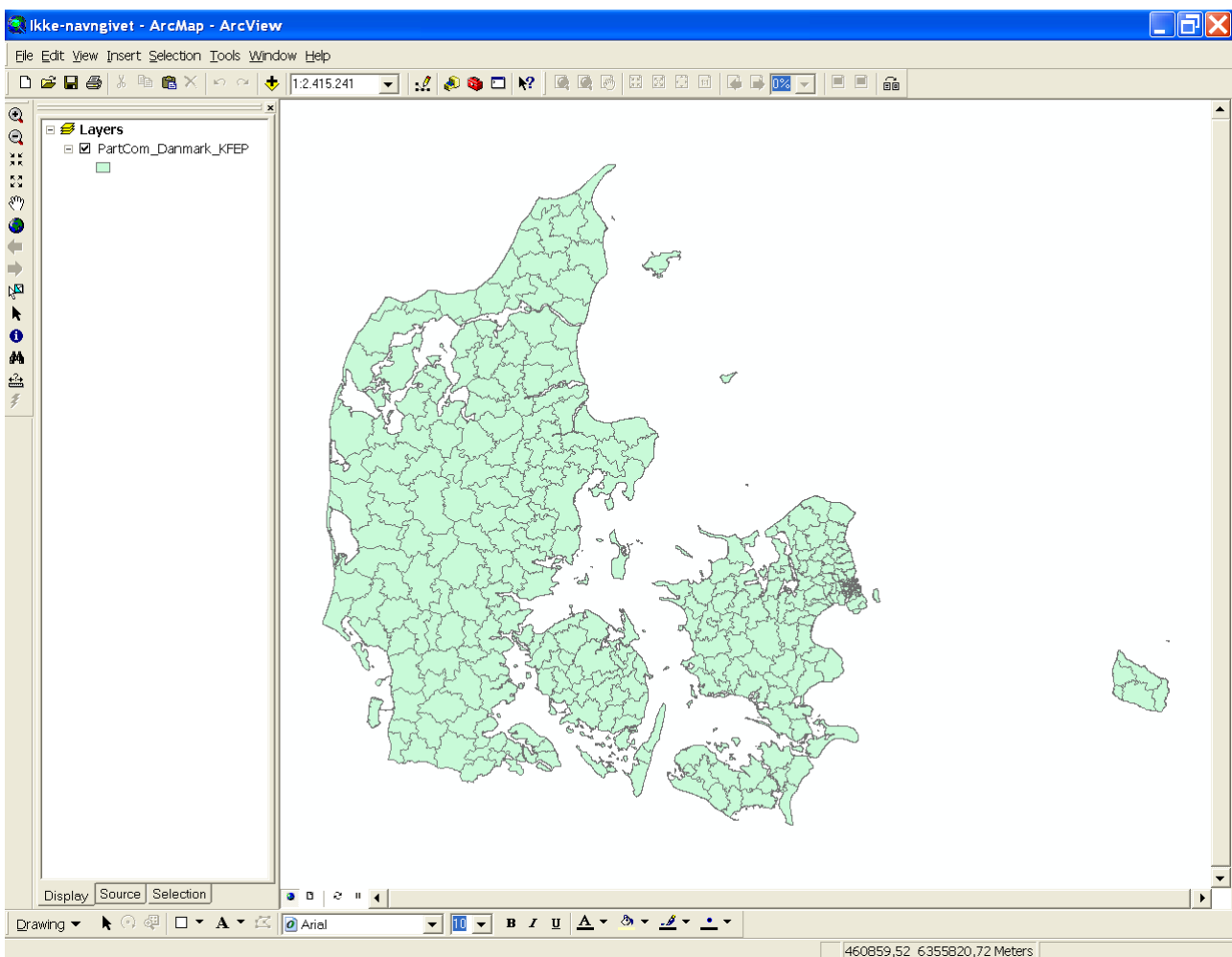


How To (2): Manipulate Existing Data to Produce a Display of Aggregated Features

This short paper provides a quick step-by-step guide to generating larger geographical data units (in this instance Danish constituencies) based on a specified attribute with ArcGIS 9.1. By the end of this tutorial You should be able to dissolve geographic features, aggregate associated attributes, and – on a somewhat different note – dispose of unnecessary data. As well, You should know how to utilize these methods for better map displays.

