

Task Taxonomy for Cartograms

Sabrina Nusrat and Stephen Kobourov
Department of Computer Science, University of Arizona, USA

Abstract

Cartograms are maps in which areas of geographic regions (countries, states) appear in proportion to some variable of interest (population, income). Despite the popularity of cartograms and the large number of cartogram variants, there are few studies evaluating the effectiveness of cartograms in conveying information. In order to design cartograms as a useful visualization tool and to be able to compare the effectiveness of cartograms generated by different methods, we need to study the nature of information conveyed and the specific tasks that can be performed on cartograms. In this paper we consider a set of cartogram visualization tasks, based on standard taxonomies from cartography and information visualization. We then propose a cartogram task taxonomy that can be used to organize not only the tasks considered here but also other tasks that might be added later.

1. Introduction

A *cartogram*, or a value-by-area map, is a representation of a map in which geographic regions are modified to reflect a statistic, such as population or income. Geographic regions, such as countries, states and provinces of a map, are scaled by area to visualize some statistical information, while attempting to keep the overall result readable and recognizable. This kind of visualization has been used for many years; in fact, the first reference to the term “cartogram” dates back to at least 1868, and Émile Levasseur’s rectangular cartograms used in an economic geography textbook [Tob04]. Since then cartograms have been studied by geographers, cartographers, economists, social scientists, geometers, and information visualization researchers.

Likely due to aesthetic appeal and the possibility to visualize data and put political and socioeconomic reality into perspective, cartograms are widely used in newspapers, magazines, textbooks, blogs, and presentations. Cartograms are frequently used to illustrate population and GDP distributions, election results, and migration patterns. Popular TED talks use cartograms to show how the news media make us perceive the world [Mil08] and to visualize the complex risk factors of deadly diseases [Ros09]. Cartograms continue to be used in textbooks, for example, to teach middle-school and high-school students about global demographics and human development [Pel].

A cartogram should enable the viewer to quickly and correctly interpret the data encoded in the visualization. Therefore, it is important to clearly define the visualization goals and a set of tasks that are suitable for cartogram visualiza-

tions. There is a broad spectrum of methods for generating cartograms: some distort shapes, some replace shapes with geometric objects, some use colors in addition to the shape changes. Although there is a rich literature on generating cartograms, there is little work on evaluating the usability of cartograms and their effectiveness. In order to compare cartograms generated by different methods we need to understand the visualization goals and to explore the possible tasks suitable for cartograms. With this in mind, we consider a set of cartogram visualization tasks, based on standard taxonomies from cartography and information visualization. We then propose a cartogram task taxonomy that can be used to organize not only the tasks considered here but also other tasks that might be added later.

2. Related Work

There are many task taxonomies in information visualization and cartography. Visualization tasks have been defined and classified, often depending on the context and scope of the tasks. Wehrend [Weh93] defines “visualization goals” as actions a user may perform on her data and presents nine such goals: *identify*, *locate*, *distinguish*, *categorize*, *cluster*, *rank*, *compare*, *associate*, *correlate*. Wehrend’s work is extended by Zhou and Feiner [ZF98] by defining “visualization techniques” as low-level operations and “visual tasks” as interfaces between high-level presentation intents and low-level visual techniques without specifying exactly “how” an operation is done. Andrienko et al. [AAG03] list *identify* and *compare* as cognitive operations for visualizing spatio-temporal data. Some recent taxonomies do not include *iden-*

tify and *compare*, but rather use terminology more common in statistics. For example, Amar et al. [AES05] present a list of low-level tasks, such as *retrieve value*, *filter*, *find extremum* and *sort*, that capture people's activities when using information visualization tools for understanding data. While the above discussion covers a general set of tasks for information visualization, it is often useful to categorize them across different dimensions. The typology of abstract visualization tasks proposed by Brehmer and Munzner [BM13] focuses on three questions: *why* is a task performed, *what* are the inputs and outputs, and *how* is the task performed. Schulz et al. [SNHS13] consider six questions: *why* is a visualization task performed, *how* is a task carried out, *what* does a task seek, *where* in the data does a task operate, *when* is it performed, and *who* is executing a task. These questions relate to the goals of the tasks, the means, the characteristics, the target and cardinality of data entities, the order of the tasks, and the type (expert/non-expert) of audience.

