JOLANTA KORYCKA-SKORUPA, JACEK PASŁAWSKI
University of Warsaw, Faculty of Geography and Regional Studies
Department of Geoinformatics, Cartography and Remote Sensing
j.skorupa@uw.edu.pl; jpaslaws@uw.edu.pl

The beginnings of the choropleth presentation

Abstract. The discovery in the cartographic collections of the Faculty of Geography and Regional Studies at the University of Warsaw of an original map by Charles Dupin – the first choropleth map – provided an opportunity to conduct a closer methodological analysis of the map and to investigate the subsequent development of this presentation method during the first half of 19th century. From relatively early on, the accepted principle was for choropleth map presentations to use statistical data still imprecisely referred to as relative, as well as using a distribution series as a method of generalizing data.

Keywords: Charles Dupin, choropleth map, history of thematic cartography

1. Introduction

Among the quantitative forms of cartographic presentation, the choropleth map is undoubtedly the most frequently used. For many people, especially those untrained in cartography, it seems to be a simple and effective form, the use of which produces an evocative, often colourful map. It appears, however, that the conscious and correct development of the choropleth map on the one hand, and its correct reading by the user of the map on the other, are in fact not so easy.

The history of knowledge – and above all, the context of use – of the so-called cartographic quantitative forms of presentation have not yet been sufficiently explored, although some interesting studies have been conducted on these topics (A.H. Robinson 1982). Much more is known about the choropleth map, including its creator and early applications. It is therefore worth taking a closer look at the first choropleth maps from the point of view of the present-day knowledge of this form of presentation.

2. From history

The study of the history of thematic maps has long resided in the background of carto-graphers' interests. The main subject of cartographic historians' research has focused on the maps now known as 'general' geographic. Thematic maps of physical phenomena, such as geological structures, and the distribution of plants and animals, began to be developed in the second half of the 17th century, and quantitative approaches came even later, in the early 19th century (A.H. Robinson 1982, I. Kretschmer 1989). The origins of quantitative methods are quite well documented, because of the existence of scientific journals which even recorded the dates of certain events. We know that 30 November 1826 can be considered the 'birth' of the choropleth map (invented by Ch. Dupin) and 17 February 1845 as the 'birth' of isopleths (invented by L. Lalane). The authors of such innovations described them in detail – for example, A. Humbolt described the idea of isotherms, in the French journal "Annales de chimie et de physique" in 1817, with reference to the isogons on E. Halley's map of the Atlantic Ocean, complementing his deduction with a sketched map.

In the mainstream of cartographic knowledge, these events were recorded later, in the 20th century. It is necessary to mention here the extensive pre-war study of H.G. Funkhouser (1937), devoted not only to maps, but also to

the wider issues of the graphic presentation of statistical data. A.H. Robinson's (1982) book, dedicated to thematic maps, is captivatingly written and richly illustrated, and I. Kretschmer's (1989) study also deserves a mention. The studies of G. Palsky (1999, 2008) contain interesting information on the origins of quantitative methods in the French language area and

Fig. 1. Charles Dupin's map (1826)

the impact of these ideas on other Western European countries, especially Great Britain, as well as attempts to standardize the choropleth map.

3. The first choropleth map

There is no doubt that the author of the first choropleth map was Charles Dupin (1784−1872) – a civil engineer, senator and member of the French Academy of Sciences. Many biographies of Dupin are available on the Internet and in French encyclopaedias, although not because of his involvement in the first statistical map of its kind. Dupin's biographers have largely focused on his work in the fields of economics, statistics and mathematics, as well as his political activities, alluding to his contact with the mathematician G. Monge, with whom he studied descriptive geometry.

Ch. Dupin, as an engineer carrying out construction work in Mediterranean ports, had a good knowledge of maps, and his background in descriptive geometry probably also had an impact on the idea of elaboration a map of France to present statistical data in a vivid and visual way.

The choropleth map entitled *Carte figurative de l'instruction populaire* showed the schooling index: the number of people for each boy studying at school, related into 85 departments. The choropleth map was an illustration of Dupin's lecture, delivered on 30 November 1826, on the education of future workers as a factor influencing the prosperity of society and an indicator of economic development. Furthermore, it was an illustration of the situation in France. Interestingly, the analysis presented by the author during the lecture was published a year later, supplementing the map with the relevant statistical tables. The departments were marked using a value scale (shading) according to the principle 'the more, the darker' – in other words, the more people per student, the darker the department was shaded. According to the author, the map allows France to be divided into the 'enlightened' (light) departments under the influence of Great Britain, and 'dark' France, where the indicator is much higher, meaning a relatively small number of boys were sent to school. The straight line dividing these areas is connecting Geneva with the small Norman port of Saint-Malo. These are the only cities marked on the choropleth map (fig. 1).

