Pardubice Region Health Data Visualization – Steps of Geoinformatics Project

Ing. Ivana Čermáková

1st author's affiliation
University of Pardubice
Studentská 95
530 09, Pardubice, Czech
Republic

st32689@student.upce.cz

Mgr. Pavel Sedlák, Ph.D.

2nd author's affiliation
University of Pardubice
Studentská 95
530 09, Pardubice, Czech
Republic

pavel.sedlak@upce.cz

doc. Ing. Jitka Komárková, Ph.D.

3rd author's affiliation
University of Pardubice
Studentská 95
530 09, Pardubice, Czech
Republic

jitka.komarkova@upce.cz

ABSTRACT
Increasing importance about health and healthcare leads to utilization of unusual methods to present information to residents. Thematic maps, particularly set of thematic maps, which are readable and understandable is one of possibilities how to inform residents. Creating a set of thematic maps is a whole project of data visualization, because it is necessary to cover all facts and to be able to clearly present facts that are important for readers. The project of data visualization consists of the analytic phase, design phase, testing phase and the final implementation phase. Final set of maps contains 19 thematic maps and using 4 cartographic methods, e.g. directional distribution. The article describes case study: the whole project of creation of the thematic maps of Pardubice Region of the Czech Republic regarding health and healthcare. The article is focused on specific tasks solved within this project.

Keywords
Data visualization, health data, geoinformatics project, project of visualization

1. INTRODUCTION
Health and healthcare belong to the important topics about which public want to be inform. Many institutions are searching for suitable ways how to provide data in a format that will be understandable for citizens. Importance of thematic maps increases in the case of this problematic. One of disadvantages for readers is that maps are usually available for the whole world or country, but they are not available at regional level.

The main aim of the paper is to find a suitable way how to provide information about important topics regarding health and healthcare at regional level, namely on region Pardubice in the Czech Republic. Various cartographic methods are used to present information and create the final set of maps.

The paper describes a case study: creation of a set of thematic maps describing health and healthcare for the Pardubice Region is a complete geoinformatics project. It covers the whole process starting with users and aim definition and available data collection. It ends with published set of thematic maps. The project of creation of the set of maps is divided into 4 phases: analysis, design, testing and implementation. Furthermore, maps can be improved according to obtained results. The final set of thematic maps is created in the last phase (implementation).

2. PHASES OF THE PROJECT
The whole project is divided into 4 phases. Content of particular phases is based on previous geoinformatics projects ([Dob00a] and [Haj00a]), but each phase is modified according to the needs of this project, see Figure 1.
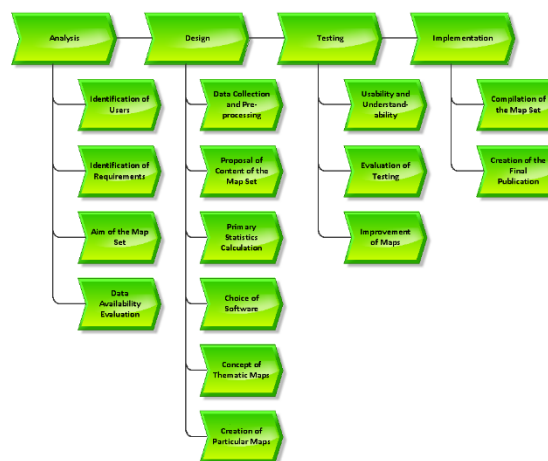

Figure 1: Phases of the Project, source: authors

Phase 1: Analysis
The analytical phase consists of definition of target group of users and their requirements and expectations, definition of aim of the set of maps and evaluation of availability of health data.

The aim of the set of maps is to describe situation of healthcare in the Pardubice Region by thematic maps to citizens. Different people (i.e. readers of the maps) have different level of knowledge so they differently understand these problematics. This fact leads to the necessity of maps readable and understandable by everyone.

