13. A total of 47 participants (17 females, aged from 19 to 57, mean = 26.8, median = 26, SD = 6.8) completed this experiment. None of their data was excluded with our preregistered exclusion criteria. Data analysis is conducted following the methodology described in Sect. 5.2.

### 8.2 Planned Analysis Results

A total of 47 participants (17 females, aged from 19 to 57, mean = 26.8, median = 26, SD = 6.8) completed this experiment. None of their data was excluded with our preregistered exclusion criteria. Data analysis is conducted following the methodology described in Sect. 5.2.

#### 8.2.1 Accuracy

Accuracy results are shown in Fig. 12, pairwise differences in Fig. 13. The small overlap between Popchart Heatmap Maps and the other two techniques in Fig. 12 shows that Popchart Heatmap Maps has a lower accuracy than the other two, performing similarly well (confirmed in Fig. 13), for Q1_map. Concerning Q2_comb, the CIs in both figures indicate no difference between the 3 techniques. For Q3_comb, the configuration of the CIs strongly indicates that Popchart Dot Maps is more accurate than Popchart 3D PrismMaps and we see evidence that Popchart Dot Maps is more accurate than Popchart Heatmap Maps with a smaller effect size. Finally, for Q4_map, we see no evidence for a difference between Popchart Dot Maps and Popchart Heatmap Maps but both are more accurate than Popchart 3D PrismMaps. This is confirmed in Fig. 13.

#### 8.2.2 Completion Time

The results are presented in Fig. 14 and pairwise ratios in Fig. 15. For Q1_map, while the overlap of CIs would suggest that Popchart 3D PrismMaps is slower than the other two techniques, individual differences in Fig. 15 leads us to argue that evidence is quite weak in this case. The configuration of CIs in both figures for Q2_comb and Q3_comb suggest that our data does not provide evidence of a difference between the 3 techniques. For Q4_map, we have strong evidence that Popchart 3D PrismMaps is almost 1.5 times slower than Popchart Heatmap Maps and approximately 1.3 times slower than popchart dot map while these two techniques seem to achieve similar performance.

### 8.2.3 Ranking

The ranking data is presented in Table 4. We can see that Popchart Dot Maps were most frequently ranked as the favorite technique. Popchart Heatmap Maps and Popchart 3D PrismMaps have similar results in this respect, but Popchart 3D PrismMaps were ranked as the least favorite by the highest number of participants (× 19) and has the worst mean. This suggests that this technique could be the least popular overall. The best mean score was obtained by Popchart Dot Maps which seems to be the overall favorite. All differences are, however, quite small.

### 8.3 Additional analysis: qualitative feedback

Participants were invited to reflect on the limitations and benefits of the techniques. The fully categorized and raw data is available on osf.io/8rxw9y/. Popchart 3D PrismMaps were praised for their capacity to increase readability (× 2 participants), to avoid overlaying information on top of the map (× 3). However, the information was considered hard to aggregate over a region (× 3), the angle of the map was reported...
as problematic (× 6), smaller cities were difficult to perceive (× 7), the 3D bars were too transparent (× 5) and height was said to be hard to compare (× 7). Popchart Dot Maps were reported as easy to read (× 13), a standard visualization (× 2) even though the circles representing cities could occlude other cities (× 18) or occlude the geographical information or even the information presented by the choropleth itself (× 18). Popchart Heatmap Maps were reported as easy to read (× 9) but participants stated that they were hard to compare (× 6), could hide geographical information (× 7) and were not visually pleasing (× 3).

8.4 Discussion
To identify the city with the highest population density, Popchart 3D PrismMaps seem to show comparable accuracy (but worse completion time) as other methods except for Q3_cont asking for geographical information to be considered. There Popchart 3D PrismMaps performed poorly. This could be explained by the extra geographical task.

For combination of unemployment data and population of some cities (Q2_cont and Q3_cont) it seems that Popchart Dot Maps might have a slight advantage in terms of accuracy—visible mostly in Q3_cont. All techniques seemed to exhibit a similar completion time.

To average the population of an entire region (Q4_cont), Popchart 3D PrismMaps performed the worst in terms of both time and accuracy, while the other two techniques seemed to perform equally well. Interestingly, Popchart Dot Maps and Popchart Heatmap Maps perform very well for such tasks (~80%), thus validating that one can display information at a finer scale while still leaving the possibility for users to aggregate per region quite accurately. We can deduce that readers of Popchart maps can still answer most of the tasks that were presented in the first experiment we conducted. This is a strong advantage over the techniques explained and studied in Sect. 3: Popchart maps can therefore provide even more precise information without sacrificing the original information conveyed by choropleth maps.

The low performance for tasks focusing on a city population and unemployment of a region (Q2_cont and Q3_cont with ~65% accuracy and, at most, ~55% accuracy) is also particularly interesting. It calls for more techniques to tackle the issue. This is supported by the qualitative feedback gathered which highlights that the technique that performs the worst (Popchart 3D PrismMaps) has the advantage over the other two of not hiding any kind of information on the map, as was identified as a drawback of both Popchart Dot Maps and Popchart Heatmap Maps.

9 Conclusion, Limitations, and Future Work

With the development of web technologies, maps can now easily be made and shared by lay people or researchers in social sciences [66]. In this context, understanding the pitfalls and benefits of specific designs for communication of research results to lay people is therefore particularly important, in particular since choropleth maps are inherently prone to lead to misinterpretation [1, 44, 64, 88, 94]. With that objective in mind we proposed several visual representations of bivariate choropleth maps and evaluated them through two studies.