It should be noted that Ch. Dupin used the map primarily as a research tool, the same function performed by later choropleth maps modelled on this map. He was convinced that educated workers were the prerequisite to a prosperous society, based on the example of Great Britain, where he had spent two years studying the British economy. The choropleth map has no legend, as will be discussed below.

4. Map analysis

The information contained on the map allows it to be analysed methodically. It contains two elements: the statistical data and the different shading steps of the 85 departments (tab. 1, columns 2 and 3), but there is no legend. We can therefore attempt to explain what rule the author used to visualize the statistical data.

The graphic array of the statistical data (fig. 2) constructed on the basis of the data shown on the map indicates a statistical skewed distribution. With the intention of explaining how to visualize the statistical data on the choropleth map, we measured the value (shading) of the individual departments' patterns (tab 1). The second column contains the statistical data, and the third shows the measured value of each department's surface. Columns 4 and 5
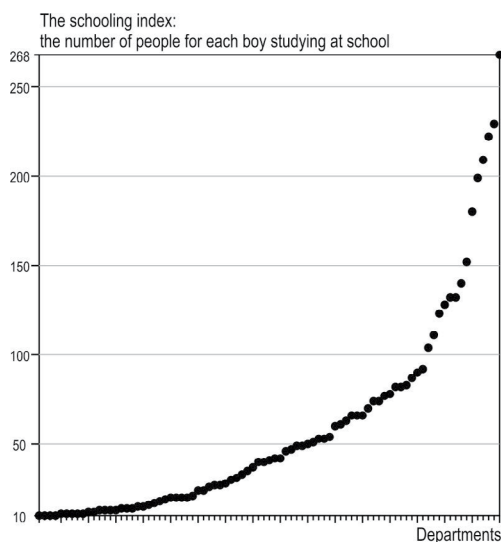

Fig. 2. Graphic array of the statistical data of the 85 departments on the map of France

Tab. 1. Statistical data, value of shading and normalized values for the departments on the map of France

