

How To: Make a map displaying comparable voting districts through time using 0-tolerance data from Den Danske Valgdatabase and ArcMap 9.3

This paper will present to you the procedures and tools which are necessary to master when making a map of voting districts that are comparable through time using ArcMap 9.3. The Political Ecological Database Archive (PEDA) project provides you with a unique opportunity to get table data from two or more elections that are geographically comparable and therefore, it is hypothesized, demographically comparable as well. This can be accomplished by downloading so-called 0-tolerance data between any two or more elections from Den Danske Valgdatabase¹. The 0-tolerance data will be described in more detail in section 1 when the theoretical argument for using 0-tolerance data is explained more thoroughly. Section 2 takes you through the primary steps of downloading data from Valgdatabasen and making it ready for further treatment and map making in ArcMap 9.3, whereas section 3 deals with the subsequent procedures in ArcMap that will lead to the actual creation of a 0-tolerance shapefile. Throughout these sections an illustrative example will be given, namely of the creation of a 0-tolerance shapefile based on the polling districts at the November 2001 and February 2005 Danish National Parliamentary elections. In the end, in section 4, this 0-tolerance shapefile will be used to create a map showing the change in logit-transformed vote percentage for the Danish party Det Radikale Venstre in every comparable (0-tolerance) district between these two elections.

1) Why use 0-tolerance election data?

Every now and then the Danish polling districts change in number and extent (shape). The cause of these changes can be many, one of them being the impact of administrative reforms as witnessed in January 2007 when a major structural reform took place and changed the administrative landscape

¹ Den Danske Valgdatabase is part of the PEDA project. It is a scientific database containing Danish election data for each party running for election in each type of election, be it a National Parliament election, a European Parliament election or a national referendum, held in Denmark since 1979. It is regularly updated with the most recent election data.

of Denmark. At that time for instance the 103 former Danish nomination districts were replaced by 92 new ones and the number of Danish municipalities went from 175 to 98. Since the Danish authorities have strived to align the polling districts to the borders of these more aggregated administrative units, many polling districts also changed their shape at the time of the reform. Nevertheless, at least going back to the Danish National Parliamentary election in 1994 greater or minor changes have occurred between each two neighboring elections, national structural reform or not. The trend has gone in the direction of making larger and fewer polling districts over time. For example at the 1994 Danish National Parliamentary election there were 1916 polling districts, at the 2001 election there were 1899 polling districts, at the 2005 election there were 1855 polling districts and at the 2007 election, after the structural reform had taken place, only 1645 polling districts existed. Some of these have simply been merged together two and two maintaining their outer borders, but it is not always that simple.

Because of this, it is not straightforward to compare election data between two different elections. One cannot just compare an original polling district on a given variable, e.g. the logit-transformed vote percentage for a party in this polling district, at two different elections, if meanwhile the polling district has changed its extent since then the demographics of the original polling district will probably have changed as well. So, to be able to compare election data on smaller geographic units through time², we need to make these units geographically (i.e. demographically) comparable. Using the 0-tolerance function when downloading election data from Den Danske Valgdatabase will do that for us. In general terms, what the 0-tolerance function does is that it uses road record data for every polling district at both elections and on the basis of that traces down all polling districts which have changed their road record, i.e. changed their shape between the two or more elections in question. It then aggregates these polling districts until the final geographic area is comparable between the elections in question. An illustration is given below in figure 1-4 from the Northern coast of Jutland where three 2001 polling districts (with black borders highlighted, see figure 1) were changed geographically between the 2001 and 2005 Danish National Parliamentary elections such that at the 2005 election only two polling districts (with red borders highlighted, see figure 2) were encompassed in the same geographic area. In the final 0-tolerance shapefile, which will be constructed in the following sections, these polling districts will be merged into one larger area using the 0-tolerance table data since this larger area (with dark green borders highlighted, see figure 3 and 4) is comparable between the two elections.

² Instead of treating Denmark as one big geographic region.

Figure 1: Three polling districts from the 2001 Danish National Parliamentary election is highlighted

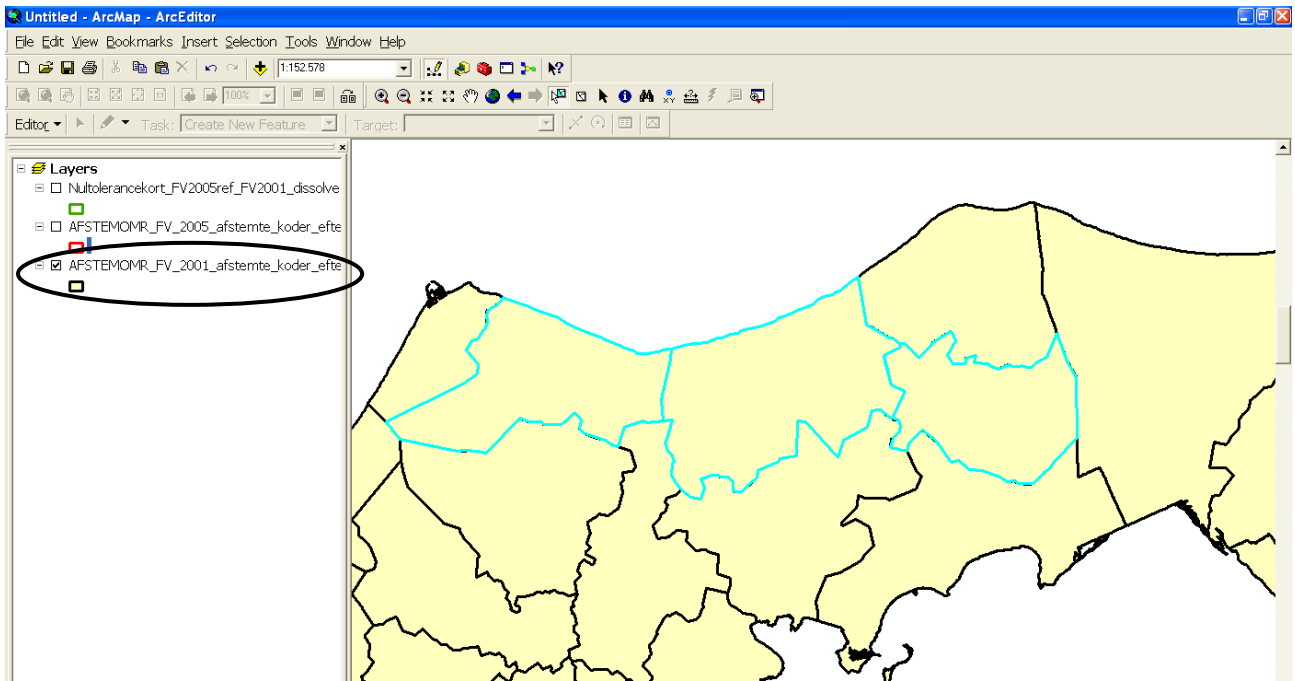


Figure 2: Before the 2005 Danish National Parliamentary election the three polling districts in figure 1 were modified and turned into only two polling districts

