# How To (1): Produce a Display of Some Interesting Data Attributes with a Spatial Distribution

This short paper provides a quick step-by-step guide to displaying desired attributes of geographical data (in this instance politically stable regions of California) with ArcGIS 9.1. By the end of this tutorial You should be able to add data to an empty screen, join table-attributes to geographic boundary files, and create a simple choropleth map – that is, a map using different colors or shadings to show varying values or classes over space – useful for suggesting any spatial patterns in your data. Also, You should know how to best incorporate this map into a text document.

Example Research Objective: To establish whether the counties of California do indeed cluster according to some predefined classification of politically stable regions. In other words, has this procedure actually identified coherent (i.e. contiguous) regions?

Modus Operandi: Fire up ArcGIS 9.1 by choosing the ArcMap icon from where you access the program package on your computer and select to start with a new empty map. These few actions should present you with the following screen:
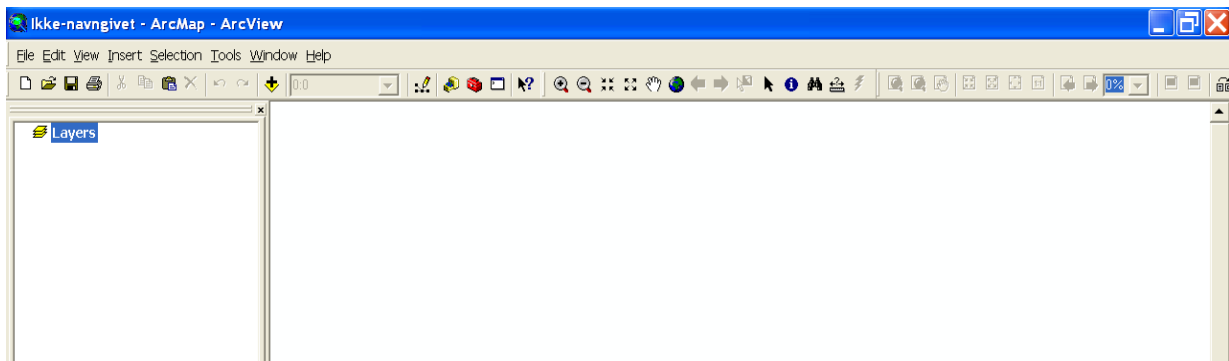
Here the Tools and Layout toolbars have been repositioned above the main panel; if the program has never been used before you may choose to do the same or leave them floating (else the program will start up with the settings of the latest client, but you can always find any toolbar through the View\ Toolbars drop-bars from the menu bar at the top).

Next, take the following steps to finally produce a choropleth map of the presumed politically stable regions of California.

1) To add data to your map, either select Add Data through the File drop-bar from the menu bar at the top or just press the Add Data icon (depicting a black addition sign on a yellow tilted square) next to the Map Scale field above the main panel. From the new pop-up screen navigate to "O:\Kristian (2005)\ROAD\californien\californien boundary files" and select "caliroadcounties.shp" as shown in the picture below:
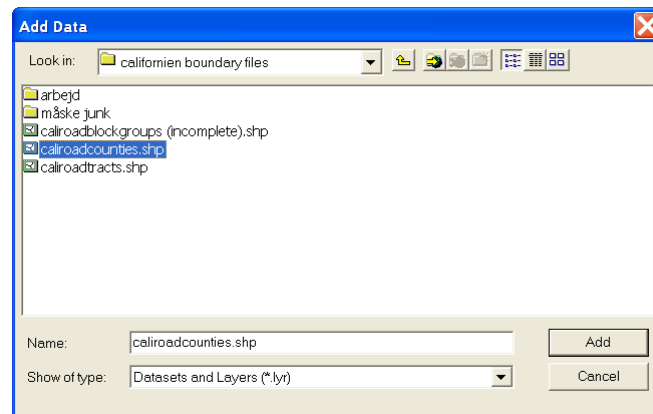
If unable to immediately access the O-drive, you will need to establish a connection in the program's file ordering system first by pressing the Connect To Folder icon in the pop-up screen (depicting a yellow arrow pointing to a globe) and choosing the appropriate folder. Note that if all your data is stored in the same folder you would naturally choose to connect to that particular folder rather than specifying a more general connection such as the O-drive or My Computer[1]. Now add the data (ignore for present purposes any potential warnings of a lack of spatial reference information).