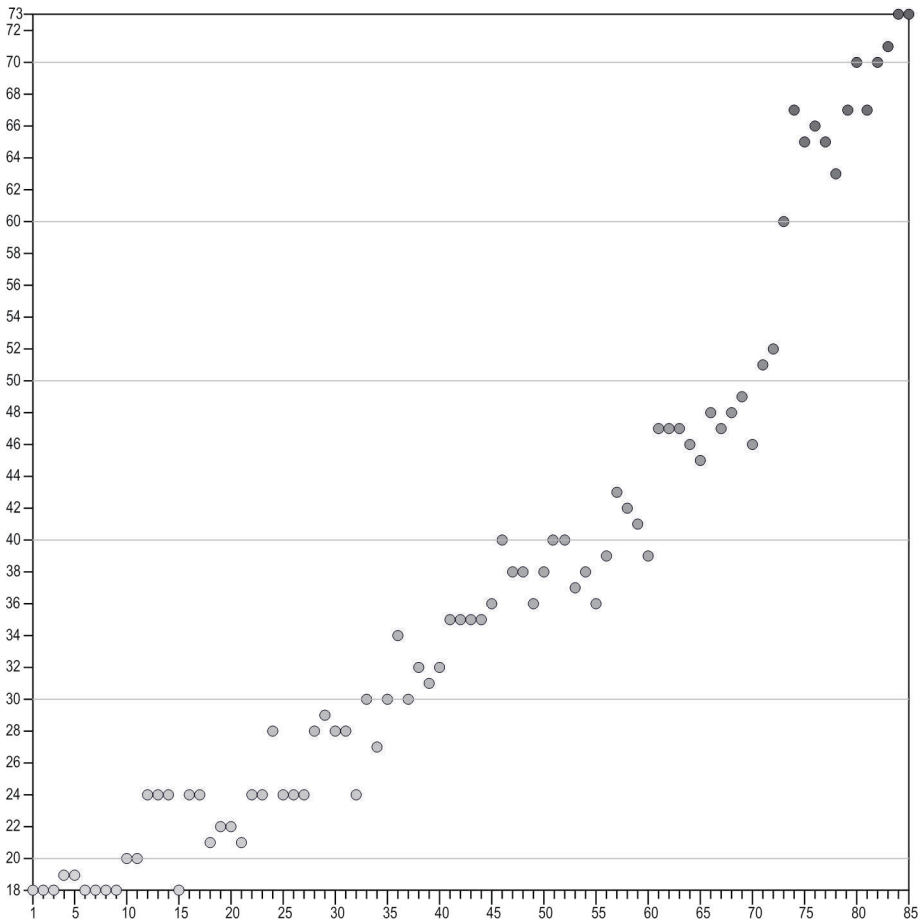| Departament	Stati-stical data	Value of shading	Normalization		Departament	Stati-stical data	Value of shading	Normalization	
			Stati-stical data	Value of shading				Stati-stical data	Value of shading
1	2	3	4	5	1	2	3	4	5
Aube	10	18	0,000	0,000	Orne	42	35	0,124	0,309
Cote –d Or	10	18	0,000	0,000	Var	42	36	0,124	0,327
Marne	10	18	0,000	0,000	Seine	46	40	0,140	0,400
Moselle	10	19	0,000	0,018	Gers	47	38	0,143	0,364
Bas-Rhin	11	19	0,004	0,018	Basses Alpes	49	38	0,151	0,364
Doubs	11	18	0,004	0,000	Bouches du Rhone	49	36	0,151	0,327
Haute Saone	11	18	0,004	0,000	Haute Garonne	50	38	0,155	0,364
Haute Marne	11	18	0,004	0,000	Ardéche	51	40	0,159	0,400
Oise	11	18	0,004	0,000	Charente	53	40	0,167	0,400
Jura	12	20	0,008	0,036	Vendée	53	37	0,167	0,345
Somme	12	20	0,008	0,036	Niévre	54	38	0,171	0,364
Haut Rhin	13	24	0,012	0,109	Sarthe	60	36	0,194	0,327
Ardennes	13	24	0,012	0,109	Lot	61	39	0,198	0,382
Aisne	13	24	0,012	0,109	Gironde	63	43	0,205	0,455
Seine et Marne	13	18	0,012	0,000	Loire	66	42	0,217	0,436
Meurthe	14	24	0,016	0,109	Pyrénées Orientales	66	41	0,217	0,418
Meuse	14	24	0,016	0,109	Tarn et Garonne	66	39	0,217	0,382
Pas-de-Calais	14	21	0,016	0,055	Vienne	70	47	0,233	0,527
Basses Pyrénées	15	22	0,019	0,073	Creuse	74	47	0,248	0,527
Yonne	15	22	0,019	0,073	Indre	74	47	0,248	0,527
Hautes Pyrénées	16	21	0,023	0,055	Aveyron	77	46	0,260	0,509
Eure et Loir	17	24	0,027	0,109	Mayenne	78	45	0,264	0,491
Vosges	18	24	0,031	0,109	Cher	82	48	0,279	0,545
Seine et Oise	19	28	0,035	0,182	Tarn	82	47	0,279	0,527
Drome	20	24	0,039	0,109	Loiret	83	48	0,283	0,545
Hautes Alpes	20	24	0,039	0,109	Haute Vienne	87	49	0,298	0,564
Isére	20	24	0,039	0,109	Maine et Loire	90	46	0,310	0,509
Nord	20	28	0,039	0,182	Lozére	92	51	0,318	0,600
Gard	21	29	0,043	0,200	Dordogne	104	52	0,364	0,618
Eure	24	28	0,054	0,182	Ille et Vilaine	111	60	0,391	0,764
Seine Inférieure	24	28	0,054	0,182	Arriege	123	67	0,438	0,891
Landes	26	24	0,062	0,109	Corréze	128	65	0,457	0,855
Calvados	27	30	0,066	0,218	Loir et Cher	132	66	0,473	0,873
Vaucluse	27	27	0,066	0,164	Loire Inférieure	132	65	0,473	0,855
Deux Sévres	28	30	0,070	0,218	Allier	140	63	0,504	0,818
Manche	30	34	0,078	0,291	Cotes du Nord	152	67	0,550	0,891
Hérault	31	30	0,081	0,218	Puy de Dome	180	70	0,659	0,945
Charente Inférieure	33	32	0,089	0,255	Finistere	199	67	0,733	0,891
Saone et Loire	35	31	0,097	0,236	Cantal	209	70	0,771	0,945
Ain	37	32	0,105	0,255	Morbihan	222	71	0,822	0,964
Lot et Garonne	40	35	0,116	0,309	Indre et Loire	229	73	0,849	1,000
Rhone	40	35	0,116	0,309	Haute Loire	268	73	1,000	1,000
Aude	41	35	0,120	0,309					