Topicality of maps requires adequate and up-to-date datasets. Available and accessible datasets must be identified and evaluated. In the case the following data sources are used: ÚZIS, Regional Public Health Authority of Pardubice Region and NADĚJE. ÚZIS is the key institute providing health data. The institute is an organisational component of the Czech Republic, founded by Ministry of Health. [Uzi00a] Regional Public Health Authority of Pardubice Region is institute which provides health data only for its region of interest. NADĚJE is a non-profit non-governmental organization that provides religious, social, medical, counselling, educational, enlightenment and pedagogical services etc. [Nad00a].

Phase 2: Design

At first, data are collected from the above listed sources. They are evaluated from the point of view of their type, e.g. ordinal, cardinal, etc. This step is necessary for choice of proper statistical, analytical and cartographic methods in next steps.

Next, detailed list of included topics must be proposed according to the aim of the set of maps and according to the available data. In our case, there are chosen 19 topics which are divided into 3 basic parts. First part is focused on general characteristics and it contains 3 topics. The main part is focused on the state of health and healthcare (12 topics desrcribed by 16 maps). The last part of the set of maps is focused on health from other points of view. This part covers 4 particular topics in total: revenues and costs of hospitals, employees in health care and Public Health Control System.

Data pre-processing is the next step. Collected data must be verified and non-complete and non-reliable data must be discarded. Data from primary sources are in different formats so data must be transformed into format suitable for the next steps.

Statistical evaluation is the next step. All the data has one type – ordinal type. It means that features are measures by quantity [Rob00a]. Basic summary statistics, e.g. average, minimum, maximum and other measures, is calculated. These measures are necessary for choosing a suitable method for visualization and dividing data into the intervals.

ArcGIS for Desktop version 10.2.2 is used to create maps because it is one of mostly used software for data analyses and consequent maps making.

CorelDRAW (version X5) is used for design of the resulting set of maps.

Concept of thematic maps is following: maps contain topographical background and thematic content. Topographic dataset ArcČR 500 is used as a topographic background. The dataset contains streets, rivers, lakes, forests or administrative boundaries [Top00a].

Design and creation of a template of maps follows. Template must be suitable for the problematic and must be attractive and interesting for map users.

Main advantage of creating thematic maps is that the thematic maps are published as a whole set. A comprehensive set of suitable symbols is proposed for the set. The symbols are unique and increase importance and understandability of maps because they are proposed in an associative way.

For the whole making the set of the maps it is necessary to describe the general characteristics of health and healthcare in the Pardubice Region. It means create maps of population, age index and health centres. Of course, mostly there are thematic maps use traditional cartography methods, for example diagram map and choropleth maps. Of course there are some special maps use more advanced methods, e.g. dots method or spatial statistics (every type of visualization is explained in the following parts).

2.1.1 Directional distribution method

This method is suitable in cases when concentration of features is higher in particular directions [Arc00a]. This method is used for one map only. This map compares births and deaths. Figure 2 shows difference between ellipse of deaths and births. The map clearly shows that both easternmost regions (Moravskoslezský a Jihomoravský) have higher number of deaths than births. It is shown by prolongation of the ellipse to the east.
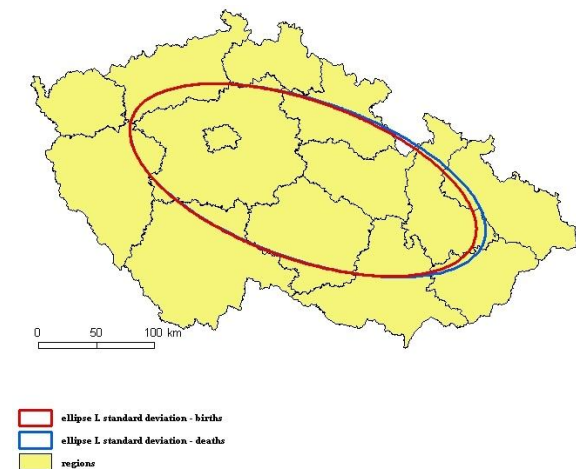

Figure 2: Comparison of deaths and births in 2012, source: [Cer00a]