Cartography is the science and practice of making and using maps. Roth [Rot13] classifies taxonomical frameworks into three types: objective-based, operator-based, and operand-based. The first type focuses on user intent, or what the user wishes to perform. Examples include *identify*, *compare* etc. Taxonomies discussed in the previous paragraphs are mostly of this type. Operator-based taxonomies focus on operators in cartographic interfaces that support the objective of users. Example operators include *brushing* [She95, Dyk97], *focusing* [BCS96, ME00], *zooming* [Shn96, EAAB08], and *linking* [BCS96, DE98]. In operand-based taxonomies, the focus is on the operand, or the object with which the user is interacting. In the context of interactive cartography, the taxonomy provided by Andrienko et al. [AAG03] is noteworthy for both operator-based and operand-based taxonomies.

There is a wide variety of methods to generate cartograms, broadly categorized in four types: contiguous, non-contiguous, Dorling, and rectangular [KS07]. In contiguous cartograms the original geographic map is modified (by pulling, pushing, and stretching the boundaries) to change the areas. Among these cartograms, the most popular method is the diffusion-based method proposed by Gastner and Newman [GN04]. Others of this type include the rubber-map method by Tobler [Tob73], contiguous cartograms by Dougenik et al. [DCN85], CartoDraw by Keim et al. [KNPS03], constraint-based continuous cartograms by House and Kocmoud [HK98], and medial-axis-based cartograms by Keim et al. [KPN05]. More recent are circular arc cartograms [KKN13]. Non-contiguous cartograms are generated by starting with the regions of the given map and scaling down each region independently, so that the desired size/area is obtained [Ols76]. Dorling cartograms represent regions in the map by circles [Dor91]. Data values are realized by circle size: the bigger the circle, the larger the data value. Rectangular cartograms, as their name in-

dicates, use rectangles to represent the regions in a map. Rectangular cartograms have been used for more than 80 years [Rai34]. More recent rectangular cartogram methods include [BSV12, KS07]. Other topological variants include rectangular hierarchical cartograms [SDW10] and rectilinear cartograms [dBMS10, ABF*13].

Quantitative measures for evaluating different cartogram types have been proposed [AKVar] and a good survey of cartogram methods can be found in [Tob04].

3. Task Taxonomy for Cartograms

Although there are many excellent task taxonomies in cartography, information visualization and human-computer interaction, visualization goals and tasks are not clearly defined for cartograms. With this in mind, we adapt existing tasks from cartography and information visualization and add new tasks, particularly suitable for cartograms. We categorize these tasks along four dimensions, based on the questions *why*, *how*, *what*, and *where*. We believe our list of visualization tasks and their classification can be used in formal evaluations of various cartogram generation methods, and the analysis of the goals and tasks suitable for cartograms can potentially improve future cartogram design.

3.1. Analytic Tasks and Visualization Goals

Most cartograms are modified geographic maps which combine two features typically not present in other maps and charts: (1) they contain geographical statistical information, (2) they contain location information. Therefore, cartograms have the advantage of allowing traditional cartographic tasks, as well as information visualization tasks about the encoded statistic. Through discussions with information visualization experts and using the affinity diagramming approach we put together a set of cartogram tasks. Some of the tasks are adapted from existing information visualization and cartography taxonomies; others are particularly relevant to cartograms. We list the tasks below, along with a general description and specific examples.

1. Detect change (compared to a base map): This is a new task proposed for cartograms that is not present in other taxonomies. In cartograms the size of a country is changed in order to realize the input weights. Since change in size (i.e., whether a region has grown or shrunk) is a central feature of cartograms, the viewer should be able to detect such change. According to Dent [Den75], this is a crucial aspect of effective cartogram communication.

Example Task: Given a population cartogram of the USA, can the viewer detect if the state of California has grown or shrunk compared to its size in geographic map?