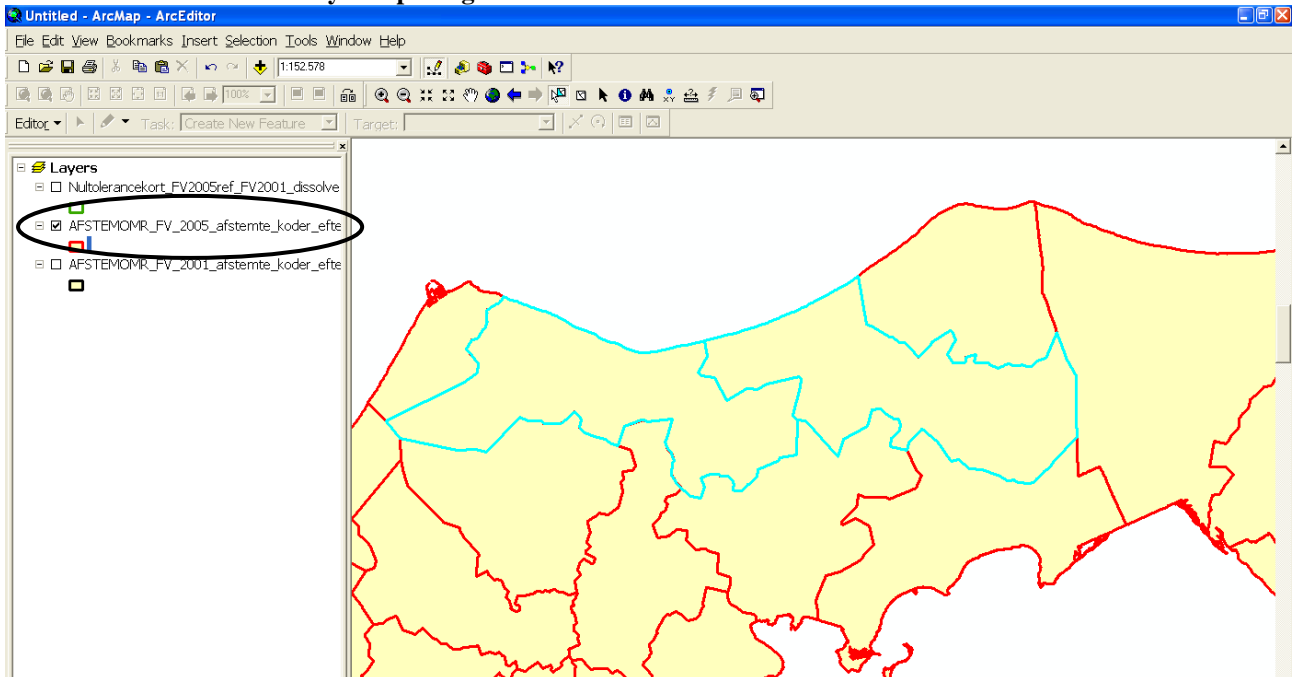


Figure 3: The modified polling districts in figure 1 and 2 have been aggregated in the 0-tolerance table data such that a 0-tolerance shapefile (to be created below) will contain only one larger area covering them all.

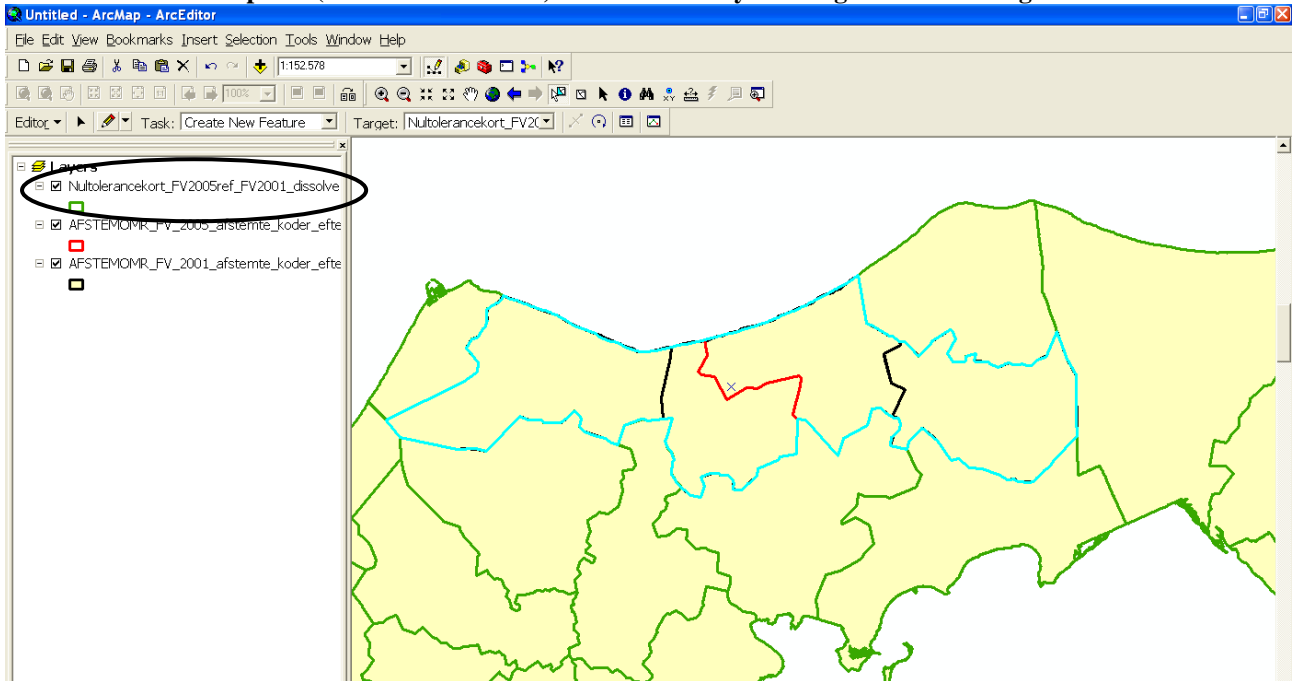
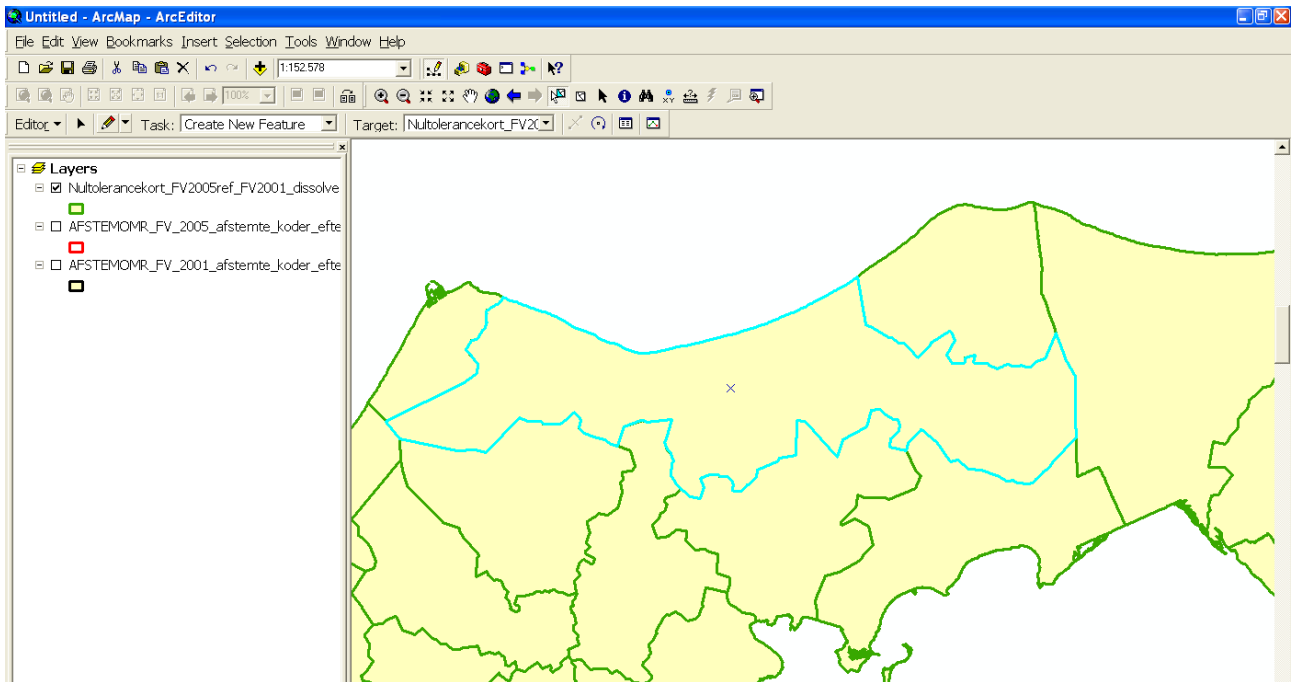


Figure 4: The resulting 0-tolerance district is highlighted in the final 0-tolerance shapefile



2) Downloading 0-tolerance data from Den Danske Valgdatabase and making it ready for ArcMap 9.3

In Den Danske Valgdatabase, when choosing which two or more elections you want comparable election data from you must remember to set the tolerance option called “Maks. divergens” to 0, see figure 5 below. Moreover, you need to choose one election as your reference election on which you want the 0-tolerance computations to be based. In this case, the two elections chosen for 0-tolerance computations will be the 2001 and 2005 Danish National Parliamentary elections where the 2005 election will serve as reference election, see figure 5 below. Finally, you need of course to choose polling districts as the basic geographic unit on which the 0-tolerance computations will be based, see figure 6 on page 6.

