Taylor & Francis
Taylor & Francis Group

DATABASE MODELLING IN CARTOGRAPHY FOR THE "ATLAS OF ELECTION"

Zdena Dobesova

Department of Geoinformatics, Faculty of Science, Palacký University,
Tr. Svobody 26, 771 46 Olomouc, Czech Republic
E-mail: zdena.dobesova@upol.cz

Abstract. Today maps are prepared in Geographic Information Systems (GIS software) and based on data stored in a database. In the stage of the conceptual database design, the graphic editor of a database model is recommended. The structure of data is often under the influence of the cartographic requests. For example, new data may be added only for visualization purposes. All database structures for a base data and a cartographical data can be defined in a conceptual database model before creation of a physical database model. Database modelling is demonstrated in ArcGIS Diagrammer software in this article. Design of a cartographic database model for the book "Atlas of Election to the Olomouc Region Council" is used as an example. Moreover, steps of a model creation, detail structure and relationships in the model are also mentioned. The cartographical database model of the "Atlas of Election" illustrates cartographical influence to the database structure.

Keywords: atlas, cartography, spatial data, election, GIS, spatial model.

1. Introduction

Geographic information systems are very often used for the creation of cartographic outputs (maps). Reality is represented in GIS as a system of thematic overlapping layers (Longley *et al.* 2001). In ArcGIS software (producer by the ESRI company), the layers are organized as feature classes in a geodatabase. Feature classes are collections of points, lines, polygons, or map texts managed as feature tables. A geodatabase is a type of database for storing spatial data. Every feature class has geometrical and lexical attributes. The geodatabase can also contain tables and other structures (relationship feature classes, value valid tables, raster data sets, etc.).

The design of a geodatabase model should reflect openness and usability for many similar cases. One purpose can be the creation of the same type of map only for various areas (district units) or for different time periods. The design of a geodatabase is started by creating a definition for a set of themes (layers). Subsequently, a detailed definition of each layer is required (spatial representation – point, line, polygon shape; data types of attributes, etc.). The database model can be shared by users for the same type of cartographic project. The difference may be only in values of data or the geographic location (other district unit). The structure of the database is the same for various maps. The database model for the book "Atlas of election to the Olomouc Region Council" (Vozenilek *et al.* 2009) is an illustrative case.

2. Database modelling

Conceptual database modelling is one of the starting phases of the design. Process of all designs can be divided into eleven steps in GIS (ESRI 2002).

2.1. Steps of GIS database design

1. Identification of the key thematic layers.
2. Specification of the scale ranges for data use and specification of the spatial representations for each scale range.
3. Decomposition of each representation into data sets: feature classes, tables, raster datasets, etc.
4. Identification of the attribute fields for each class
5. Specification of any valid values and the relationships between classes and tables.
6. Identification of subtypes for the feature classes.
7. Definition of spatial relationships, integrity rules, and behaviour (i.e., topologies and geometrical networks).
8. Proposition of a geodatabase design.
9. Implementation, reviewing, and refining of the first design.
10. Design of work flows for building and maintaining each layer.
11. Documentation of the design using appropriate methods (e.g., database schema, UML diagrams).

The first seven steps are the same for any GIS. The last four steps depend on the physical implementation.

A graphic modelling editor is very useful in the stage of conceptual and logical design (Arctur, Zeidler 2004).

2.2. Software for database modeling

Universal programs exist for database modelling. They are commonly named CASE Tools (Computer-Aided Software Engineering). In the case of ArcGIS software, three CASE tools can be considered. The first recommended software is Microsoft Visio, the second is Rational Rose (Rational Software Architect, Rational Software Modeller) and the third is ArcGIS Diagrammer (ESRI 2011a, b). Microsoft Visio and Rational Rose software serve for modelling of a non-spatial database. ArcGIS Diagrammer software is limited only to ArcGIS geodatabase model design. Microsoft Visio and Rational Rose are more universally used than ArcGIS Diagrammer in conceptual database modelling. On the other hand, the empty elements are directly prepared for the ArcGIS geodatabase structures (feature class, dataset, relationship class, annotation class, domains, tables, geometric network, etc.) in ArcGIS Diagrammer software. These special structures are missing in MS Visio and Rational Rose. It is necessary to import them into software. ArcGIS Diagrammer can be freely downloaded from ESRI web pages. MS Visio and Rational Rose are commercial software. A future complication is that ESRI has stopped updating support for the Rational Rose software. For these reasons, the ArcGIS Diagrammer is more useful than two CASE tools mentioned above.