2. Locate: The task is to search and find a country in a cartogram. In some taxonomies this task is denoted as *locate* and in others as *lookup*. Brehmer and Munzner [BM13]

differentiate between *locate* and *lookup* tasks: in the context of cartograms, if the viewer is familiar with the USA, she can simply lookup California while an unfamiliar viewer has to search and locate California first. Since cartograms often drastically deform an existing map, even if the viewer is familiar with the underlying maps, finding something in the cartogram might not be a simple lookup.

Example Cartogram Task: Given a population cartogram of the USA, locate the state of California.

3. Recognize: One of the goals in generating cartograms is to keep the original map recognizable, while distorting it to realize the given statistic. Therefore, this is an important task in our taxonomy. The aim of this task is to find out if the viewer can recognize countries from the original map when looking at the cartogram.

Example Cartogram Task: Given the shape of a state from the original map and shapes of two states from the cartogram, find out which of the two cartogram states corresponds to the state from the original map.

4. Identify: The *identify* task has been used in many taxonomies but conveys slightly different meanings. It can mean geographic search in space, e.g., “identify your house based on an aerial image in Google Earth” or temporal search, e.g., “when will the bluff erosion reach my house?”, or an attribute search, e.g., “what is the range of the endangered species?” In our taxonomy we use *identify* for attribute or characteristic search, as in Brehmer and Munzner [BM13]. Identify focuses on a single object.

Example Cartogram Task: In a red-blue cartogram, identify the winning candidate for the state of California.

5. Compare: The *compare* task is another very commonly used one in objective-based taxonomies [Rot13, WL90, Web93]. This task has also been used in a qualitative study of cartograms [War98]. This task is unambiguous, and usually asks for similarities or differences between attributes. We use it in the same way in our taxonomy.

Example Cartogram Task: Given a population cartogram of the USA, compare two states by size.

6. Find top-k: This is another commonly used task in visualization. Here the goal is to find k entries with the maximum (or minimum) values of a given attribute. This is a generic task that covers specific tasks, such as “Find extremum”, where the goal is to find the data with the extreme value [AES05]; and “Sort”, where all the data entries are ordered based on the value of a given attribute.

Example Cartogram Task: Given a population cartogram, find out which state has the highest/lowest population.

7. Filter: The *filter* task asks to find data cases satisfying some criteria about a given attribute, e.g., [AES05]. That is, the viewer can filter out examples that fail the criteria. We use this task in the same way in our taxonomy.

Example Cartogram Task: Find states which have higher population than the state of California.

8. Find adjacency: Some cartograms preserve the given topology, others do not. In order to understand the map characteristics properly, it is important to identify the neighboring states of a given state. Thus, this is an important new task for visualizing cartograms.

Example Cartogram Task: Given a cartogram, find all states adjacent to California.

9. Cluster: The goal of the *cluster* task is to find objects with similar attributes. We use it the same way.

Example Cartogram Task: In a cartogram showing obesity rates, find states with similar obesity rate as California

10. Summarize (Analyze / compare distributions): Cartograms are most often used to convey a “big picture”. The *summarize* task is one that asks the viewer to see the big picture. This task is associated with summarizing/overviewing the data shown in the cartogram, as well as with finding global distribution patterns.

Example Cartogram Task: In a red-blue election results cartogram, find if it is a close election, or a “landslide win”.

3.2. Classification of Tasks

We categorize the possible tasks for visualizing and interpreting information in cartograms along four dimensions: *goals, means, characteristics, and cardinality*; see Table 1 for a summary. Our classification is based on three foundational typologies by Bertin [Ber83], Brehmer and Munzner [BM13] and Schulz et al. [SNHS13].

Goals: why is a task performed? The goal, or objective of a task does not define the task itself, rather the reason why it is being performed. We identified five goals for cartograms.

1. Query: Tasks in this group are usually local tasks; they focus on one or two objects. Some of the tasks may require comparing a state in the cartogram with the state in the original map. These tasks do not require searching through the map, for e.g.: *recognize, detect change*.
2. Explore: Tasks in this group require searching through the cartograms, comparing data, and finding relation among datasets, for e.g.: *find extremum* and *cluster*.
3. Extract: Some tasks require extracting metadata; such tasks fall in this group. An example task is *identify*.