Figure 5: Downloading 0-tolerance data from Den Danske Valgdatabase

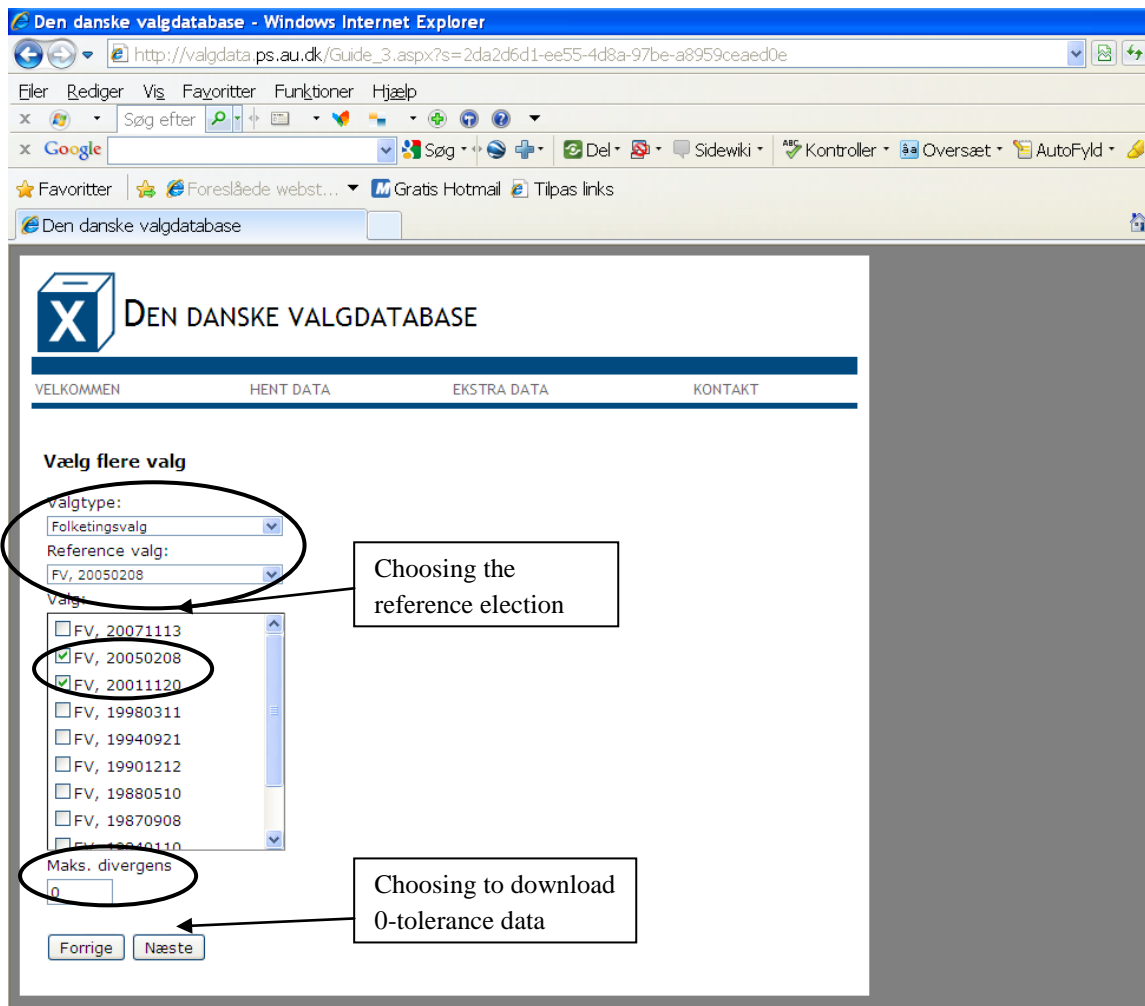
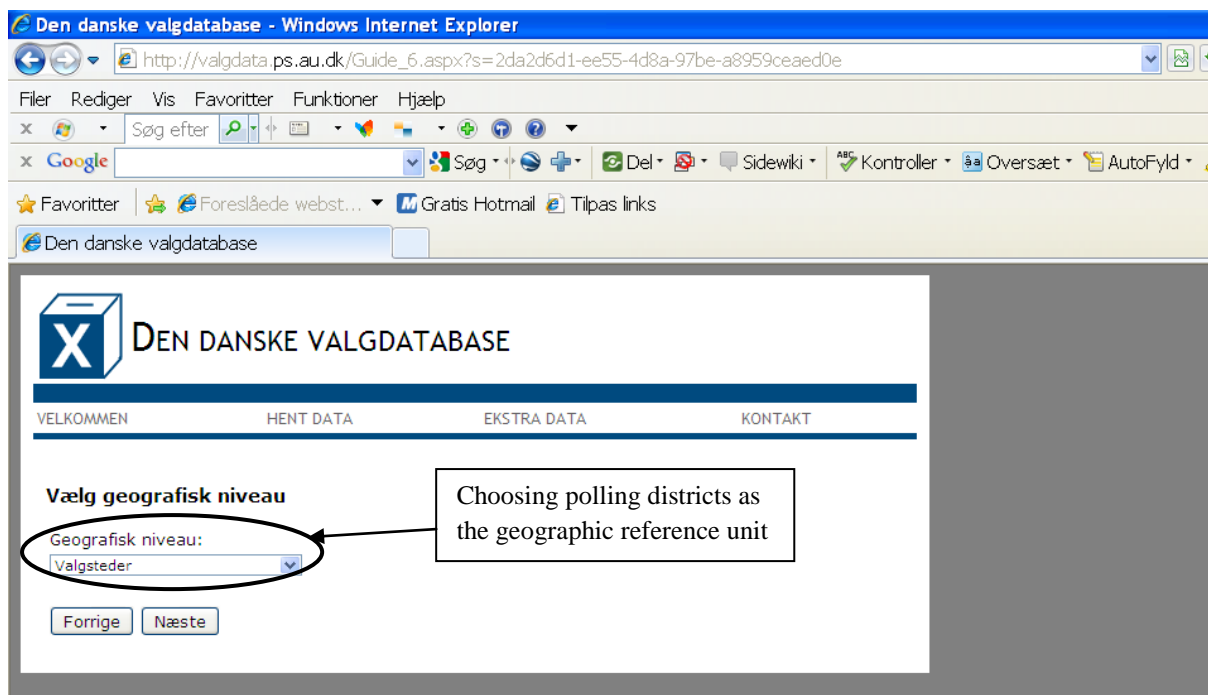


Figure 6: Choosing polling districts as the geographic unit of reference for the 0-tolerance calculations



Whenever you download data from Den Danske Valgdatabase, data will be delivered in an Excel-file in long format with party specific voting data from each election placed after each other in the same columns. This is also the case when requesting 0-tolerance data. In our example, we get an Excel-file with the 2001 party specific election data followed by the 2005 data in the same columns and with a row for each new 0-tolerance district created on the basis of the 0-tolerance computations. Since 1774 0-tolerance districts were created in total³ we get a file containing $2 \times 1774 = 3548$ rows with 0-tolerance election data⁴. Because it is somewhat easier to work with when later on we look at logit change for Danish party Det Radikale Venstre between 2001 and 2005, we want this long format to be transformed into wide format, i.e. to have party specific election data for each election serve as independent cases (i.e. each election year must have its own columns).

³ Remember that at the 2005 Danish National Parliamentary election there were 1855 polling districts in total. Thus, the 0-tolerance function has merged a large number of districts together because of polling district changes between 2001 and 2005.