ArcGIS Diagrammer software offers the empty geodatabase elements that are represented by boxes. They differ by colour: datasets are brown, feature classes are green, tables are yellow (Fig. 2), relationship classes are grey, and domains are violet. All entities can be dragged and dropped from the palette to the drawing area. The hierarchy of elements is depicted by blue arrows. It is simply to recognize which green feature class belongs to the brown dataset. The properties of datasets and feature class (name, spatial representation,...) are filled and recorded in the property window of each element. Properties could be optionally shown in element boxes in the main window of the schema diagram. Pack and unpack of elements (double arrows in corners) bring detailed information about the structure. The interface of ArcGIS Diagrammer is very intuitive.

A clear structure of the geodatabase is designed in a graphical way in ArcGIS Diagrammer. Graphic expression is more comprehensible than text description (Dobesova 2011a, b). ArcGIS Diagrammer software is suitable for geodatabase modelling. There is good lucidity of the data structure. Database schema from ArcGIS Diagrammer represents useful documentation in a graphic form.

Transfer of the geodatabase schema to another project or user is quick. It is not required to copy the geodatabase with all the old data then delete all that data and replace it with new data. The schema of geodatabase can be transferred alone. ESRI web page offers data models for free utilization in various themes (hydrology, transportation, atmosphere, etc.) (ESRI… 2010). Additionally, the creation of the physical geodatabase from the schema is also quick. Importing it into ArcCatalog automatically creates all the structure from the XML file. It is quicker to remake the database schema in ArcGIS Diagrammer and import them to new geodatabase than to remake physical geodatabase directly.

3. Model for the "Atlas of Election"

The book "Atlas of Election to the Olomouc Region Council" (Vozenilek *et al.* 2009) arose from a cooperation of the Department of Geoinformatics, Department of Political Science at Palacký University in Olomouc and the Czech Statistical Office in 2009. The regional subbranch of the Czech Statistical Office (CZU) in Olomouc provided data about results of elections. The Department of Geoinformatics designed all maps for the "Atlas of Election". Expert descriptions of election were written by political scientists from the Department of Political Science. The Atlas is an example of cartographic visualization of spatial aspects of an election.

The Atlas is divided into three parts according to three election years (2000, 2004, 2008). The number of the valid votes, the success of the largest parties and the election turnout were monitored. All information was collected in each election year and for all districts in the Olomouc Region. The administrative division of the Olomouc Region is into the OPR districts. Furthermore, each ORP district is divided into villages in a detail level. Results of elections were monitored and visualized in the "Atlas of Election" for both levels of the administrative division.

3.1. The logical model of the "Atlas of Election"

The first draft of the geodatabase model considered one dataset with five feature classes. They were three polygon feature classes (village, ORP district boundary, and region boundary), one point feature classes (ORP district) and one annotation class with labels of ORP district. The results of the election were recorded in three attribute tables for every year (Election2000t, Election2004t, Election2008t). The tables with the election results had a various structure of attributes. The attribute "IdVillage" was the primary key in the "election tables". The name and number of political parties differed from year to year. The table contained the number of votes and results in percentage (suffix _REL) for each political party. The example of the structure of the attribute table "Election2000t" is shown in Fig. 1.

It was also recognized that spatial data will differ a little from year to year. In some cases, a boundary district remake was made by the state authority. Subsequently, boundaries of ORP districts and Olomouc region changed their shapes. It was necessary to remake the first geodatabase logical model in ArcGIS Diagrammer. Two new datasets were added. Every dataset was aimed at one election year. Datasets were named Election2000, Election2004 and Election2008. Datasets were used in the geodatabase only as an organizational structure.

Fig. 1. Structure of the table with results of election for the year 2000

All datasets contained the same five feature classes of villages, ORP districts boundary, region boundary, ORP towns and annotations. Names of feature classes were marked by the suffix of the year (Village2000, Village2004, etc.). Structures of feature classes are the same for all three years. The multiplication of feature classes was needed by various shapes of boundaries (real data in the Atlas geodatabase). The final model of geodatabase is in Fig. 2.