Fig. 3. Graph showing the relationship between the value (shading) of departments' areas (vertical axis) and the order of statistical data. The horizontal axis shows the departments according to the order of the data, not maintaining the relationship of differences between the statistical data

contain normalized data for the statistical data and the shading of the surfaces, converted into a range of 0–100%.

The graph created to illustrate the shading of the departments (vertical scale), arranged according to the order of data, gives an interesting picture (fig. 3). Admittedly, there is a clear relationship between the increase in statistical data and the shading, but distinct 'irregularities' can also be seen. At the beginning of the graph, departments 6-9 have the lowest shading, despite their statistical data being slightly higher than the first three. In many cases, different statistical data were assigned the same levels of shading (for example, departments 41-44), and vice versa – different degrees of shading correspond to the same statistical data (for example, departments 27 and 28). There are also examples of complete compatibility between the statistical data and shading (for example, departments 12-14).

Discrepancies between the variability of the statistical values and the variability of the shading are also well illustrated in figure 4. The mentioned normalized data of these features are marked on both axes. It is not difficult to imagine how this graph would look if the variability of value scale and the statistical data
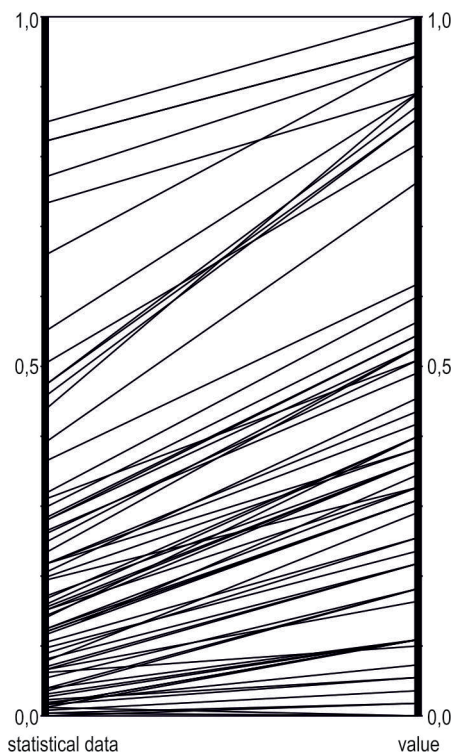
Fig. 4. The axes show the statistical data of the 85 departments and their corresponding value (shading) on the map. Both data strings were normalized in the range 0−1

Tab. 2. Class selection according to geometric progress

Class limits	Class intervals	Number of departments
10–18	8	22
19–30	12	14
31–51	20	15
52–90	38	19
91–155	64	9
156–268	112	6

ments marked with the same level of shading. We can therefore confirm with certainty that, according to the present state of knowledge, Dupin's choropleth map was intended by the author to be a continuous (unclassed) choropleth map.

In possession of the statistical data and a map with the administrative divisions of France from that time, it is hard to resist the desire to see what image will be obtained using our modern--day knowledge and technical capabilities. Figure 7 comprises three choropleth maps, created according to the rules explained in the title. As mentioned, a skewed distribution (fig. 6) usually prompts the cartographer to develop classes with intervals changing according to one of the progressive rules, which would result in similar quantities. The general idea of using progressive methods is to obtain classes with variable, mathematically defined intervals, 'striving' to maintain similar quantities of space units each class. In the case of this map, class designation according to geometric progress can be used (G.F. Jenks, M.R.C. Coulson 1963). These classes would have the following limits: 10–18, 19–30, 31–51, 52–90, 91–155, 156–268 (tab. 2). These seemed to be the most suitable classes due to the fact that the graph's line of the schooling index is closest to the curve of geometric progress (fig. 6). The choropleth map developed on the basis of these classes (fig. 7B) seems quite similar to Dupin's map. The 6-class choropleth map depicts more clearly a compact block of the departments with relatively low values, defined by the author as 'enlightened' France.