2.1.2 Dots method

Dots method of visualization allows highlighting number of features better than single features. [Rob00a] For example, Figure 3 shows occurrence of tuberculosis in Pardubice Region by dots method (1 dot =1 occurrence).
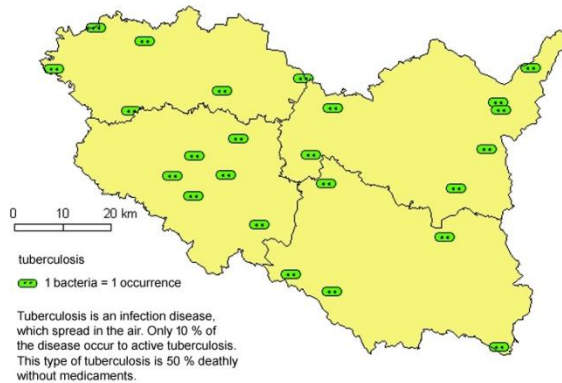

Figure 3: Tuberculosis in the Pardubice Region in 2012, source: [Cer00a]

Dots method is used for map of incidence of chickenpox also. Red dot is used for representation of number of people with chickenpox. Red colour is used because it is visible character of this disease. One red dot represents 90 patients. Icon of green bacteria is used in the case of map of incidence of tuberculosis instead of dots. Green bacteria symbol is a typical symbol for this disease. Dots method is used in the map describing state medical and map of epidemiological investigations.

2.1.3 Diagram map method

This method is suitable for some maps and some topics only. Diagram map method is used for example for thematic map regarding hospitalized persons. In this case, diagram uses enlarging icon of white bed with a man on an orange background and size of the symbol represents number of hospitalized persons. Different sizes represent different numeric intervals. Choice of a suitable number of intervals is important to adequately represents differences between regions on one side and to keep readability on the other side. Statistical analysis helps to set suitable ranges of intervals. Three intervals are chosen for each thematic map describing situation at the national level. Another problem is how data divide into intervals. It is individual for every map. Natural breaks (jenks) are used for most of maps. More about each map and method used for dividing data into intervals is available in [Cer00a].

Diagram maps method is used in the case of 12 maps. One is described above.

2.1.4 Choropleth mapping method

Maps are used for phenomena with spatial variation which can be divided by boundaries to the intervals. [Ter00a] Map of women in working age with hormonal contraceptives is divided to two ranges: < 50 % and > 50 %. Saturated colour (shades of pink in the case of this map) has values with higher intensity. Map regarding hospitalized persons at the regional level used for each particular value one colour (as a whole 4 colours). Map regarding births, described in previous section, uses choropleth mapping method too. Map at the national level uses 3 intervals (yellow, blue, pink) and map at the regional level uses only 2 intervals (blue and pink) – one colour for 2 values. Map is available like Figure 4.
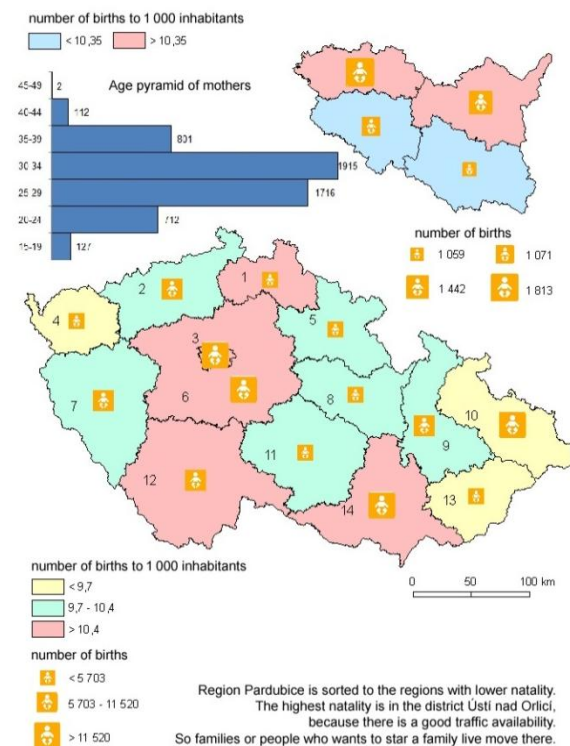

Figure 4: Births in the Czech Republic in 2012, source: [Cer00a]

Map about abortion at the national level uses 3 intervals symbolized by shades of blue. The last map describes employees working in health care. There are used 4 colours in this map – one colour is used for each particular value, i.e. each district. Map is designed in shades of green.