The geodatabase model also contained two relationship classes in every dataset. For example, the first relationship class Election2008 defines the relationship between the attribute table Election2008t and the feature class Village2008. The cardinality is 1:1. Primary key IdVillage was the identifier of a village. It is unique for every village in Olomouc District. The relationship class is coloured grey in Fig. 3. The ORP districts were labelled by name on the map (Fig. 5, 6 – Hranice, Konice, Olomouc, etc.). Label texts were stored in the annotation feature class M2008. An annotation feature class is used for the label placement at the same position on various maps. The second relationship class Annotation2008 (grey colour) joined the annotation class M2008 with the feature class Orp2008. The cardinality of relationship was also 1:1. The existence of the annotation class brought four coded domains to the model. They are four violet boxes in Fig. 2. The coded value included the allowed values of the horizontal and vertical alignment of texts on the map. The final geodatabase model contains all the necessary tables with the election results and structures in great detail.

The datasets contained the same five feature classes:
 – **Village** – polygon feature class of villages in the Olomouc region.
 – **RegionBoundary** – polygon feature class of the whole Olomouc region.
 – **OrpBoundary** – polygon feature class of the ORP districts in the Olomouc region.
 – **Orp** – point feature class of the ORP towns.
 – **M** – annotation class with the text labels of the ORP districts.

3.2. The physical model and cartographic methods

Finally, the geodatabase schema was exported from ArcGIS Diagrammer to the XML file (Publish function). This XML file was imported to a new empty geodatabase in the ArcCatalog application. ArcCatalog is a manager of spatial data for ArcGIS software. All geodatabase structures, previously designed into the model, are automatically created by import. The structure of geodatabase after import is shown in Fig. 4.

Real data were imported to the empty structure of the geodatabase from shapefiles. Subsequently, maps were constructed in ArcGIS Desktop - ArcMAP. Seven maps described the election for one election year in the "Atlas of Election to the Olomouc Region Council". The total number of maps was 21.

Examples of two maps are provided in Fig. 5 and Fig. 6. These maps demonstrate two types of cartographic methods – a choropleth map and a diagram map (Kanok 1999; Vozenilek *et al.* 2011). Both methods were used