An attempt could also be made to determine classes based on 'natural' breaks in the graph

were completely consistent – the straight lines connecting the two axes would be more or less parallel and should not intersect.

Taking note of the above, we should bear in mind the technical capabilities of an engraver in the 1830s. However, the presented graph precludes the idea that the author and engraver of the map were aiming to generalize the data in the form of classes. Perhaps the author expected the engraver to differentiate the shading according to the values, but it should be stressed that, had that been the case, 57 steps of shading would have been required to correspond to the different statistical data on the map! According to the method of measurement used, the map has 37 steps of shading, as shown in figure 5. The dots representing the departments are arranged here according to order of shading. The dots positioned next to each other in parallel represent the depart-
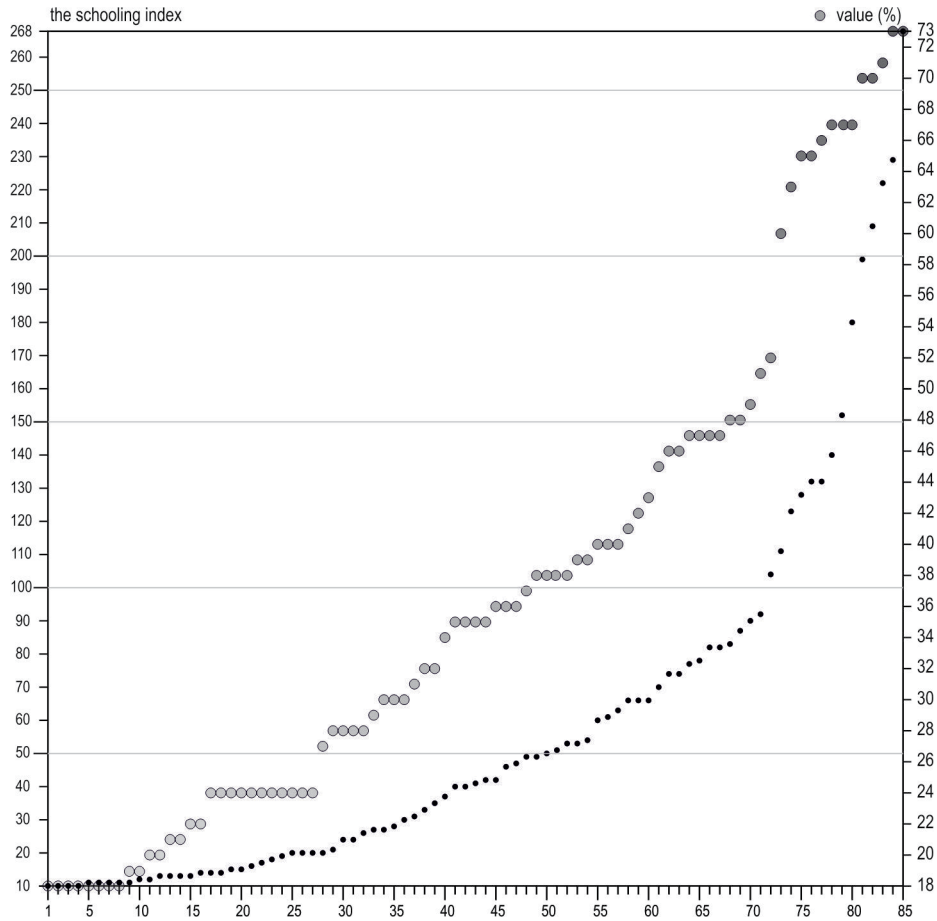
Fig. 5. Statistical data and value (shading). The engraver's 37 steps of shading are visible

– for example: 10–22, 23–45, 46–55, 56–100, 101–160, 161–268 (tab. 3). Based on this division, the first class would contain 25% of the departments, and the second and fourth would contain 15% (fig. 7). Differences between the maps are visible only in the lowest values (in the first class) and the medium values (in the third and fourth classes).