Means: how is a task carried out? The means of visualization tasks do not define the tasks themselves, but rather explain how the tasks can be performed [SNHS13]. We have identified three different means.

1. Navigation: One of the methods for performing visualization tasks is to navigate or browse through the dataset. Example navigation tasks are: *locate, find adjacency*.

	Goals			Means				Characteristics		Cardinality		
	Query	Search	Extract	Map Relation	Data Relation	Navigation	Derive	Low Level	High-Level	Single	Multiple	All
Recognize	✓	×	×	✓	×	×	×	✓	×	✓	×	×
Detect Change	✓	×	×	✓	×	×	×	✓	×	✓	×	×
Compare	✓	×	×	×	✓	×	×	✓	×	×	✓	×
Find top- <i>k</i>	×	✓	×	×	✓	×	×	×	✓	×	×	✓
Filter	×	✓	×	×	✓	×	×	×	✓	×	×	✓
Cluster	×	✓	×	×	✓	×	×	×	✓	×	×	✓
Locate	×	✓	×	×	×	✓	×	×	×	×	×	✓
Find Adjacency	×	✓	×	×	×	✓	×	×	✓	×	✓	×
Summarize	×	×	✓	×	×	×	✓	×	✓	×	×	✓
Identify	×	×	✓	×	×	×	✓	✓	×	✓	×	×

Table 1: Tasks and their dimensions

2. Relation: This includes all means to find some relation (e.g., similarity or difference). For cartograms, we further subdivide Relation into:
 - a. Relation across geography or data-relation: these require finding a relation in the data. Example data relation tasks are: *compare*, *cluster*.
 - b. Relation across visualization or map-relation: these require finding a relation between the original map and the cartogram. Example tasks are: *recognize*, *detect change*.
3. Derive: The tasks in this group are performed by extraction of information, or abstraction of the data. This often involves augmentation, reduction, or filtering of data. Example derive tasks are *identify* and *summarize*.

Characteristics: what are the features of a task? This dimension does not define the task itself, rather identifies what is the level of complexity of the visualization task. Characteristics or features of a visualization task depend highly on the type of information that the task aims to reveal [SNHS13]. In the context of cartograms, these characteristics can be divided into two categories:

1. Low-level data characteristics: involve simple tasks that can be performed by observation from the visualization. Example tasks for cartograms: *identify*, *locate*, *compare*.
2. High-level data characteristics: involve more complex tasks that need to be deduced from the visualization. Example tasks for cartograms include: *filter*, *cluster*, *sort*.

Cardinality: where in the data a task operates? The cardinality of a task specifies where the task operates. This dimension directly relates tasks with the components of data. The reading levels by Bertin [Ber83] contain three types: *elementary*, *intermediate* and *overall*, and they deal with a single data element, multiple elements and all elements, respectively. Similar differentiation is made by Schulz et al. [SNHS13] and Yi et al. [YEL10]. Thus, the cardinality of a cartogram task differentiates the number of regions that are investigated by a task: a single region, multiple re-

gions, or the entire map. Example tasks in cartograms that consider a *single instance* are: *detect change*, *recognize*. Example tasks with *multiple instances* are *compare*, *find adjacency*, and those with *all instances* are *summarize*, *cluster*.

4. Discussion, Limitations and Conclusions

Based on existing taxonomies from cartography and information visualization, we propose a taxonomy specifically designed with cartograms in mind. We categorize tasks in multiple dimensions that can be useful in the evaluation of different types of cartograms. For example, the first two tasks, “recognize” and “detect change”, have similar goals, means, characteristics and cardinality as they both deal with the shapes and sizes of regions; see Table 1. Based on similar patterns we have grouped “compare”, “find top-*k*”, “filter”, and “cluster” as they deal with size comparison. Finally, “locate” and “find adjacency” form a group and “summarize” and “identify” form a group.

As in other taxonomies, there are tasks that are compound and depend on simpler tasks. For example, we have tasks that are “low-level” and tasks that are “high-level”. In order to pursue high-level tasks (e.g. “find top-*k*”) we often need to perform multiple low-level tasks (e.g., “compare”).