⁴ A great way of checking whether the 0-tolerance computations are sound is to check if the number of voters in a 0-tolerance district agrees with the number of voters you find when adding the number of voters in each of the original polling districts that have been aggregated to create the 0-tolerance district in question (using e.g. the file "Valgdata_afstomr_FV2005".xls in O:\PEDA\GIS\Afstemningsområdekort i GIS\FV_2005).

Having done this, our next task consists in creating the above mentioned variable containing logit changes for Det Radikale Venstre in our Excel-file. In other words, we want to construct a variable showing the difference in logit-transformed vote percentage for Det Radikale Venstre in each 0-tolerance district between the 2001 and 2005 elections. This can also be written $\text{LNpB}(FV2005) - \text{LNpB}(FV2001)$ ⁵ which equals the natural logarithm to the odds ratio for Det Radikale Venstre between the two elections, i.e. $\text{LN} \left(\frac{pB(FV2005)}{pB(FV2001)} \right)$. Creating this variable is quite straightforward since our 0-tolerance Excel-file contains voting data for Det Radikale Venstre for both the 2001 and 2005 elections as well as data on the total number of valid votes in each 0-tolerance district for both elections. Thus, having made this variable showing the change in logittransformed party percentage for Det Radikale Venstre between the 2001 and 2005 Danish National Parliamentary elections, we are almost ready to import our data into ArcMap 9.3 for making the 0-tolerance shapefile. First however, we need to convert the Excel-file into an ArcMap-friendly file format. The dBase IV-file format is a good choice. Using the software StatTransfer makes this conversion operation quite an easy task to perform. When having converted our Excel-data into dBase-file format, we are now ready to launch ArcMap 9.3 and add both the relevant shapefile and our dBase-file into a new empty map.

3) Creating a 0-tolerance shapefile in ArcMap 9.3

Now the question is: *which shapefile should we use to construct our 0-tolerance map?* So far only our table data is 0-tolerance compatible. Our next task therefore consists in modifying one of our existing shapefiles of polling districts to make it resemble the 0-tolerance districts in our dBase-file named “0tolerance_2005reference_2001_opdateret”.dbf (in O:\PEDA\GIS\0-tolerancekort FV\0-tolerancekort FV2005ref_FV2001). Since we began this example by choosing the 2005 Danish National Parliamentary election as reference election for the 0-tolerance computations, we will use the 2005 polling districts shapefile named “AFSTEMOMR_FV_2005_afstemte_koder_etter_tilretning”.shp (in O:\PEDA\GIS\0-tolerancekort FV\0-tolerancekort FV2005ref_FV2001) as the one to modify.

⁵ Where LN stands for the natural logarithm and pB is the vote percentage for Det Radikale Venstre (B) of the total valid vote in each district.

A reasonable way of creating the 0-tolerance shapefile begins by joining our 0-tolerance election data to the 2005 shapefile. For a thorough guide on how to use ArcMap 9.3's "Join-function" consult the papers by Malene Rode Larsen "How To: Make a map displaying party support using ArcMap 9.3".doc and "How To: Prepare for, carry out and display a cluster analysis in ArcMap using the Ward's cluster analysis technique".doc (in O:\PEDA\GIS\ Introductory papers to ArcMap 9.3). Our unique key variable during the join-operation will be a variable which contains the unique district codes. In our 2005 shapefile's "Attribute Table" this variable is called "VALGSTEDID" whereas in the dBase-file we will use the variable "GRUPPE" containing the 0-tolerance district codes. The reason why we can use these two different variables to join on is that the 0-tolerance district codes have been given codes which are equal to some of the original 2005 district codes. For instance, every time two or more polling districts have been merged in the 0-tolerance computation the resulting 0-tolerance district will get one of the original polling district codes whereas the other original polling district codes will be left out in the 0-tolerance table. Figure 7 below provides an example. The two original 2005 polling districts "Rådhuset" (461023) and "Sct. Knuds Gymnasium" (461024) have been merged into a 0-tolerance district. This has got one of the original polling district codes, namely 461024, whereas the other original code, 461023, has been left out in the new "GRUPPE"-variable.

Figure 7: The dBase-file containing 0-tolerance districts with election data

OJD	VALGID	GRUPPE	VALGSTEDID	KREDSNR	STORKREDSII	LANDSDELSII	STEMMEBERE	AFGIVNE ST	BLAIKE STE
706	196	461017	461017	51	10	2	5122	4554	26
707	196	461018	461018	51	10	2	4524	4057	17
708	196	461019	461019	51	10	2	3114	2644	8
709	196	461020	461020	51	10	2	3624	3282	21
710	196	461021	461021	51	10	2	1614	1422	9
711	196	461022	461022	51	10	2	416	375	1
712	196	461024	461023;461024	52	10	2	13303	11408	84
713	196	461025	461025	52	10	2	5586	4746	27
714	196	461026	461026	52	10	2	8277	7185	37
715	196	461027	461027	52	10	2	364	328	0
716	196	461028	461028	52	10	2	3276	2917	12
717	196	461029	461029	52	10	2	4847	4193	23
718	196	461030	461030	52	10	2	6908	5907	29
719	196	461031	461031;461033	52	10	2	5657	5041	23
720	196	461032	461032	52	10	2	667	616	3
721	196	461034	461034	52	10	2	477	442	3
722	196	461035	461035	50	10	2	5511	3703	20

The original polling district 461023 no longer exists in the new 0-tolerance district variable ("GRUPPE"), it has been merged into district 461024.

Therefore, when joining the shapefile and the dBase-file there will naturally be some rows in the shapefile's "Attribute Table" which will contain "null" values instead of election data, since these are the rows containing the original district codes that have been left out in the 0-tolerance table. An illustration is given below for the district codes mentioned just above. As should be clear from looking at figure 8, no election data is available for the original polling district "Rådhuset" with district code 461023.

Figure 8: Some of the original district codes do not match with the new 0-tolerance codes, therefore no election data will get attached when joining.