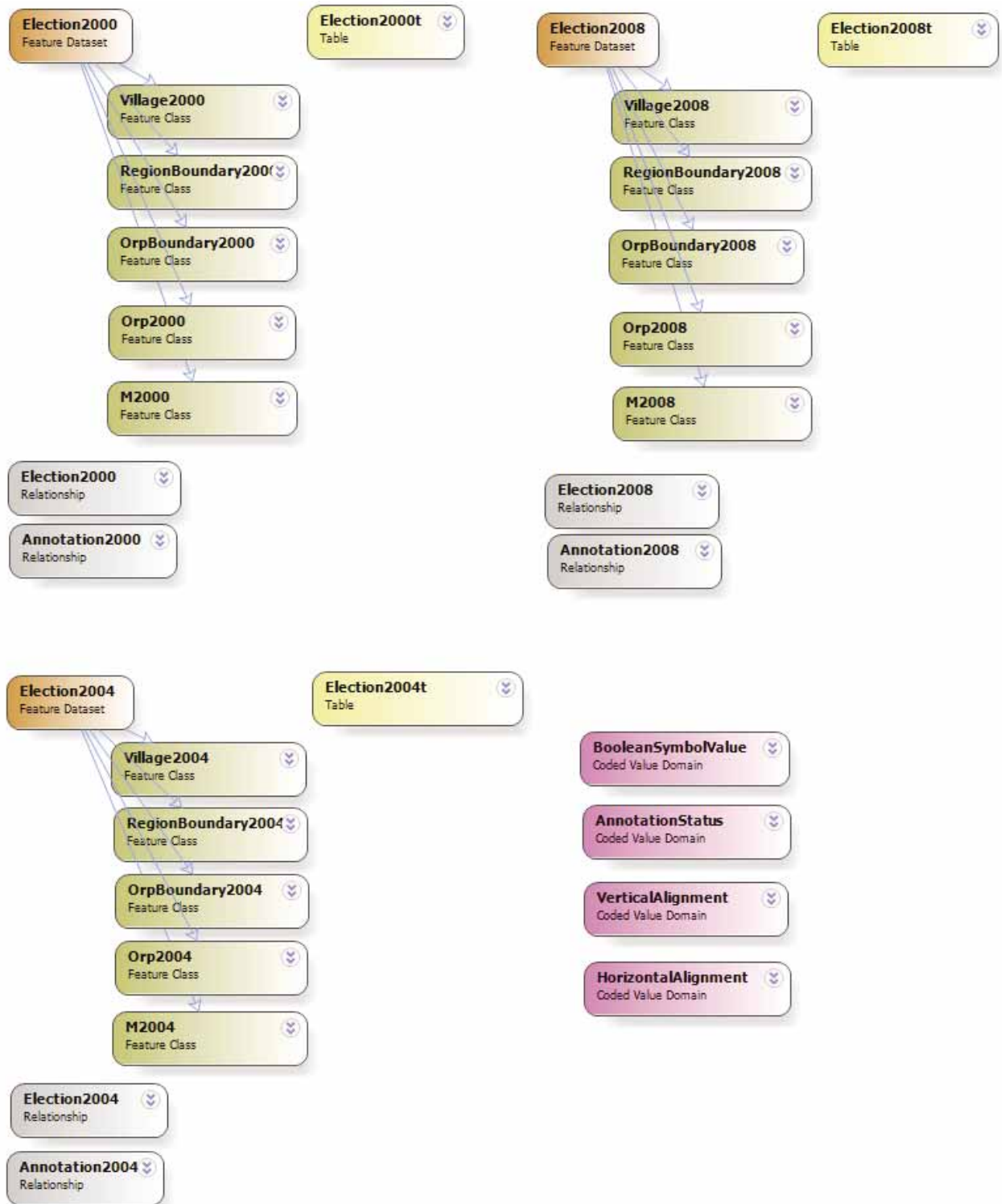**Fig. 2.** The structure of "Election2008" dataset

Fig. 3. Database model of the "Atlas" in ArcGIS Diagrammer

for expression of election results in the "Atlas of Election". The first map (Fig. 5) expresses the portion of valid votes for one party ODS in all villages. The portion of votes oscillated from 0% to 40% in villages. The choropleth map expresses the value in tones of blue (Brewer 2002). Blue colour is "the political colour" of the ODS party. Author of the map reflected this fact in the map.

