

Choropleth map – which data type should I use?

Tomasz Nowacki ^{a, *}

^a POLAND, University of Warsaw, Faculty of Geography and Regional Studies, Chair of Geomatics and Information Systems, Department of Geoinformatics, Cartography and Remote Sensing, t.nowacki@uw.edu.pl

* Corresponding author

Keywords: Data types, Quantitative data, Absolute data, Relative data, Choropleth map, User study

Abstract:

The rules for choropleth map design have not been fully established. One of the controversial issues is the characteristics of the quantitative data which is accepted for the creation of choropleth maps. There are two ways of expressing data: absolute and relative.

Absolute data should be understood as pure data which is not related to any other data. This type of data is the result of measuring or counting. Relative data is data which is associated with other data and expressed as a fraction. The division of data into the two categories is widely accepted by cartographers, but the way in which they are interpreted varies. Determining what absolute data is, seems simple. On the other hand it is not so evident how to choose between the several different terms that are used to determine how relative data is described.

The dividing of the data types into more detailed parts at first leads to separation into four main types of data:

- absolute unprocessed – e.g. number of Toyota cars sold (Fig. 1a);
- absolute processed – e.g. percentage of Toyota cars sold (Fig. 1b);
- relative unprocessed – e.g. average points scored of the Spanish language exam (Fig. 1c);
- relative processed – e.g. average percentage of points scored in the Spanish language exam (Fig. 1d).

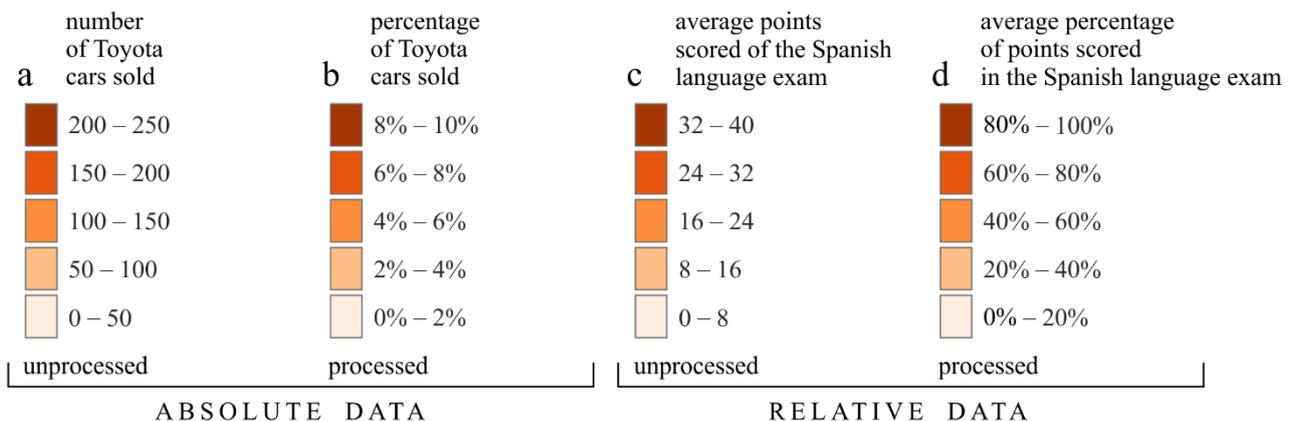


Figure 1. Main data types presented on choropleth maps legends.

Some data types cannot be classified either as relative or absolute. This approach leads to indirect data types, for example:

- data that can be considered to be absolute processed, and relative unprocessed (e.g. the number of cars per family in Poland, treated as the total number of cars related to the number of families or the average number of cars per family),
- data that can be classified as absolute processed, and relative processed (e.g. the increase in the average number of cars per family in Poland in the last decade).

The occurrence of a problem with an ambiguous data type classification leads to a situation in which different types of data are used to create choropleth maps. It is necessary to make an extension to the classification of data types, and indicate which of these categorically cannot be used for choropleth maps.

For choropleth maps, cartographers accept the use of different types of relative data, which can also be unrelated to area. At the same time, many cartographers are convinced that absolute data should be presented using a proportional symbol map. Visualising absolute data using choropleth maps seems to be inappropriate due to difficulties in interpreting the data

resulting from the different sizes of enumeration units. When reading a choropleth map based on absolute data, the size of the enumeration units should be omitted.

It seems that the perception of choropleth maps created for relative data, and which are not calculated taking into account the area of the enumeration units, should be just as difficult as the perception is of choropleth maps created for absolute data.

The opinion that absolute data cannot be used to elaborate choropleth maps is unacceptable. In this study, the relationship between the difficulty in using absolute data presented on choropleth maps and the profit resulting from graphical continuity, which is a great advantage of this form of cartographic presentation, was considered.

This study contains empirical research on the choropleth maps created for absolute and relative, unprocessed and processed data. The research was conducted in Poland in 2019. There were over 500 participants. They were asked to analyse maps and solve different types of tasks. The time and correctness of response were examined to look for differences between the maps created for different data types. The research verified whether relative and absolute data work better when are presented by choropleth maps or proportional symbols.

This work was supported by the Polish National Science Centre [Grant number UMO-2016/23/B/HS6/03846], "Evaluation of cartographic presentation methods in the context of map perception and effectiveness of visual transmission."