A comparison of the three choropleth maps (fig. 7) highlights the fact that both the author, Ch. Dupin, and his engraver were able to visualize statistical data very well. The third choropleth map (natural breaks) appears to present 'enlightened' France most clearly. These choropleth maps can be treated as a contribution to the discussion on methods of designating class. It is therefore important to focus on how Dupin's followers coped with this methodical problem; there is no doubt that the authors of

Tab. 3. Class selection using the natural breaks method

Class limits	Class intervals	Number of departments
10–22	12	29
23–45	22	16
46–55	9	9
56–100	44	17
101–160	59	8
161–268	107	6

J - curve of linear progress
A - curve of arythmetic progress
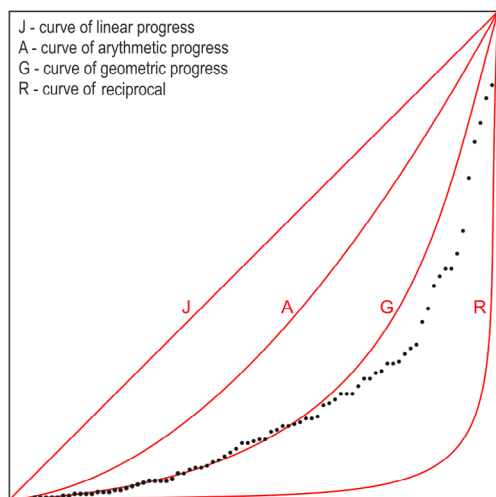G - curve of geometric progress
R - curve of reciprocal

Fig. 6. Curves used to determine the class

choropleth maps from the French circle were merely followers, as can easily be demonstrated by carefully reviewing the reproductions of the maps in A.H. Robinson's (1982) book.

5. Method development

Other choropleth maps quickly appeared which were modelled on Dupin's map. Most likely in 1827, H. Sommerhausen, a bookseller from Brussels, published a map with the same title as Dupin's map, but including Pays-Bas – today's Belgium, the Netherlands and Luxembourg. The map was reproduced by A.H. Robinson (1982, p. 63). Similarly to Dupin's first choropleth map, this map also lacks a legend for the thematic scale.

The choropleth map number 13 in *Administrativ-statistischer Atlas vom Preussischen Staate* (consisting of 22 maps) is of great interest. The first edition of the atlas, marked 1827/1828, appeared at the time of the publication of Dupin's map, but the both choropleth maps are completely different in character.

In the *Atlas of Prussia*, on a map at the scale of approximately 1:2,000,000, over 330 counties (Kreis) were marked and each was assigned one of 17 colours according to the scale placed under the map. The theme of the map is population density, which is shown in this way on a map for the first time. The map takes the form of a hand-coloured copperplate. A reprint

was published in 1990, which is the source of the map section shown here (fig. 8). In a commentary on the reprint, W. Scharfe (1990) quotes a review of the atlas written by H. Berghaus in which the reviewer drew attention to this map and the number of distinguished classes. It should be added that, in the second edition of the atlas from 1845, the number of classes on the map of population density was reduced to 13.

The author of the map is unknown. A.H. Robinson (1982, p. 113) suggests that it could be C. von Rau, a topographer and major of the Prussian army, while I. Kretschmer (1989, p. 18) cites the opinion that the map may have been inspired by C. Ritter, the well-known German geographer and author of a set of thematic maps entitled *Sechs Karten von Europa* (1806). The topic of one of these maps is population density, which is described by figures alone. The intention of the author of this map was to create a colourful scale with volume varying from light (low population density) to dark. The cities with the highest population density were marked in black. The 'contemporary' structure of the colour scale draws the attention, and the description reveals that the author tended to use classes with similar intervals and "round numbers".

A reprint of the atlas was published with a commentary for each map. The map entitled 'Ratio of population to area' (*Verhältnis der Bevölkerung zur Grundfläche*) was discussed in the commentary by W. Scharfe (1990, pp. 133−142), who included a table of data on population density for 1825.