This cartographic method is suitable for exploring the electoral preferences in individual villages. The geometric part of the map was taken from the feature class "Village2008", and numeric data was taken from the attribute table "Election2008". Point feature class "Orp2008" was used for visualization of the main towns of ORP districts. Districts were labelled from annotation feature class "M2008".
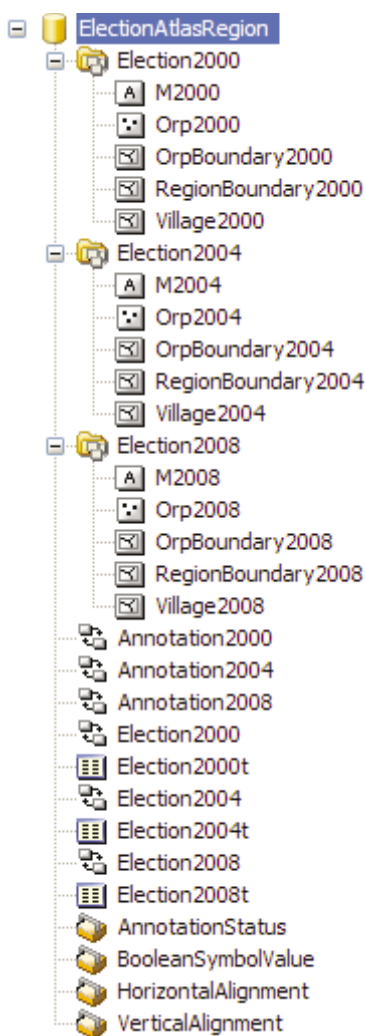
Fig. 4. Structure of the file geodatabase in ArcCatalog after import from the XML file
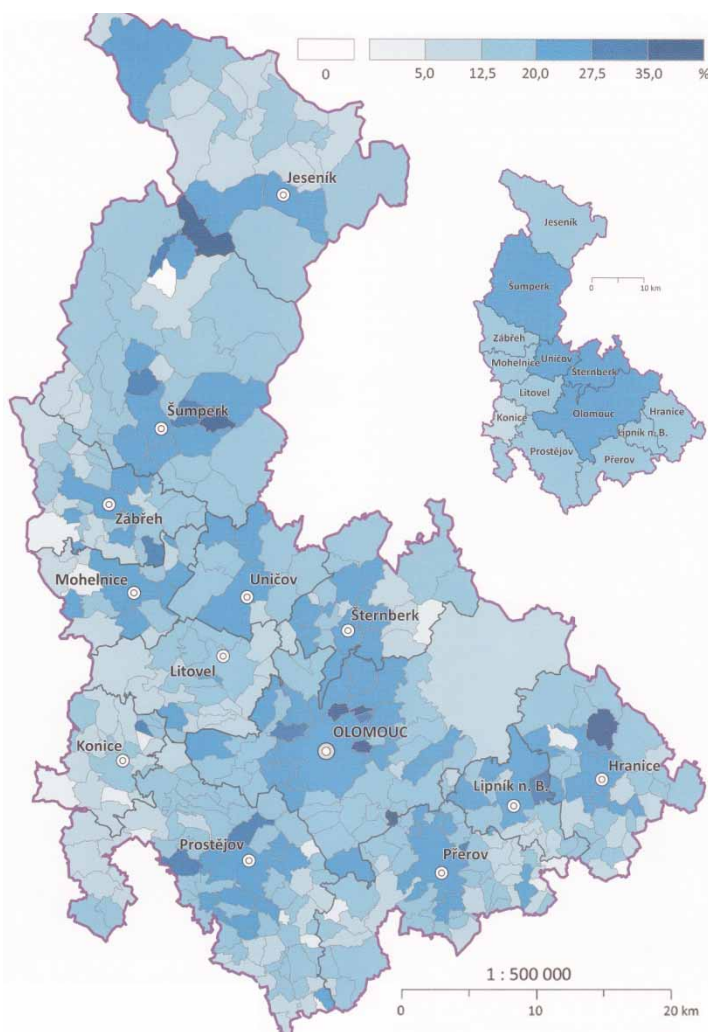

Fig. 5. Portion of valid votes in villages for the ODS party (2008)

The second map (Fig. 6) contains bar diagrams with aggregate results of five political parties at the administrative level of the ORP districts. Reader of the map can compare the results of the parties in diagrams. The colours of bars correspond to the "political colours" of parties. Geometric part of the map was taken from the feature class "OrpBoundary". The feature class "RegionBoundary" was used for the small inset map in the right upper corner. Numeric data was taken from the attribute table "Election2008".

For the map outputs, two cartographical methods were used – diagram map and choropleth map. Nevertheless, other cartographical diagram methods can be used in ArcGIS program and with Diagram Map Creator - custom extension of ArcGIS (Dobesova, Valent 2011). Suitable cartographical methods for expression of the results were chosen according to the cartographical literature and expert-cartographers (Vozenilek *et al.* 2011; Brus *et al.* 2009; Dobesova, Brus 2011).

One geodatabase for the "Atlas of Election" concentrated all source data for all maps in the book in the time series. More than 21 maps were created from one data source. The structure was suggested to consider all differ-

ences in shapes of territory. For that purpose, data were organized into the separate logical parts – three datasets according to the three election years.