Our experiments of course have some limitations. First of all we could have considered additional techniques such as colour-based bivariate maps. Maps have been studied quite extensively and it was impossible to be exhaustive. Another limitation lies in the loading time of the maps which influenced the completion time. To avoid this issue we could have used simple screenshots but we would have lost all interactivity (e.g., providing labels on hovering) and opting for such a solution would have made the experiment much harder. On a similar note, interaction could have changed the performance of some techniques (notably for techniques relying on 3D graphics). This is something we intend to investigate in future work. While giving the possibility to rotate/drag the map and change the viewing angle would help overcome most of the limitations of maps relying on 3D graphics (in our case popchart 3D PrismMaps) they could also lead to problems in interpretation. Indeed, arbitrarily-rotated maps could be more difficult to make sense of if the user cannot recognize the geography. Giving the possibility for participants to interact may also have led to longer completion times, or a much higher cognitive load if such interactions are carried out through classical input techniques [13]. We hypothesize, however, that constrained interaction allowing only small adjustments of the pitch/viewing angle and small rotations can lead to ideal performance by overcoming the occlusion issue while keeping the map close to a specific orientation that does not impede the user’s sense-making process and would limit the impact on cognitive load. Finally, we have currently tested our technique with a single dataset and studies with different datasets should further confirm our results.

While 3D representations are seldom used in information visualization, our study results are consistent with previous studies and highlight the potential of 3D CartoMaps, performing as well as deformed cartogram maps in particular for tasks involving summarizing information from different regions to get a big-picture message, for which deformed cartogram maps have been praised [74]. For both dasymetric and regular choropleth maps, the results we have obtained are promising but the interpretation of population data is still an impossible to be exhaustive. Another limitation lies in the loading time of the maps which influenced the completion time. To avoid this issue we could have used simple screenshots but we would have lost all interactivity (e.g., providing labels on hovering) and opting for such a solution would have made the experiment much harder. On a similar note, interaction could have changed the performance of some techniques (notably for techniques relying on 3D graphics). This is something we intend to investigate in future work. While giving the possibility to rotate/drag the map and change the viewing angle would help overcome most of the limitations of maps relying on 3D graphics (in our case popchart 3D PrismMaps) they could also lead to problems in interpretation. Indeed, arbitrarily-rotated maps could be more difficult to make sense of if the user cannot recognize the geography. Giving the possibility for participants to interact may also have led to longer completion times, or a much higher cognitive load if such interactions are carried out through classical input techniques [13]. We hypothesize, however, that constrained interaction allowing only small adjustments of the pitch/viewing angle and small rotations can lead to ideal performance by overcoming the occlusion issue while keeping the map close to a specific orientation that does not impede the user’s sense-making process and would limit the impact on cognitive load. Finally, we have currently tested our technique with a single dataset and studies with different datasets should further confirm our results.

While 3D representations are seldom used in information visualization, our study results are consistent with previous studies and highlight the potential of 3D CartoMaps, performing as well as deformed cartogram maps in particular for tasks involving summarizing information from different regions to get a big-picture message, for which deformed cartogram maps have been praised [74]. For both dasymetric and regular choropleth maps, the results we have obtained are promising but the interpretation of population data is still an

### Acknowledgments

We thank Jean-Daniel Fekete for his feedback and our participants.

### References

[1] M. J. Alam, S. G. Kobourov, and S. Veeramoni. Quantitative measures for cartogram generation techniques. *Computer Graphics Forum*, 34(3):351–360, 2015. doi: 10.1111/cgf.12467

[2] T. Baguley. Standardized or simple effect size: What should be reported? *British Journal of Psychology*, 100(3):603–617, Aug. 2009. doi: 10.1348/000712608X377117

[3] M. Baker. Statisticians issue warning over misuse of P values. *Nature*, 531(7593):151, Mar. 2016. doi: 10.1038/nature.2016.15903

[4] A. Balbi and A.-M. Guerry. Statistique comparée de l'état de l'instruction et du nombre des crimes, 1829.

[5] L. V. Barroso, R. P. Pérez-Machado, C. Small, and W. Cabral-Miranda. Changing spatial perception: dasymetric mapping to improve analysis of health outcomes in a magacity. *Journal of Maps*, 12(5):1242–1247, 2016. doi: 10.1080/17445647.2015.1101403

[6] J. Bertillon. *Atlas de statistique graphique de la ville de Paris, année 1889*, tome II. Préfecture du département de la Seine, Secrétariat général, Service de la statistique municipale, 1891.

[7] J. Bertillon. *Résultats statistiques du dénombrement de 1891 pour la ville de Paris et le département de la Seine et renseignements relatifs aux dénombrements antérieurs*, p. 73. Préfecture de la Seine, 1891.

[8] J. Bertillon. *Résultats statistiques du dénombrement de 1891 pour la ville de Paris et le département de la Seine et renseignements relatifs aux dénombrements antérieurs*, p. 84. Préfecture de la Seine, 1891.

[9] J. Bertin. *Sémioséologie graphique: les diagrammes*. Montréal, 1967.

[10] L. Besançon, M. Ammi, and T. Isenberg. Pressure-Based Gain Factor Control for Mobile 3D Interaction using Locally-Coupled Devices. In...