Returning to the authors from the French circle, the maps developed by the geographer and statistician A. Balbi and lawyer A.M Guerry constituted an extension of the choropleth map's subject. In 1829, they published three maps of France on a single plate, presenting schooling as on Dupin's map) in addition to the spatial diversity of thefts and murders. This was a subject related to the research developed later on in the field of 'moral statistics', the aim of which was to understand and combat the social pathologies that were rampant at the time. The subject of this map resulted from the interests of both authors. A.M. Guerry is the author of the 1833 book entitled *Essai sur statistique morale de la France.* The term 'moral statistics' used in the title became quite popular and the further use of choropleth maps was related to this subject.
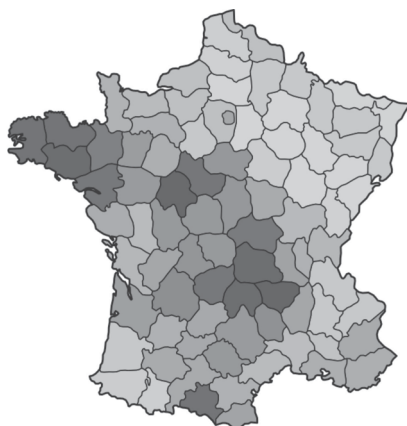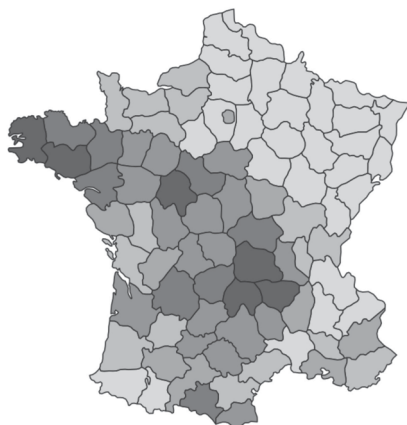
A. choropleth map according to Dupin

B. 6-class choropleth map developed according to geometric progress

C. 6-class choropleth map developed using the natural breaks method

Fig. 7. Different class selection based on Ch. Dupin's map

Fig. 8. A map section from *Administrativ-statistischer Atlas vom Preussischen Staate*

Also of interest are the maps of A. d'Angeville. He developed 16 maps to illustrate his extensive paper on the French population, which was published in 1836. The maps illustrated topics related to the social and demographic situation, as well as the theme of 'moral statistics'. They can be treated as the first, very successful, compilation that allowed the comparison of maps on various topics. The first map presents population density, while the next one shows various social themes. The data for map no. 13, entitled 'Crimes', relates to the number of

people per accused criminal in each department. The lower the number (the greater the occurrence of crime), the darker the pattern the department is assigned. The influence of Dupin's first choropleth map, which used an index with a similar structure, is evident here. Map no. 10, showing the living conditions of the French population, is similarly developed, based on an index of the number of doors and windows per 100 inhabitants. A.H. Robinson (1982, p. 183) probably reproduced this map because it also clearly presents Dupin's 'enlightened' and 'dark' France.

A. d'Angeville's maps were elaborated in an intriguing way. Although they have no legend, and therefore no thematic scale, they contain a tabular data set. In each map, the departments are arranged from the lowest value to the highest and divided into five classes consisting of of 17 departments each. The individual departments are numbered on the maps, not with statistical data like Ch. Dupin's map, but with the successive numbers provided in the table. As mentioned, there are 16 of these edited maps illustrating the variability of phenomena – generalized, according to one rule (equal quantities) into five classes – and beyond that, allowing the reader to identify specific data for the department of interest.

Based on the review of the early choropleth maps cited by the authors of the above-mentioned historical studies and the reproductions included in A.H. Robinson's (1982) book, we observe that just ten years after the publication of Ch. Dupin's concept, the basic idea of the choropleth map was known as a form of visualizing and generalizing statistical data.

6. Statistical congresses

According to H.G. Funkhouser (1937), the period 1830−1850 saw two decades of dynamic development in the field of statistics. Statistical offices and societies were established, and specialist statistical journals were launched. At the same time, the need for international cooperation in data collection and processing was recognised.

On the initiative of A. Quetelet (1796−1874), the first international statistical congress took place in Brussels in 1853. The issues of graphic presentation were discussed during the 3rd Congress in Vienna in 1857 and at the 8th Congress in St. Petersburg in 1872. Thanks to the publication of congress materials, the proceedings and the ongoing discussions reported in the quoted literature (H.G. Funkhouser 1937, A.H. Robinson 1982, I. Kretschmer 1989, G. Palsky 1999) are quite well known.

The dominating problem, comparable to the case of collecting statistical data, was the issue of the standardization of graphic forms. Opinions on this subject ranged greatly, especially concerning specific issues such as methods of class selection, their number, and colour scales. Congressional meetings were a good opportunity to present various ideas. For example, the proponents of 'natural' division into classes debated with the supporters of equal quantities, and another group of statisticians were in favour of determining the arithmetic mean (originating from A. Quetelet) and the class intervals corresponding to the standard deviation of the set. In this case, the colour scale would consist of shades of red (above average) and blue (below average). Today, this type of thematic scale is called 'divergent'. The advantage of this kind of class selection would be the facilitation of comparisons between the contents of choropleth maps, since the same colour on different maps would indicate the similar position of a given areal unit in the overall statistical set. This principle had previously been used by the American J. Fletcher in a publication from 1849 (H.G. Funkhouser 1937).