From the cartographical point of view, the two geodatabase structure had the cartographical contribution. The first is the definition of the relationships with tables, because any map project in ArcGIS Desktop can take the predefined relationships from geodatabase. These relationships connected numerical data to the polygons (village) as a source data for the choropleth maps and the bar diagrams. Another possibility was to create the relations directly in the map projects in ArcGIS. In that situation, it was necessary to create a relationship in every map project repeatedly. The existing geodatabase relationships could be used repetitively for other new map projects. The definition of the geodatabase relationships were a better solution than the definition of the map relationships. The second cartographic contribution was the definition of annotation classes in every datasets in the geodatabase. Annotation classes contained names of ORP districts, information about text size, font and position of text in the map. Using the annotation class as the map layer, the text positions were the same on all maps.
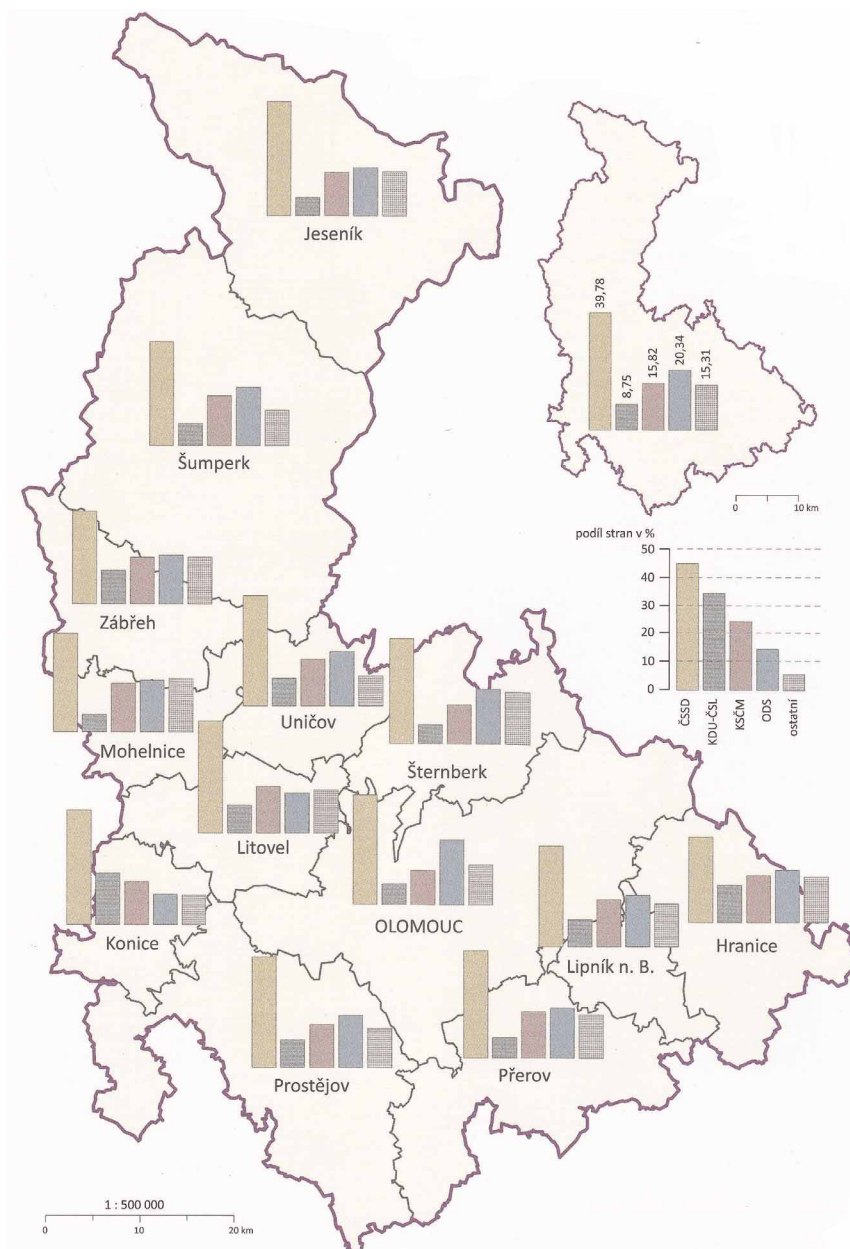
Fig. 6. Diagram map with the results of five political parties (2008)

Therefore, all maps had the same appearance in the labelling arrangement. Both geodatabase structures that were caused by the cartographical requests brought assets in the map creation.

4. Conclusion

The process of a logical modelling of a geodatabase is very useful. ArcGIS Diagrammer was used as a CASE tool for ArcGIS software. It was applied to stages of identification of the key layers and structures, as well as to processes of refining and documentation of the model. The cartographical requests introduced relationship and annotation classes into the model of the geodatabase. The geodatabase structure also respected the time series of the election results and the changes of the district shapes. The case study of the "Atlas of Election" shows that the structure of the geodatabase is under the influence of cartographical purpose.