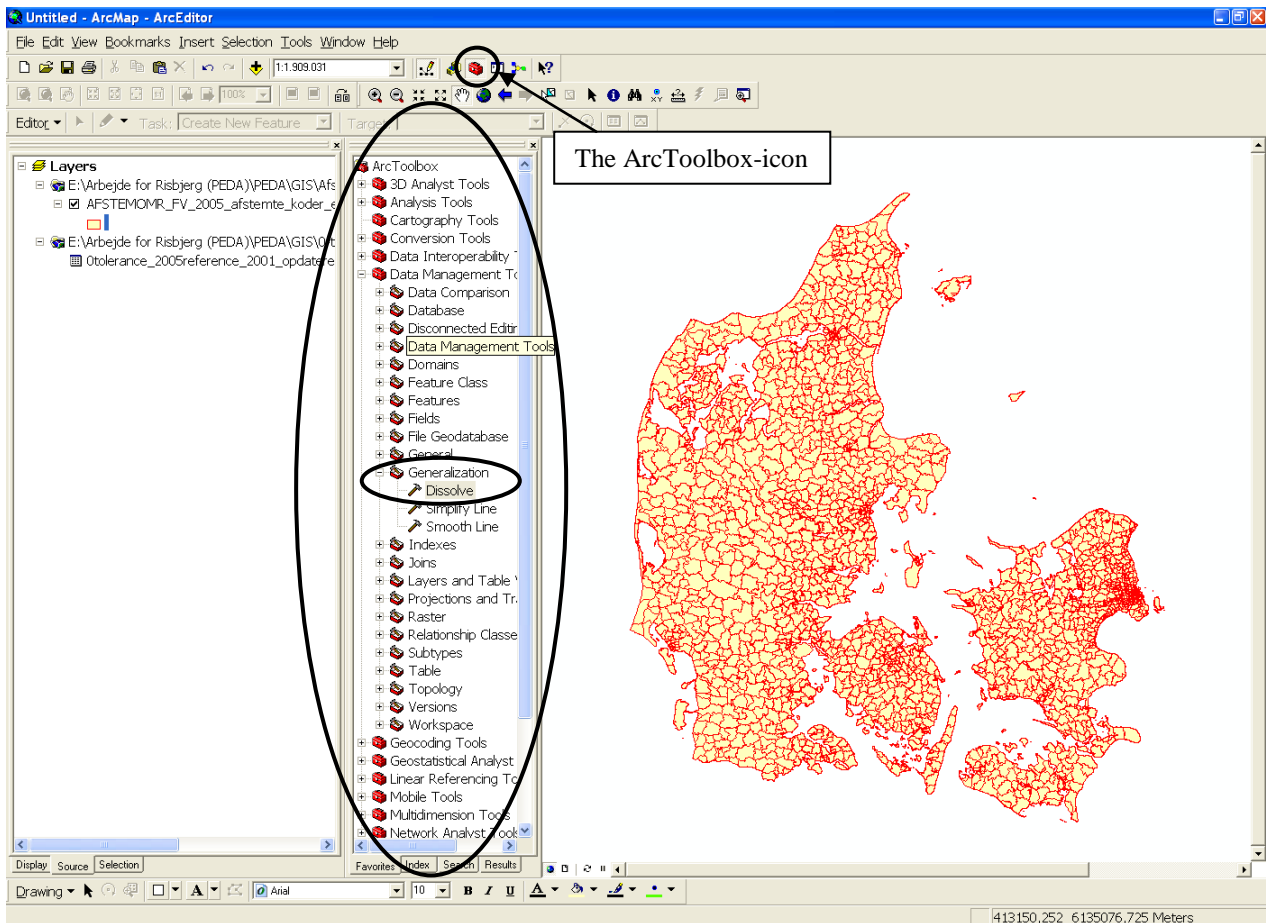
VALGSTEDID	VALGSTEDNA	STORKREDSH	LANDSDELSH	KREDSNUM	OID	VALGID	GRUPPE	VALGSTEDID	KREDSNR	STORKREDSH	LANDSDELSH	STEMMEBERE	AFGIVNE ST
461011	Jernbanegades Skole	10	2	51	700	196	461011	461011	51	10	2	6732	5657
461012	Højstrupskolen	10	2	51	701	196	461012	461012	51	10	2	8016	6506
461013	Lumby Forsamlingshus	10	2	51	702	196	461013	461013	51	10	2	628	733
461014	Slige Skole	10	2	51	703	196	461014	461014	51	10	2	4770	4071
461015	Allesø Forsamlingshus	10	2	51	704	196	461015	461015	51	10	2	327	307
461016	Spurvelundskolen	10	2	51	705	196	461016	461016	51	10	2	2223	1941
461017	Kroggårdskolen	10	2	51	706	196	461017	461017	51	10	2	5122	4554
461018	Tarup Skole	10	2	51	707	196	461018	461018	51	10	2	4524	4057
461019	Paarup Forsamlingshus	10	2	51	708	196	461019	461019	51	10	2	3114	2644
461020	Koruphallen	10	2	51	709	196	461020	461020	51	10	2	3624	3282
461021	Ubberudhallen	10	2	51	710	196	461021	461021	51	10	2	1614	1422
461022	Ravneshøj Forsamling	10	2	51	711	196	461022	461022	51	10	2	446	375
461023	Rådhuset	10	2	52	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
461024	Sct. Knuds Gymnasium	10	2	52	712	196	461024	461024	52	10	2	13303	11408
461025	Rosengårdskolen	10	2	52	713	196	461025	461025	52	10	2	5586	4746
461026	Sanderumhallen	10	2	52	714	196	461026	461026	52	10	2	8277	7185
461027	Brændekilde Forsamling	10	2	52	715	196	461027	461027	52	10	2	364	328
461028	Rasmus Rask Skolen	10	2	52	716	196	461028	461028	52	10	2	3276	2917
461029	Tinglækkeskolen	10	2	52	717	196	461029	461029	52	10	2	4847	4193
461030	Dalumskolen	10	2	52	718	196	461030	461030	52	10	2	6908	5907
461031	Stenløse Forsamlingshu	10	2	52	719	196	461031	461031,461033	52	10	2	5657	5041
461032	Fangel Forsamlingshus	10	2	52	720	196	461032	461032	52	10	2	667	616
461033	Højby Skole	10	2	52	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
461034	Tinglækkeskolen	10	2	52	721	196	461034	461034	52	10	2	477	442
461035	Humlehaveskolen / Æld	10	2	50	722	196	461035	461035	50	10	2	5511	3703
461036	Tornbjerg Gymnasium	10	2	50	723	196	461036	461036	50	10	2	3592	2990
471001	Hjaldstrup	10	2	55	724	196	471001	471001	55	10	2	570	515
471002	Krogsbølle	10	2	45	725	196	471002	471002	55	10	2	816	738
471002	Krogsbølle	10	2	55	725	196	471002	471002	55	10	2	816	738
471002	Krogsbølle	10	2	55	725	196	471002	471002	55	10	2	816	738
471002	Krogsbølle	10	2	55	725	196	471002	471002	55	10	2	816	738
471002	Krogsbølle	10	2	55	725	196	471002	471002	55	10	2	816	738
471002	Krogsbølle	10	2	55	725	196	471002	471002	55	10	2	816	738
471002	Krogsbølle	10	2	55	725	196	471002	471002	55	10	2	816	738
471002	Krogsbølle	10	2	55	725	196	471002	471002	55	10	2	816	738
471002	Krogsbølle	10	2	55	725	196	471002	471002	55	10	2	816	738
471002	Krogsbølle	10	2	55	725	196	471002	471002	55	10	2	816	738
471002	Krogsbølle	10	2	55	725	196	471002	471002	55	10	2	816	738

What we should do now is edit in the shapefile's "Attribute Table" such that every original polling district code contained in the "VALGSTEDID" variable (the left most variable in figure 8) that was left out following the 0-tolerance computations is assigned the same district code as the code belonging to the 0-tolerance district to which it has been merged. For instance, the original polling district "Rådhuset" with district code 461023 should be assigned the new 0-tolerance code 461024 (and probably also its original name Sct. Knuds Gymnasium just to avoid confusion), such that the geography of these two original polling districts can subsequently be merged in ArcMap creating the new 0-tolerance district. Editing in the shapefile's original attribute data is done by using

ArcMap 9.3's "Edit"-tool from within the "Editor"-toolbar. For a detailed description on (in O:\PEDA\GIS\ Introductory papers to ArcMap 9.3) to proceed, consult the paper "How To Edit existing shapefiles using ArcMap 9.3".doc by Malene Rode Larsen. Whenever this district code-editing is done we should end up having the exact same 0-tolerance district codes as in our dBase-file, namely 1774 unique codes⁶. However, our 2005 shapefile still consists of the 1855 original geographic areas. To finally create a shapefile showing the 1774 0-tolerance districts only we will make use of one of ArcMap's most distinctive features, that is, the "Dissolve"-function. Using the "Dissolve"-tool, every feature that has the same characteristic (in this case the same 0-tolerance district code) are aggregated into one larger area as long as these areas share a common border, i.e. are geographic neighbors. In our case this logically must be the case since the 0-tolerance districts naturally consist of smaller neighboring polling districts. To get to the "Dissolve"-tool you should select the ArcToolbox-icon depicting a red box positioned in the main panel of your screen, see figure 9 below. When clicking the ArcToolbox, a new window will emerge on your screen between your "Table of Contents" and your "Data window", see figure 9 below. The ArcToolbox contains numerous possibilities for manipulating shapefiles and their attribute data in ArcMap. In our case, however, we are only interested in finding the "Dissolve"-tool. We select first the "Data Management Tools"-option and next the "Generalization"-sub-option and from here we can choose the "Dissolve"-tool.