The wording of the recommendations prepared by the section chaired by A. Quetelet during the congress in Vienna in 1857 are interesting. The name of the presentation form was not used in the text, although they are clearly choropleth maps. The authors recommend using no more than 12 classes, or fewer, depending on the requirements, and providing a legend next to the map; this should contain statistical data with reference areal units which should be "The absolute numbers which have served as bases for the proportional numbers" that was used to (according to H.G. Funkhouser 1937, p. 315).

It follows that the use of relative data for a choropleth map presentation was adopted from the earliest use of the choropleth map method. Based on the materials of the Vienna Congress of 1857, it can be concluded that the choropleth map was treated at that time as a basic map illustrating statistical data, and that this data was relative data.

A later compilation by G. Mayr, prepared for the congress in St. Petersburg (1872) and published in the congress materials, proposing a classification of the graphic forms useful for the presentation of statistical data, distinguishes choropleth maps illustrating 'statistical averages' among the forms used on maps. Over time, the conviction of the choropleth map presentation of relative data alone became established (H.G. Funkhouser, p. 335).

In the end, none of the congresses adopted a resolution on the standardization of elaboration choropleth maps or other graphic methods. The congressional meetings were an opportunity for discussion and a stimulus for publications which can be described nowadays as methodical.

Following the origins of the dissemination of the choropleth map method, it can be observed that the authors used this graphic form to analyse the presented phenomena. Dupin's choropleth map was an illustration of a thesis on the diversity of France based on the level of education of workers. A similar explanatory function was served by the choropleth maps in A. d'Angeville's paper. Cartographic presentation was of interest to doctors. Choropleth maps were still used in the first half of the 19th century to identify relationships between the occurrence of diseases (such as hernias) and other phenomena, such as selected crops.

7. Final remarks

In the early years of the development and use of choropleth maps, the conviction about using relative values and the generalization of data using classes was already fixed. Initially, the principle of "equal quantities" was applied in the French language circle, but the choropleth map showing the population density of Prussia, which was developed around the same time, used the principle of "round numbers" of class limits. Despite the numerous discussions that took place at the statistical congresses held in the second half of the 19th century, the properties of the various statistical methods for determining classes were not adequately clarified, much less the rules of standardization accepted.

Literature

Funkhouser H.G. 1937, *Historical development of the graphical presentation of statistical data*. "Osiris" Vol. 3, pp. 269−404.

Jenks G.F., Coulson M.R.C. 1963, *Class intervals for statistical maps*. "International Yearbook of Cartography" Vol. 3, pp. 119−134.

Kretschmer I., 1989, *Die Entwicklung der Methodenlehre der thematischen Kartographie bis in die 1960er Jahre*. "Berichte und Informationen" Nr 12. Österreichische Akademie der Wissenschaften Institut für Kartographie, pp. 1−55.

Palsky G., 1999, *The debate on the standardization of statistical maps and diagrams (1857−1901). Elements of the history of graphical language*. "Cybergeo. European Journal of Geography" (source: http://journals.openedition.org/cybergeo/148).

Palsky G., 2008, *Connection and exchange in European thematic cartography. The case of 19-th century choropleth maps.* "Belgeo", No. 3/4, pp. 413−426.

Robinson A.H., 1971, *Genealogy of isopleth*. "The Cartographic Journal" Vol. 8, No. 1, pp. 49−53.

Robinson A.H., 1982, *Early thematic mapping in the history of cartography*. Chicago, London: The University Chicago Press.

Scharfe W., 1990, *Administrativ-statistischer Atlas vom Preussischen Staate. Erläuterungsband*. Publikationen Historischen Kommission zu Berlin, Reihe "Kartenwerk zur Preussischen Geschichte". „*Administrativ-statistischer Atlas Preussischen Staate", Werk, Verfasser und Urheber,* pp. 6−18. *Verhältnis der Bevölkerung zur Grundfläche,* pp. 133−142. Berlin: Kiepert Verlag.