The database schema of the "Atlas of Election" represents a universal model for the Czech Republic. This model was primarily created for the Olomouc Region. Nevertheless, the same situation exists in the other 13 regions of the Czech Republic. The elections took place in other regions in the same years as well. All regions are divided into the same administrative system of districts and villages. The advantage of models is design sharing between users and projects. The model for the "Atlas of Election" can be used again and again in cartographical tasks. Small changes can be made for specific projects. For example, maps about next election in 2012 need only the addition of the new attribute table "Election2012t" with actual political parties and actual results of the election. Final geodatabase can be utilized as the database for the new map outputs.

The presented geodatabase schema would be the inspiration for cartographers who create maps using GIS.

References

Arctur, D.; Zeidler, M. 2004. *Designing Geodatabase, Case Studies in GIS Data Modeling*. ESRI, Readlans, USA. 393 p. ISBN 1-58948-021-X.

Brus, J.; Dobesova, Z.; Kanok, J. 2009. Utilization of expert systems in thematic cartography, in *International Conference on Intelligent Networking and Collaborative Systems, INCoS*. Barcelona, 285–289. ISBN 978-076953858-7.

Brewer, C. A. 2002. *Designing Better Maps*. New York, NY: Environmental Systems Research Institute. 203 p.

Dobesova, Z.; Brus, J. 2011. Coping with cartographical ontology, in *Conference Proceedings SGEM 2011, 11th International Multidisciplinary Scientific GeoConfrence. STEF92 Technology Ltd.* Sofia, Bulgaria, 377–384. ISSN 1314-2704.

Dobesova, Z. 2011a. Visual programming language in geographic information systems, recent researches in applied informatics, in *Proceedings of the 2nd International Conference on Applied Informatics and Computing Theory, AICT '11*. Prague, NAUN/IEEE.AM, WSEAS Press, 276–280. ISBN 978-1-61804-034-3.

Dobesova, Z. 2011b. Programming language python for data processing, in *Proceedings of International Conference on Electrical and Control Engineering (ICECE), Yichang, China, Volume 6*. Institute of Electrical and Electronic Engineers (IEEE), CFP 1173J-PRT, 4866–4869. ISBN 978-1-4244-8163-7.
http://dx.doi.org/10.1109/ICECENG.2011.6057428

Dobesova, Z.; Valent, T. 2011. Program extension for diagram maps, *Geodesy and Cartography* 37(1): 22–28.
http://dx.doi.org/10.3846/13921541.2011.558330

ESRI Support. 2010. Geodatabase – Data Models. Available from Internet: http://support.esri.com/en/downloads/data-model

ESRI. 2002. Status report on ArcGIS data models, ArcNews 24(3): 6–7.

ESRI. 2011a. Building Geodatabases with CASE Tools. Available from Internet: http://resources.arcgis.com/content/building-geodatabases-case-tools

ESRI. 2011b. ArcGIS Resource Center. ArcGIS Diagrammer. Available from Internet: http://resources.arcgis.com/gallery/file/arcobjects-net-api/details?entryID=F12ADF8F-1422-2418-34B2-C276C6BCCF98

Kanok, J. 1999. *Thematic Cartography*. University of Ostrava, Ostrava. 318 p. (in Czech).

Longley, P. A.; Goodchild, M. F.; Maguire, D. J.; Rhind, D. W. 2001. *Geographic Information. Systems and Science*. Wiley. ISBN 0471495212.

Vozenilek, V.; Saradin, P.; Nemcova, Z.; Demlova, N.; Brychtova, A.; Ptacek, P. 2009. *The Atlas of Election to the Olomouc Region Council 2000, 2004 and 2008*. Publishing house of Palacký University, Olomouc. 65 p. ISBN 978-80-244-2300-5 (in Czech).

Vozenilek, V.; Kaňok, J., *et al.* 2011. *Methods of thematic cartography. Visualization of spatial data*. Publishing house of Palacký University. 216 p. ISBN 978-80-244-2790-4 (in Czech).

Zdena DOBESOVA. PhD Doctor. She holds a PhD (since 2007) degree from Technical University in Ostrava, Faculty of Mining and Geology. Research interests: GIS, spatial database, digital cartography, programming, scripting in Python for ArcGIS.