Phase 3: Testing

After creating all maps, it is necessary to find out if the maps are usable and if the readers understand content of the maps.

Several people, representatives of both public and specialists, are chosen for testing thematic maps. The main objective of testing phase is brief user testing of usability and understandability of thematic maps.

Representatives of both groups tested understandability of thematic maps and design of the maps also. All comments and identified problems were evaluated and used to improve maps.

Phase 4: Implementation

Compilation of the final set of thematic maps is the next step. As it was described previously, the set is divided into 3 parts and it contains graph (age pyramid) and 4 tables with important information about monitored area as well. Additional text must be included to describe main aim of the publication, content of the set of maps and target group of readers. The whole set of maps is available in [Cer00a].

3. CONCLUSION

The main aim of the paper is to identify suitable ways how to visualize health and healthcare data at the regional level in a comprehensive set of maps. Proposed methods are applied to the Pardubice Region in a form of a case study.

Creation of the comprehensive set of maps is a whole project. It is necessary to identify target group of users, their requirements, to collect data, to appropriately process and visualize them. The project is divided to the following parts: analytic phase, design phase, testing phase and final implementation phase.

The comprehensive set of thematic maps contains 19 thematic maps, 4 tables and 1 graph. It represents the main result of the case study. Thematic maps are used as a suitable way for visualization of particular topics and their distribution over the area of interest (Pardubice Region). Four cartographic methods for visualization are used to create thematic maps: directional distribution, dots method, diagram map and choropleth mapping. A new set of associative map symbols is proposed for the final set of maps to present health and healthcare data in an understandable way.

The main contribution of the paper is show a possibility how visualize the health data on the regional level, because no many researches are focused on. And it is very important topic in these days.

The output set of maps can be extended in the future by using other methods of visualization, by including new important topics or by combination of both possibilities. Creating other sets of thematic maps on the regional level can compare health and healthcare from various viewpoints also.

5. REFERENCES

[Arc00a] ArcREVUE: 2/2014 [online]. Praha: ARCDATA Praha s.r.o, 2014. ISSN 1211-2135. Available: http://www.arcdata.cz/Publikace/casopis-arcrevue/ Access Date: 9 September, 2015

[Cer00a] Čermáková, I. Visualization of selected health data from Pardubice region. Pardubice, 2015. Diploma thesis. Univerzita Pardubice. Available in Czech only, title: Vizualizace vybraných dat ze zdravotnictví Pardubického kraje

[Dob00a] Dobesova, Z. Geographic information systems for botanical garden - Steps of design and realization in Conf. proc. SGEM, pp.377-384. 2012

[Haj00a] Hajkova, L. et. al. GIS and cartography for bioclimatological research: Atlas of phenological conditions of Czechia in Conf.proc. SGEM, pp.1065-1072, 2012

[Nad00a] Naděje. [online]. 2014. Available: http://www.nadeje.cz/vyrocni_zpravy Access Date: 19 September, 2014

[Rob00a] Robinson, Arthur Howard. Elements of cartography. 6th ed. New York: John Wiley & Sons, 1995, 674 s., [8] s. obr. příl. ISBN 0471555797.

[Ter00a] Terry, A. Slocum. Thematic cartography and geovisualization. 3th ed. United States of America: Pearson Prentice Hall, 2010, 561 s., ISBN 0-13-801006-4

[Top00a] Topografická mapa. Atlas map České republiky - GEOČR500 [online]. 1998. Available: http://www.geology.cz/demo/cd_geocr500/stranky/p_topograficka.html Access Date: 30 November, 2014

[Uzi00a] Ústav zdravotnických informací a statistiky ČR [online]. 2014. Available: http://www.uzis.cz/nas Access Date: 19 September, 2014