Given the many different types of cartograms, it is impossible to impose uniform cartogram requirements, but a comprehensive collection of tasks should make a fair evaluation possible. To cover a variety of cartogram-specific tasks in such an evaluation, it would suffice to pick one task from each of the four groups, but a thorough evaluation will require at least seven tasks (as the last three groups have two distinct goals/means/characteristics/cardinality patterns). While a single taxonomy is rarely complete and covers all possible tasks and task dimensions, the proposed taxonomy can be a useful guideline for the design and evaluation of cartograms and we have recently used it in an evaluation of four major types of cartograms [NAK15].

References

- [AAG03] ANDRIENKO N., ANDRIENKO G., GATALSKY P.: Exploratory spatio-temporal visualization: an analytical review. *J. of Vis. Lang. and Comp.* 14, 6 (2003), 503–541. 1, 2
- [ABF*13] ALAM M. J., BIEDL T. C., FELSNER S., KAUFMANN M., KOBOUROV S. G., UECKERDT T.: Computing cartograms with optimal complexity. *Discrete & Computational Geometry* 50, 3 (2013), 784–810. 2
- [AES05] AMAR R., EAGAN J., STASKO J.: Low-level components of analytic activity in information visualization. *IEEE Symp. on Info. Vis.* (2005), 111–117. 2, 3
- [AKVar] ALAM M. J., KOBOUROV S. G., VEERAMONI S.: Quantitative measures for cartogram generation techniques. *Computer Graphics Forum, special issue on EuroVis'15* (2015 (to appear)). 2
- [BCS96] BUJA A., COOK D., SWAYNE D. F.: Interactive high-dimensional data visualization. *Journal of Computational and Graphical Statistics* 5 (1996), 78–99. 2
- [Ber83] BERTIN J.: *Semiology of Graphics: Diagrams, Networks, Maps*. The University of Wisconsin Press, Madison, 1983. 3, 4
- [BM13] BREHMER M., MUNZNER T.: A multi-level typology of abstract visualization tasks. *IEEE Transaction on Visualization and Computer Graphics* 19, 12 (2013), 2376–2385. 2, 3
- [BSV12] BUCHIN K., SPECKMANN B., VERDONSCHOT S.: Evolution strategies for optimizing rectangular cartograms. In *GIS'12* (2012), vol. 7478, pp. 29–42. 2
- [dBMS10] DE BERG M., MUMFORD E., SPECKMANN B.: Optimal BSPs and rectilinear cartograms. *Int. J. Comp. Geo. & App.* 20, 2 (2010), 203–222. 2
- [DCN85] DOUGENIK J. A., CHRISMAN N. R., NIEMEYER D. R.: An algorithm to construct continuous area cartograms. *The Professional Geographer* 37, 1 (1985), 75–81. 2
- [DE98] DIX A., ELLIS G.: Starting simple - adding value to static visualisation through simple interaction. In *Working Conf. on Ad. Vis. Interfaces* (1998). 2
- [Den75] DENT B. D.: Comm. aspects of value-by-area cartograms. *The American Cartographer* 2, 2 (1975), 154–168. 2
- [Dor91] DORLING D.: *The vis. of spatial structure*. PhD thesis, Univ. of Newcastle-upon-Tyne, 1991. 2
- [Dyk97] DYKES J.: Exploring spatial data representation with dynamic graphics. *Comp. & Geosc.* 23 (1997), 345–370. 2
- [EAAB08] EDSALL R., ANDRIENKO G., ANDRIENKO N., BUTTENFIELD B.: Interactive maps for exploring spatial data. In *Manual of Geo. Info. Sys.* (2008), pp. 837–858. 2
- [GN04] GASTNER M. T., NEWMAN M. E. J.: Diffusion-based method for producing density-equalizing maps. In *Nat. Aca. of Sc. of the USA* (2004), vol. 101, pp. 7499–7504. 2
- [HK98] HOUSE D. H., KOCMOUD C. J.: Continuous cartogram construction. In *IEEE Visualization* (1998), pp. 197–204. 2