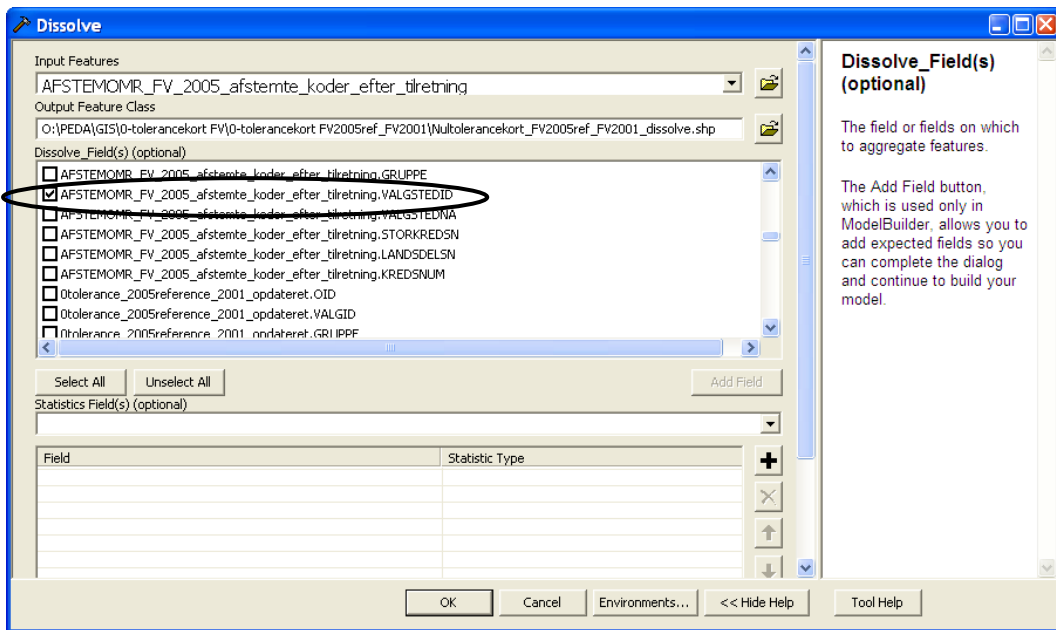
⁶ However, there are many more codes (rows) in the shapefile (3501 in total, see figure 8 above) since at this time quite a few district codes have doublets or triplets, which is partly a result of the editing.

Figure 9: Finding the “Dissolve”-function from within the ArcToolbox–window



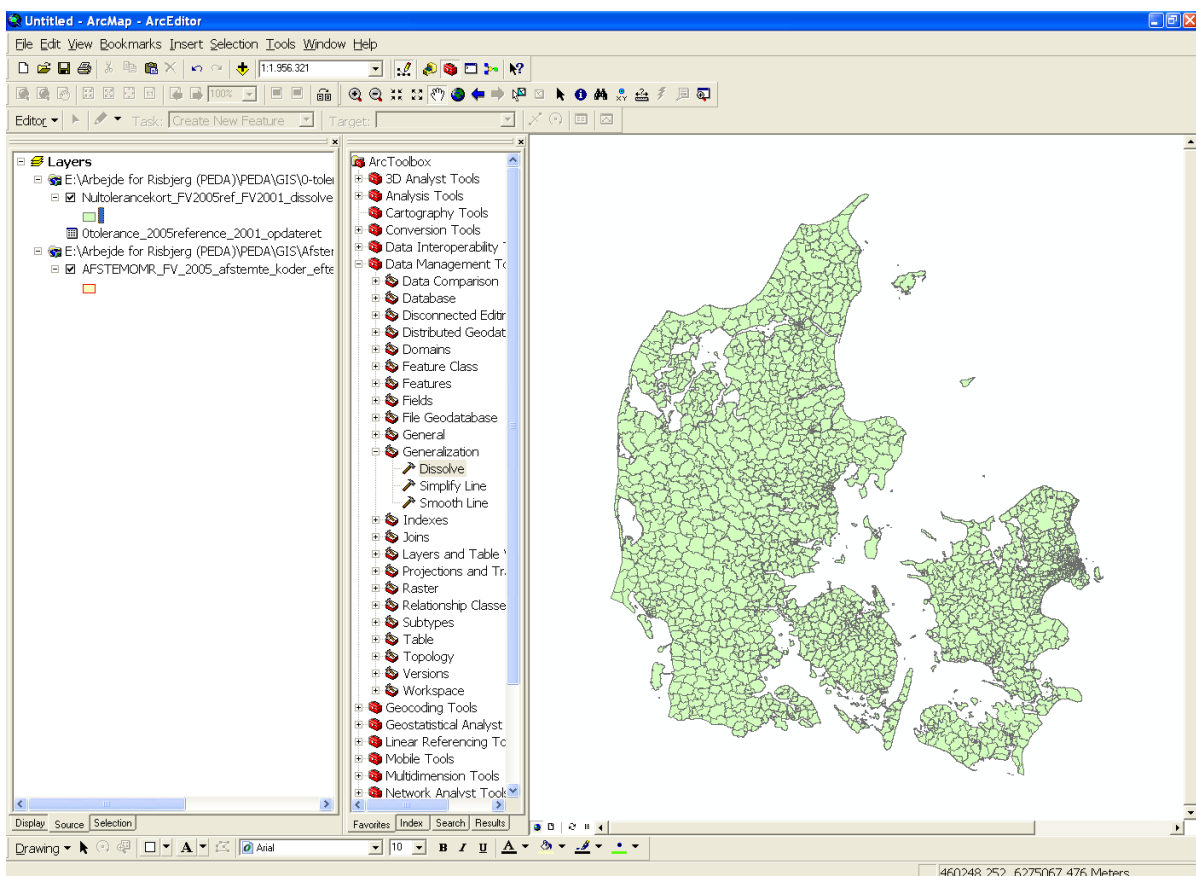
When double-clicking the “Dissolve”-tool, the following “Dissolve”-window will pop up on your screen, see figure 10 below. You should fill it out as done below; 1) choosing which shapefile you want to dissolve on (in our case the 2005 shapefile with the edited “VALGSTEDID”-variable), 2) choosing a name for your dissolved shapefile (we choose to call it “Nultolerancekort_FV2005ref_FV2001_dissolve”) and 3) clicking the field which you want to dissolve/aggregate on (in our case the edited “VALGSTEDID”-variable now containing the 0-tolerance district codes only).

Figure 10: The “Dissolve”-window



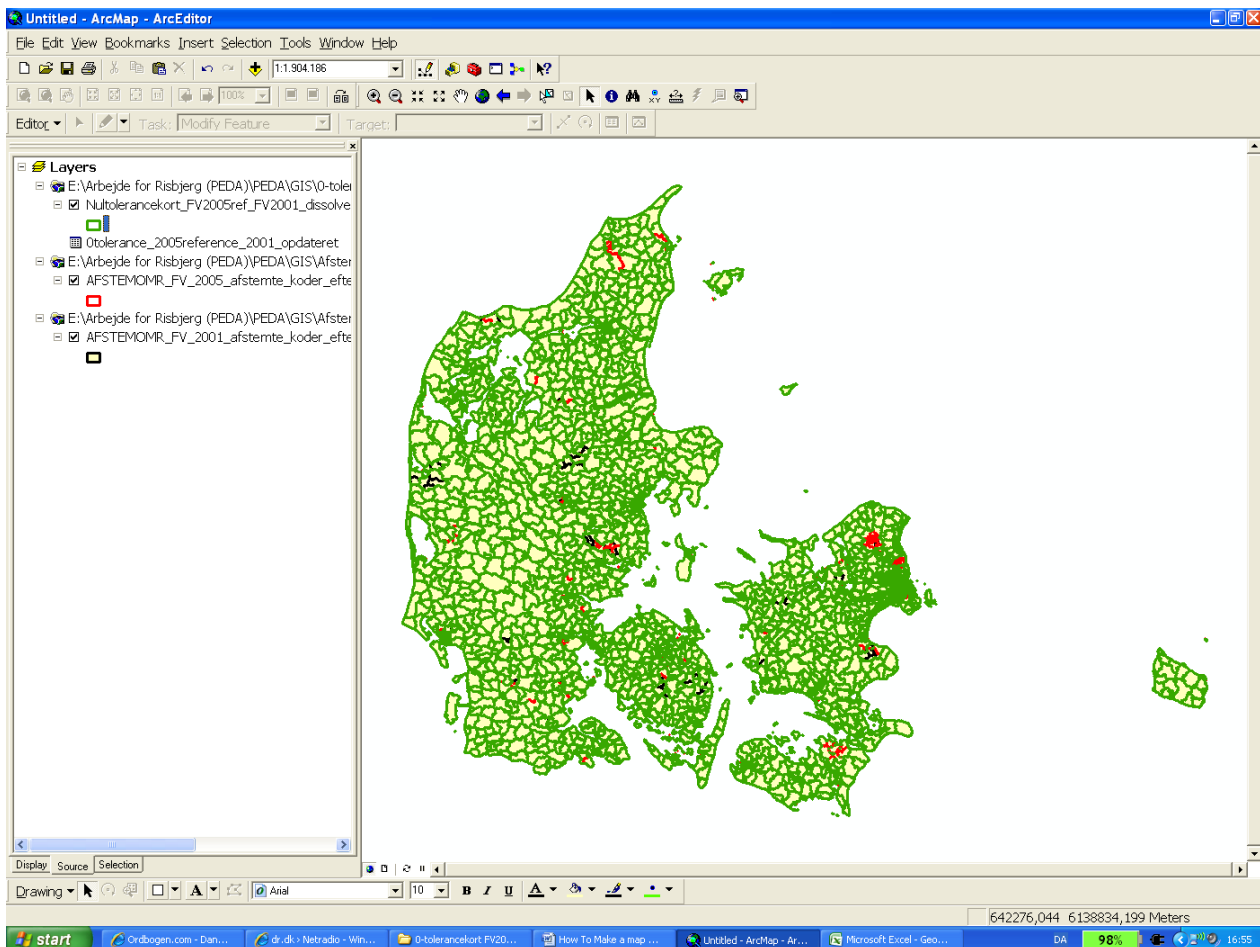
The resulting 0-tolerance shapefile “Nultolerancekort_FV2005ref_FV2001_dissolve.”shp looks like this:

Figure 11: The 0-tolerance shapefile



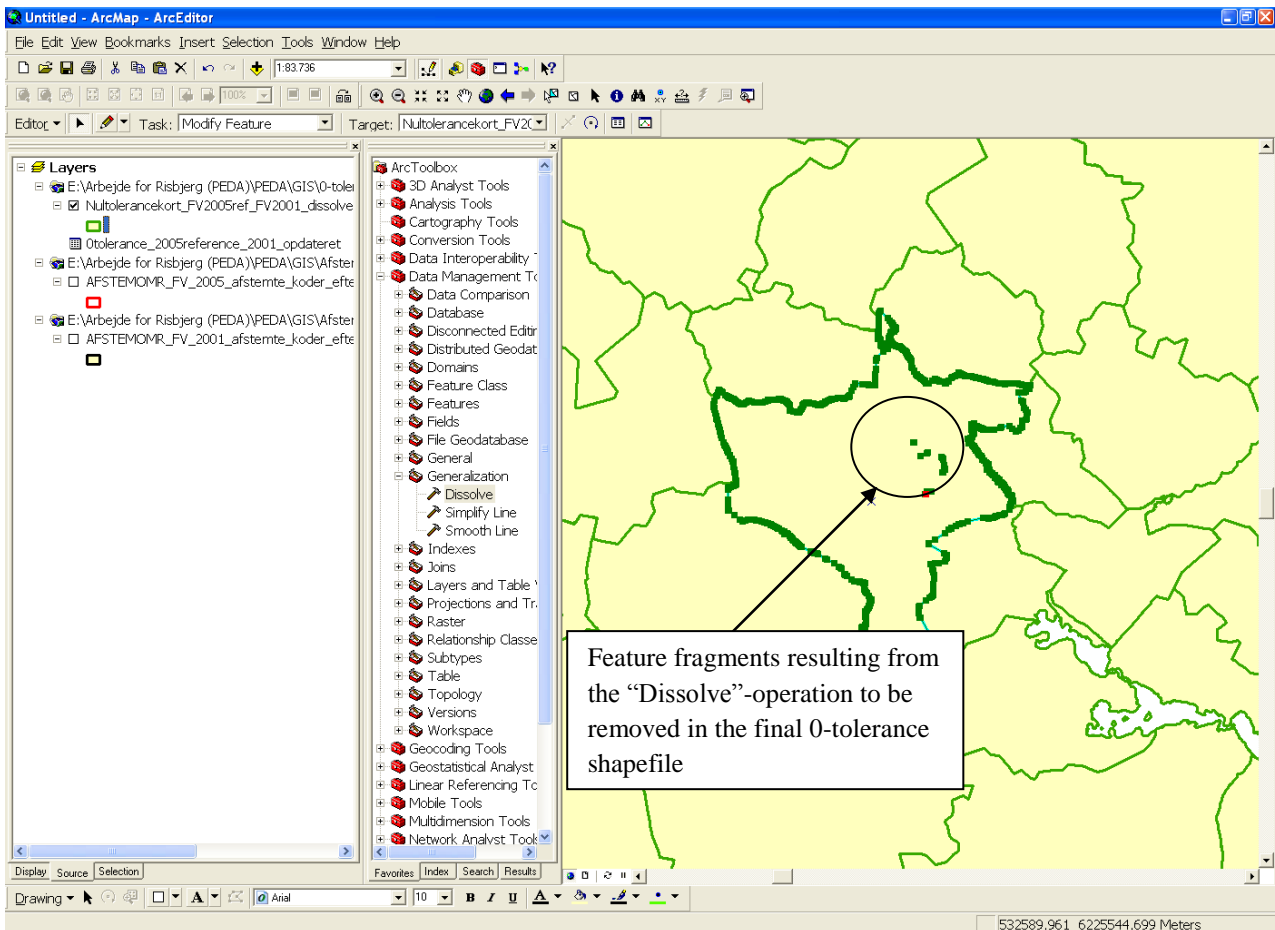
It might be difficult at first to see any obvious changes between our new 0-tolerance shapefile containing the 1774 0-tolerance districts and the original 2005-shapefile containing the 1855 polling districts. To make a better inspection of the changes made it is fruitful to add to the data frame the shapefile containing the polling districts of the 2001 Danish National Parliamentary election as well. The 2001-shapefile should lie in the bottom of the three shapefiles, the 2005-shapefile in the middle and on top we should have the 0-tolerance shapefile we have just created. Furthermore, the two topmost layers should be made transparent and each should be given a different outline colour for their borders. Doing this will result in the following, see figure 12 below. Here it is much easier to see which of the original 2001 and 2005 polling districts have been aggregated to 0-tolerance districts to make them comparable.

Figure 12: The 0-tolerance shapefile with the 2001- and 2005-shapefiles lying underneath making it possible to see the 0-tolerance aggregations



Actually, we might not be completely done yet. This is because sometimes, when we use the “Dissolve”-tool, small fragments of the original features will not have been removed as they should have been. An example of this is provided in figure 12 below where we use the “Edit”-tool together with the “Modify Feature”-function from inside the “Editor”-toolbar to remove unwanted fragments from the dissolved feature in question (on how to find and use these ArcMap-features consult the paper “How To Edit existing shapefiles using ArcMap 9.3”.doc by Malene Rode Larsen (in O:\PEDA\GIS\ Introductory papers to ArcMap 9.3)). If you have the time, you should go through the whole shapefile to see if this type of error has occurred in other areas where features have been dissolved.