Example Research Objective: None at present.

Modus Operandi: Start up ArcMap on your computer and add “PartCom Danmark KFEP.shp” from “O:\Kristian (2005)\PARTCOM\færdige kort” to an empty map. These few actions should present you with the following screen (toolbars may be positioned differently):



Considering the .shp (shape) file extension, the added data not surprisingly is a geographic boundary file containing a graphic representation of the included data, specifically Danish municipalities outside of the capital and polling districts within it. As you may notice, however, the latter features are quite difficult to distinguish until you zoom in on the black blotch on the east coast of the rightmost of the three larger land masses called Zealand (if unable to see the zoom icons depicting a magnifying glass open the Tools toolbar through the View\ Toolbars drop-bar from the menu bar at the top). Moreover, it is not immediately obvious, which constituency either of the two types of features belong to unless you are exceedingly familiar with the Danish administrative and electoral systems.

The following steps will address each issue in turn.

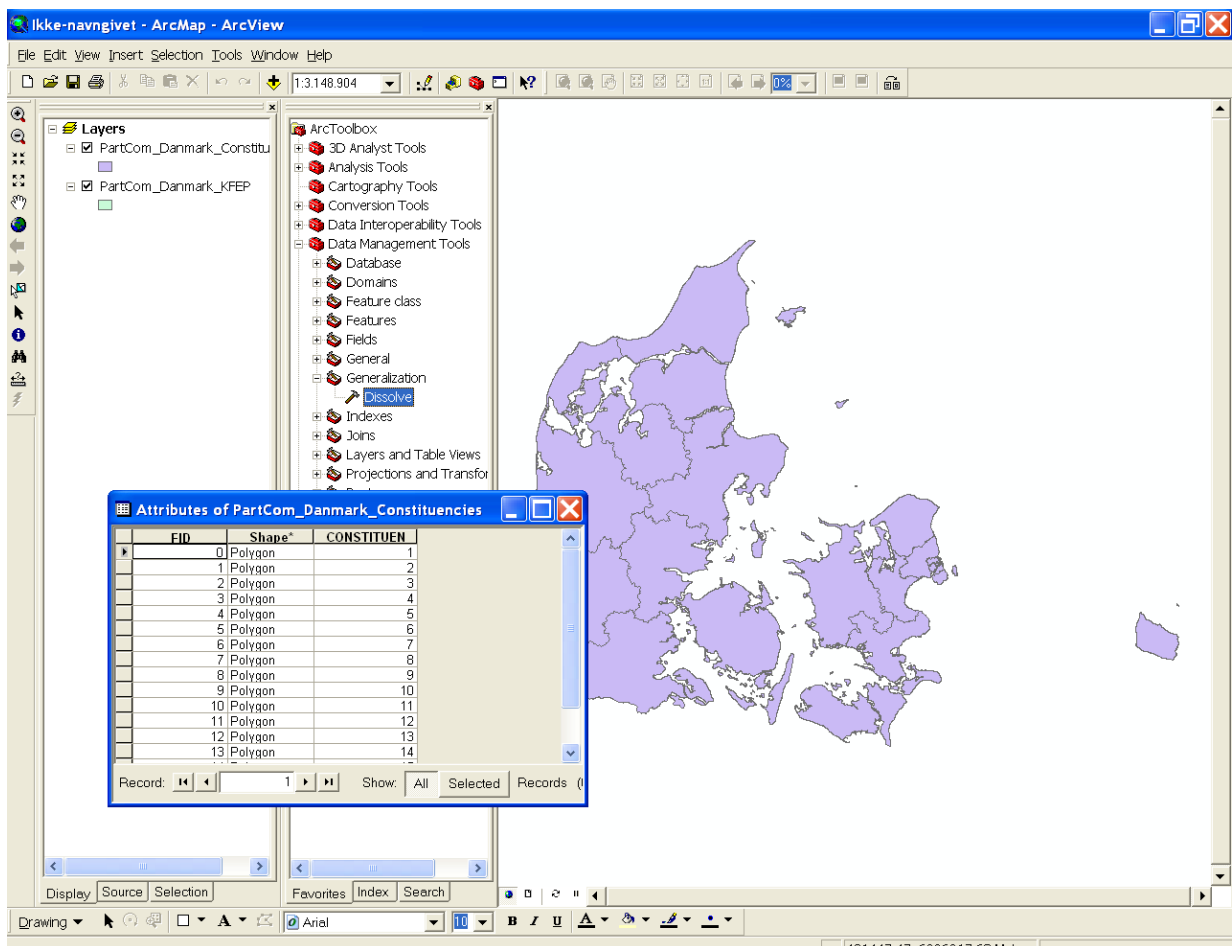
- 1) If – for some reason – what you’re really interested in is at the constituency level, turning smaller geographical units into larger units is a relatively simple task. All that is needed is a set of values identifying each unit’s individual class and ArcGIS will do the rest. Hence, open the attribute table of the layer by right-clicking “PartCom Danmark KFEP.shp” in the table of contents to the left and select Open Attribute Table. As it is shown in the picture below, it is apparent that the required information is readily available:

FID	Shape*	SR_PARTCOM	MUNICIPALI	DISTRICT	CONSTITUEN	REGION	MAINAREA	COUNTRY	COUNTYNAME	
0	Polygon	10101	101	1	1	1	1	1	Søndre	2. Christia
1	Polygon	10102	101	1	1	1	1	1	Søndre	2. Syd
2	Polygon	10103	101	1	1	1	1	1	Søndre	2. Øst
3	Polygon	10104	101	1	1	1	1	1	Søndre	2. Vest
4	Polygon	10105	101	2	1	1	1	1	Søndre	3. Rådhus
5	Polygon	10106	101	2	1	1	1	1	Søndre	3. Øst
6	Polygon	10107	101	2	1	1	1	1	Søndre	3. Vest
7	Polygon	10108	101	3	1	1	1	1	Søndre	4. Sundby
8	Polygon	10109	101	3	1	1	1	1	Søndre	4. Øst
9	Polygon	10110	101	3	1	1	1	1	Søndre	4. Vest
10	Polygon	10111	101	4	1	1	1	1	Søndre	5. Blågård
11	Polygon	10112	101	4	1	1	1	1	Søndre	5. Nord
12	Polygon	10113	101	4	1	1	1	1	Søndre	5. Syd

Besides the usual FID and Shape* fields and the unique identifier SR_PARTCOM, a field labeled CONSTITUEN assigns each feature a value corresponding to one of the 17 Danish constituencies.

To dissolve the municipalities and polling districts on this attribute, either select ArcToolbox through the Window drop-bar from the menu bar at the top or just press the

ArcToolbox icon (depicting a red toolbox) to the right of the Map Scale field above the main panel. Then, in the new toolbox panel navigate to Data Management Tools\Generalization and double-click on the Dissolve Tool (all tools are symbolized with a hammer), which opens the Dissolve Features Wizard. In it, first, choose the input features you wish to dissolve (in the current example there is only one option); secondly, choose the output feature class, i.e. the file name you wish to store the new layer under; and thirdly, choose CONSTITUEN as the dissolve field on which to aggregate features. Keep in mind that – similar to joining tables - in principle field names do not matter as you could choose any attribute upon which to dissolve features however meaningless. Pressing OK now, the resulting layer should look like this:



Had the CONSTITUEN variable not been readily available, such a set of values would have had to be provided first either by joining another table with the information to the original layer or by adding a field to the layer's attribute table and calculating the values. Though you need not do so at this point, additional fields may be added to the attribute table through the Options drop-bar in the attribute table itself and subsequently calculated

upon by right-clicking the field label and selecting Calculate Values... (if selecting features, the Calculate Values... operation will only apply to those features).

- 2) The procedure described above only introduces half of what the Dissolve Tool can do, namely dissolving geographic features into larger units. In addition, the Dissolve Tool offers several options for aggregating associated attributes. Thus, if data is joined to the geographic boundary file, it is possible to obtain these summarizing values: Sum, Mean, Minimum, Maximum, Range, Standard Deviation, First-, and Last value. In order to see how this task is accomplished repeat the above procedure, but instead of pressing OK scroll down the wizard interface to reveal two other data entry selections. Select COUNTYNAME (identical to the names of the constituencies) in the statistics field and it will appear in the field part of the table below. The statistic type First can then be selected in the empty cell to the right thereof. Pressing OK now results in a basically identical layer, however, with a new field in the attribute table labeled FIRST_COUN:

FID	Shape*	CONSTITUEN	FIRST_COUN
0	Polygon	1 Sandre	
1	Polygon	2 Østre	
2	Polygon	3 Vestre	
3	Polygon	4 København	
4	Polygon	5 Frederiksberg	
5	Polygon	6 Roskilde	
6	Polygon	7 Vestsjælland	
7	Polygon	8 Storstrøms	
8	Polygon	9 Bornholms	
9	Polygon	10 Fyns	
10	Polygon	11 Sønderjyllands	
11	Polygon	12 Ribe	
12	Polygon	13 Vejle	
13	Polygon	14 Ringkjøbing	

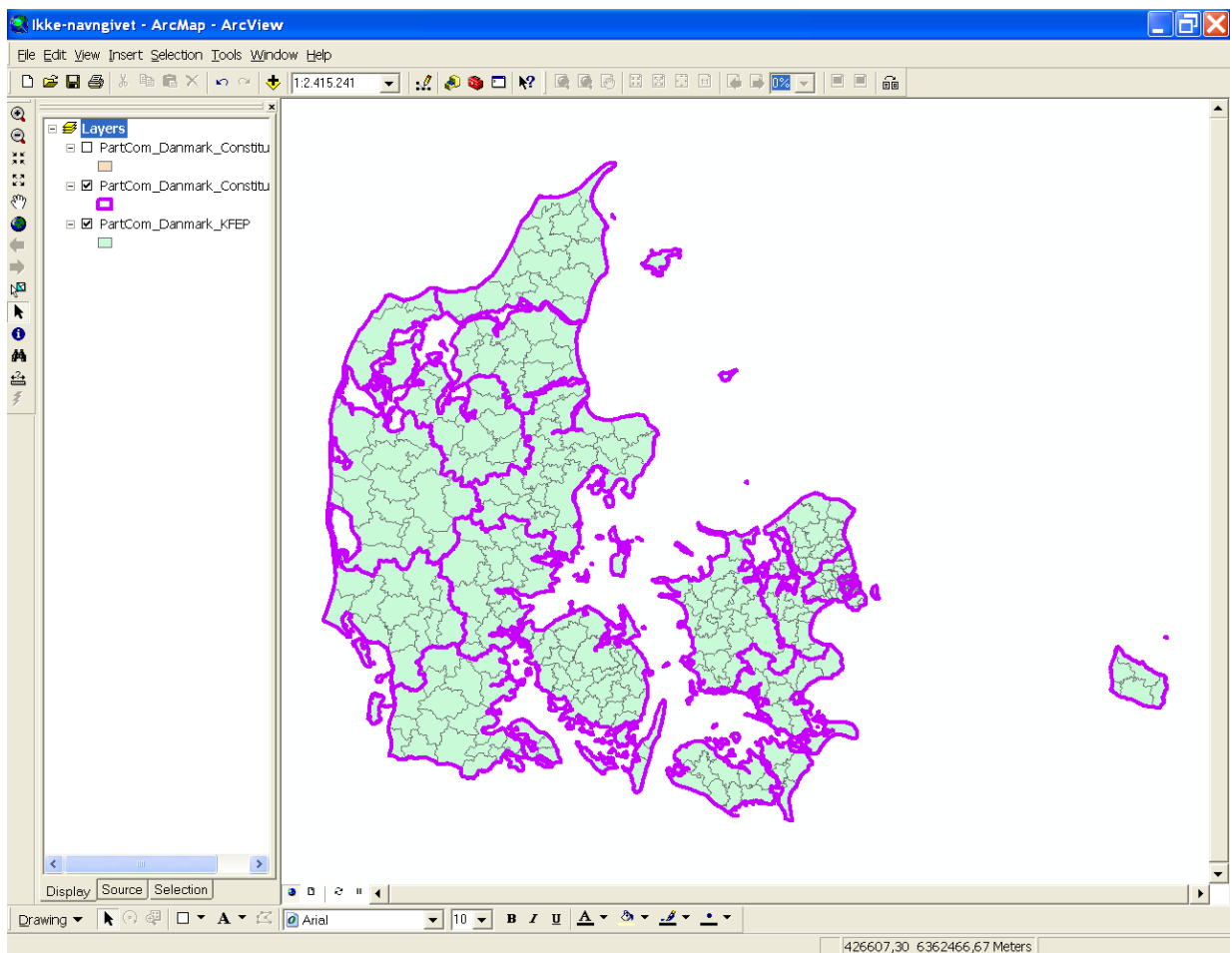
Not having to aggregate data separately constitutes an evident advantage in doing fast, limited transformations. Yet the procedure of using the Dissolve Tool quickly becomes cumbersome as the number of variables (or the number of summary statistics you want for each variable) increases. Accordingly, it may be advisable to nevertheless separate data attributes and boundary files in the incidence of wanting to keep many variables (on the other hand this solution poses the minor issue of maintaining the correct FID, see below). Furthermore, ArcGIS does not at present allow for the renaming of variables, which can have – indeed is somewhat likely to have – the unfortunate consequence of creating some confusion about the content of each field after aggregation as labels can only be ten characters long (e.g. COUNTRY and COUNTYNAME would both be