- [KKN13] KÄMPER J.-H., KOBOUROV S. G., NÖLLENBURG M.: Circular-arc cartograms. In *IEEE Pacific Vis. Symp. (PacificVis'13)* (2013), pp. 1–8. 2
- [KNPS03] KEIM D., NORTH S., PANSE C., SCHNEIDEWIND J.: Visualizing geographic information: VisualPoints vs. CartoDraw. *Info. Vis.* 2, 1 (2003), 58–67. 2
- [KPN05] KEIM D. A., PANSE C., NORTH S. C.: Medial-axis-based cartograms. *IEEE Computer Graphics and Applications* 25, 3 (2005), 60–68. 2
- [KS07] KREVELD M., SPECKMANN B.: On rectangular cartograms. *Computational Geometry* 37 (2007), 175–187. 2
- [ME00] MASTERS R., EDSALL R.: Interaction tools to support knowledge discovery: A case study using data explorer and Tcl/Tk. In *Visualization Development Environments* (2000). 2
- [Mil08] MILLER A.: TED talk: The news about the news. <http://goo.gl/pn4XLB>, 2008. [Online; accessed 10/20/2014]. 1
- [NAK15] NUSRAT S., ALAM M. J., KOBOUROV S. G.: Evaluating cartogram effectiveness. *CoRR abs/1504.02218* (2015). 4
- [Ols76] OLSON J.: Noncontiguous area cartogram. *The Professional Geographer* 28, 4 (1976), 371–380. 2
- [Pel] PELKOFER T.: Using cartograms to learn about Latin America. <http://goo.gl/yB8nQs>. [Online; accessed 10/20/2014]. 1
- [Rai34] RAISZ E.: The rectangular statistical cartogram. *Geographical Review* 24, 3 (1934), 292–296. 2
- [Ros09] ROSLING H.: TED Talk: Insights on HIV, in stunning data visuals. <http://goo.gl/DJ4LYP>, 2009. [Online; accessed 10/20/2014]. 1
- [Rot13] ROTH R. E.: An empirically-derived taxonomy of interaction primitives for interactive cartography and geovisualization. *IEEE Trans. on Vis. and Comp. Graphics* 19, 12 (2013), 2356–2365. 2, 3
- [SDW10] SLINGSBY A., DYKES J., WOOD J.: Rectangular hierarchical cartograms for socio-economic data. *J. of Maps* 6, 1 (2010), 330–345. 2
- [She95] SHEPHERD I.: *Putting time on the map: Dynamic displays in data visualization and GIS*. 1995. 2
- [Shn96] SHNEIDERMAN B.: The eyes have it: A task by data type taxonomy for information visualizations. In *IEEE Symp. on Vis. Lang.* (1996), pp. 336–343. 2
- [SNHS13] SCHULZ H.-J., NOCKE T., HEITZLER M., SCHUMANN H.: A design space of visualization tasks. *IEEE Trans. on Vis. and Comp. Graphics* 19, 12 (2013), 2366–2375. 2, 3, 4
- [Tob73] TOBLER W.: A continuous transformation useful for districting. *Annals, NY Aca. of Sc.* 219 (1973), 215–220. 2
- [Tob04] TOBLER W.: Thirty five years of computer cartograms. *Annals, Assoc. of American Geo.* 94 (2004), 58–73. 1, 2
- [War98] WARE J. A.: *Using animation to improve the communicative aspect of cartograms*. PhD thesis, Michigan State University, 1998. 3
- [Weh93] WEHREND S.: Appendix B: Taxonomy of visualization goals. In *Visual cues: Practical data visualization* (1993), Keller P. R., Keller M. M., (Eds.), IEEE Computer Society Press. 1, 3
- [WL90] WEHREND S., LEWIS C.: A problem-oriented classification of visualization techniques. In *Conf. on Visualization* (1990), IEEE Computer Society Press, pp. 139–143. 3
- [YEL10] YI J. S., ELMQVIST N., LEE S.: TimeMatrix: Analyzing temporal social networks using interactive matrix-based visualizations. *Int. J. of HCI* 26 (2010), 1031–1051. 4
- [ZF98] ZHOU M. X., FEINER S. K.: Visual task characterization for automated visual discourse synthesis. *SIGCHI conference on Human factors in computing systems* 23, 18 (1998), 392–399. 1