Figure 12: The “Dissolve”-operation has left small fragments of features which need to be removed

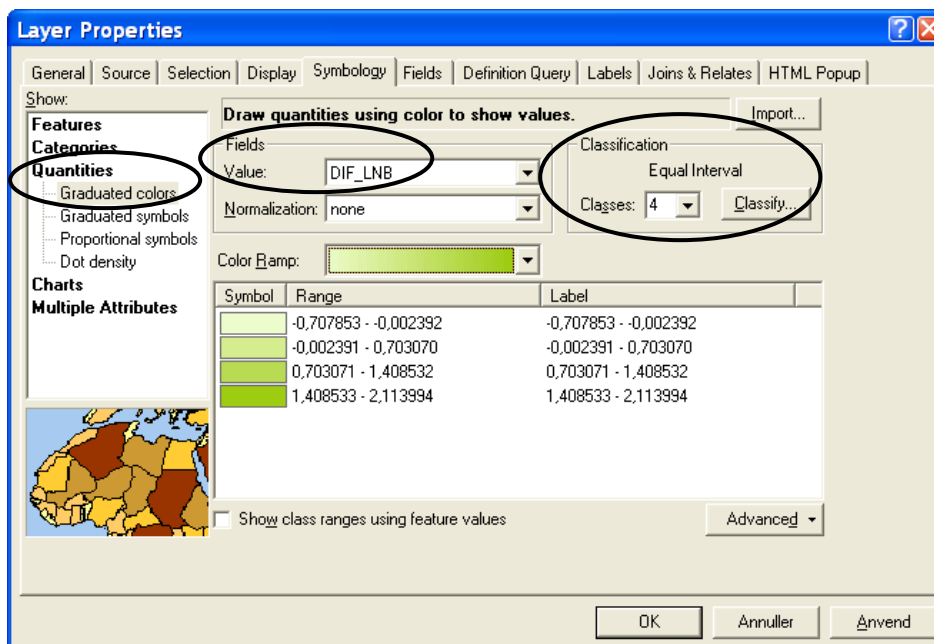


4) Making a map of the natural logarithm to the odds ratio for the Danish party Det Radikale Venstre in every 0-tolerance district between the 2001 and 2005 Danish National Parliamentary elections

Finally, having made the 0-tolerance shapefile between the 2001 and 2005 Danish National Parliamentary elections using the 2005 election as reference election for the 0-tolerance computations, we will use this shapefile to make a map depicting the natural logarithm to the odds ratio for the Danish party Det Radikale Venstre in every 0-tolerance district between the two above mentioned elections. As explained in section 2 we have already created this variable and joined it to the original 2005 polling district shapefile. However, unless otherwise indicated in the “Dissolve”-window, the attribute table of our newly dissolved shapefile “Nultolerancekort_ FV2005ref_ FV2001_ dissolve”, will, apart from the permanent “FID”- and “Shape”-attributes, contain only the variable on which the dissolve was carried out, i.e. the 0-tolerance district variable. In other words, we must once again join our election data in the dBase-file to our shapefile.

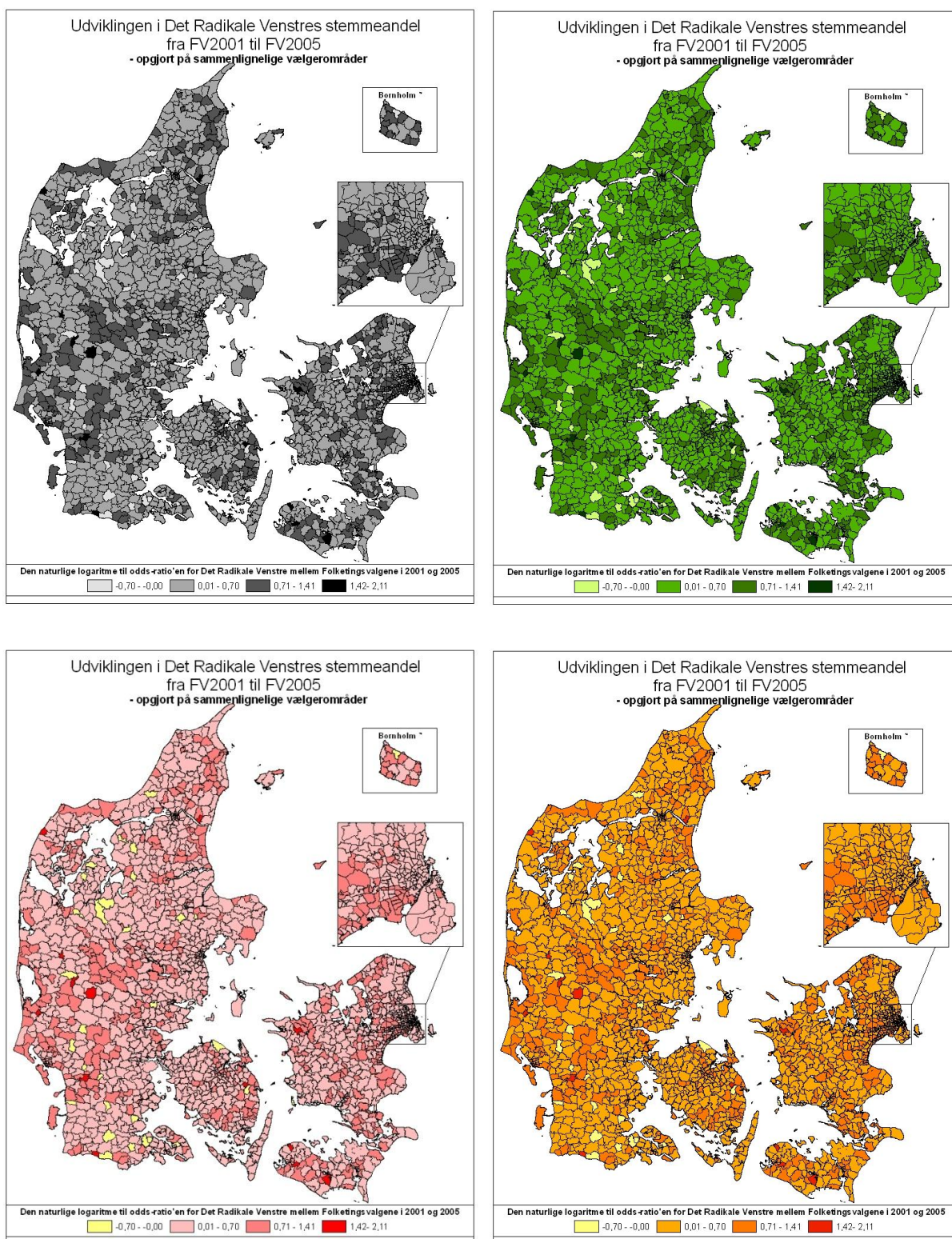
We are now finally ready to make a map showing our variable of interest, that is, the natural logarithm to the odds ratio for Det Radikale Venstre. In accordance with the paper “How To: Colour geographic features and different types of attributes using ArcMap 9.3”.doc by Malene Rode Larsen (in O:\PEDA\GIS\ Introductory papers to ArcMap 9.3) many colouring methods as well as colouring schemes seem available to us for pursuing this goal. We choose here to make a map based on the following alternatives in the “Symbology”-tab from inside the “Layer Properties”-window, see figure 13 on page 16. That is, we choose the “Quantities – Graduated Colors”-option with four equal intervals.

Figure 13: Choosing an appropriate look for the final map inside the “Symbology”-tab from within the “Layer Properties”-window



The following figures (figure 14a-d) show 0-tolerance maps of the logarithm to the odds ratio for the Det Radikale Venstre using four different colouring schemes. The first interval containing only negative numbers indicates where the logit-transformed vote percentage for Det Radikale Venstre went down from the November 2001 to the February 2005 election, whereas the three intervals containing only positive numbers indicate the strength of the progress in vote percentage for Det Radikale Venstre between these two elections. From the maps below it is evident that Det Radikale Venstre had a very successful election in 2005. The party's share of the total valid vote increased in almost all of the 1774 comparable 0-tolerance districts.

Figure 14a-c: Maps showing the natural logarithm to the odds ratio for the Danish party Det Radikale Venstre in every 0- tolerance district



Malene Rode Larsen, February 2010

References:

- Larsen, Malene Rode: “How To: Colour geographic features and different types of attributes using ArcMap 9.3”.doc (O:\PEDA\GIS\Introductory papers to ArcMap 9.3)
- Larsen, Malene Rode: “How To Edit existing shapefiles using ArcMap 9.3”.doc (O:\PEDA\GIS\Introductory papers to ArcMap 9.3)
- Larsen, Malene Rode: “How To: Make a map displaying party support using ArcMap 9.3” (O:\PEDA\GIS\Introductory papers to ArcMap 9.3)
- Larsen, Malene Rode: “How To: Prepare for, carry out and display a cluster analysis in ArcMap using the Ward’s cluster analysis technique” (O:\PEDA\GIS\ Introductory papers to ArcMap 9.3)