FIRST_COUN and consequently named FIRST_COUN and FIRST_COU1, not to mention the resulting FIRST_FIRS and FIRST_FIR1 of aggregating the same attributes twice)

Given these two caveats a few pointers on the use of Excel and the dBASE format (some of this *may* apply to SPSS as well) seem warranted: It is quite easy to change field names or recalculate entire fields in Excel. Attribute tables are saved in the dBASE format and always have a .dbf file extension by which to recognize them. Simply open them as any other spreadsheet and just ignore the warnings when trying to save or close them. But do not under any circumstances change the order of rows in a file attached to a geographic boundary file unless you are able to return to the initial order before saving! The reason is that while the dBASE file may show your unique identifier (i.e. SR_PARTCOM), which you use to tie the geographic features to your data, it does not include the program's unique identifier (i.e. the FID), which actually ties the data attributes in the attribute table to the geographic features. Sort the dBASE file in Excel without restoring order and suddenly attributes will appear to be associated with random features (this typically becomes noticeable, when you try to aggregate supposedly contiguous units). Secondly, while attribute tables in ArcGIS may store more variables, Excel can only manage 256 columns. Variables beyond number 256 will be lost in using Excel – at least when you save your dBASE file! Finally, when saving in the dBASE format only visible characters will be transferred to memory so adjust your column width to show the entire attribute length before saving; keep in mind as well that variable names never can exceed ten characters; and generally avoid special characters, spaces, and dashes besides underscore in naming variables (this caution goes for naming files too).

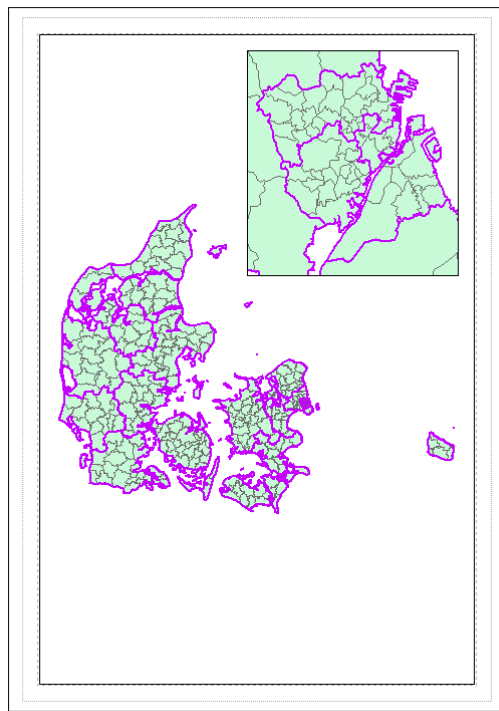
- 3) So far we've assumed that dissolving municipalities and polling districts into constituencies represented a suitable solution to the issue of determining how each of those features contributed to the constituency level. Sometimes, dissolving the smaller units is not a desirable outcome, however, as this procedure levels any internal variation (imagine a heavily populated urban area surrounded by mainly rural areas for instance). Visually, if not statistically, a better solution may thus often be to rather outline the extent of the larger units on top of the smaller ones. This idea amounts to again dissolving the

units like before, but rendering the resulting layer transparent to see the original layer beneath. Clicking on the colored field in the table of contents allows you to change the fill color of the top layer to “no color”. Similarly, you can change the color and width of the outline to distinguish it from that of the lower units, which it inherently will follow to some extent. Once satisfied, pressing OK should leave you with a display similar to this:

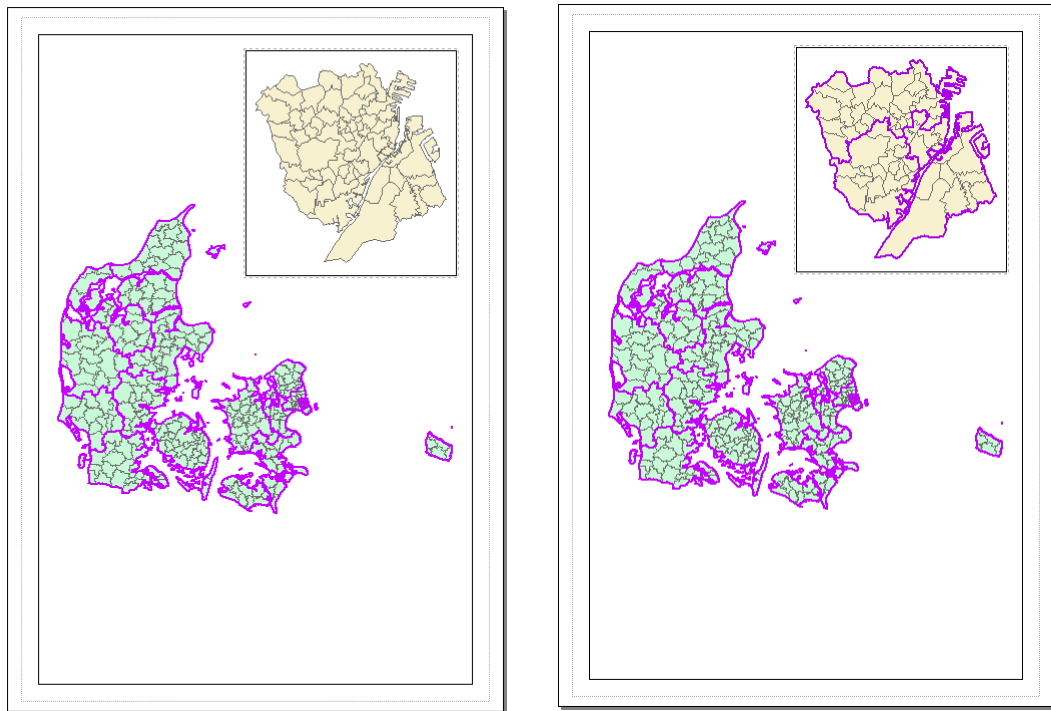


Thickening the lines, however, plainly does not make it any easier to distinguish the polling districts in the capital area. To remedy this deficiency, it is useful to remember that several picture frames may be added to the layout view making it feasible to show both the entire map and close-ups of any part of the map at the same time. As picture frames are inherently related to data frames they can be added in two ways, either by inserting empty data frame in the data view (select Data Frame through the Insert drop-bar from the menu bar at the top) or by copying a picture frame in the layout view (activate any frame and press Ctrl+c followed by Ctrl+v). We will do the latter to automatically transfer all layers from the first

frame to the second (to switch between views press the icons at the bottom of the main panel depicting a globe and a sheet of paper respectively). Next, open the attribute table of “PartCom_Danmark_KFEP” in the new frame and select all features within municipalities 101 and 147 (select each feature directly in the table or perform an SQL search by selecting Select by Attributes... through the Options drop-bar in the attribute table); then Zoom to Selected Features through the Selection drop-bar from the menu bar at the top. After clearing all selected values through the Selection drop-bar and readjusting the size of the second picture frame the layout view should now look something like this:



If preferable it is also possible to remove the surrounding “unwanted” areas altogether: Simply select the features within municipalities 101 and 147 once again, right-click the layer, choose Export Data... from the Data drop-bar, and save the layer under a new file name. The layout will look like this:



Once having made a satisfying layout keep it as a template by saving the layout to a map file, which you may access at any time to do an exactly similar display of some other interesting data attributes.

Exercises:

- Redo the map of Denmark to show districts and regions.
- Create a compact layout showing both Denmark and the capital area as well as the lone island of Bornholm to the far right with as little empty space as possible.
- Incorporate this layout into a text document to see how it looks in real size.

Kristian (2006)