Since the selected data is a geographic boundary file, it will appear both as a layer under the active data frame in the table of contents to the left and as a graphic representation of the Californian counties in the main panel. Toggle the layer on and off by use of the check box in the table of contents. You may also at any time shift between the data view, which you are currently in, and the layout view by selecting either view through the File drop-bar from the menu bar at the top, or by pressing the respective icons right below the main panel (depicting a globe and a sheet of paper). To add more data to the active data frame, that is, additional boundary files or tables, press the Add Data icon again. To add another data frame and thus generate an additional picture frame in the layout view select

---

[1] The program's file ordering system is really a feature of ArcCatalog, which may be accessed separately for large-scale data management or to edit the range of folders connected to once unhandy to wield or obsolete.

Data Frame through the Insert drop-bar from the menu bar (before adding data to this new data frame you would need to activate it by right-clicking on the data frame and select Activate; this would switch the display in the main panel).

2) At the moment, the graphic representation of the Californian counties shows all features with a single random color, but we want to ultimately color each county according to some proposed regional scheme of political stability. In order to do this, each county must have been assigned a value identifying its individual class. This information may have to be contrived from somewhere else, hence, open the attribute table of the layer by right-clicking "caliroadcounties" 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 not readily available:

FID and Shape* are attributes automatically ascribed to every shape file (i.e. every geographic data file) by ArcGIS. FID is the unique Feature IDentification number, which runs from zero through the last feature, here county, and Shape* denotes the geographical type of each such feature (other possibilities besides polygon are point and line, however, counties are inherently polygons). Neither variable can be deleted nor changed. COUNTY and STATE (perhaps FIRST_STAT) are part of another identification system, namely the FIPS, or Federal Information Processing Standards, codes, in combination uniquely identifying any county in the United States (but as we have data for one state only, COUNTY is the only relevant variable right now).

Consequently, more data must be added to the data frame and joined with the existing attribute table. ArcGIS allows for the addition of tables with no graphic representation if just these data files are in the dBASE format, and both SPSS and Excel will easily convert their files into this format (though note that the procedure does have some quirks with regards to what information is stored). So start SPSS, navigate to "O:\Kristian (2005)\ROAD\californien\californien data", and select "clusters.sav"; then simply save the file as a dBASE file type. If pressing the Add Data icon in ArcGIS you should now be able to find the table in the folder where you chose to save it, and add the data to your screen.

Since the selected data does not have a graphic representation by itself, it will only appear in the table of contents to the right, and only when the tab titled Source is at the front. Opening the new table in the same way as the county attribute table before, you are presented with this screen:

Attributes of clusters

OID	COUNTY	CLU12_1	CLU11_1	CLU10_1	CLU9_1	CLU8_1	CLU7_1	CLU6_1	CLU5_1	CLU4_1	CLU3_1	CLU2_1
0	1	1	1	1	1	1	1	1	1	1	1	1
1	3	2	2	2	2	2	2	2	2	2	2	1
2	5	3	3	3	3	3	3	3	3	3	3	2
3	7	1	1	1	1	1	1	1	1	1	1	1
4	9	3	3	3	3	3	3	3	3	3	3	2
5	11	4	4	4	4	4	4	4	4	3	3	2
6	13	1	1	1	1	1	1	1	1	1	1	1
7	15	5	5	5	5	5	5	5	1	1	1	1
8	17	3	3	3	3	3	3	3	3	3	3	2
9	19	6	6	6	6	6	4	4	4	3	3	2
10	21	4	4	4	4	4	4	4	4	3	3	2
11	23	7	7	7	7	7	6	6	5	4	3	2
12	25	8	8	8	5	5	5	5	1	1	1	1
13	27	7	7	7	7	7	6	6	5	4	3	2
14	29	6	6	6	6	6	4	4	4	3	3	2
15	31	6	6	6	6	6	4	4	4	3	3	2
16	33	1	1	1	1	1	1	1	1	1	1	1
17	35	9	9	9	8	7	6	6	5	4	3	2
18	37	8	8	8	5	5	5	5	1	1	1	1
19	39	6	6	6	6	6	4	4	4	3	3	2

Record: 1   Show: All | Selected   Records (0 out of 58 Selected.)   Options

Again, the OID (Object IDentification number) attribute is automatically ascribed to every non-geographic data file by ArcGIS and like the FID cannot be deleted or changed. COUNTY is identical to the COUNTY attribute of the boundary file and thus provides an essential link between entries in the two data files. The remaining variables describe the proposed regional groupings ranging from two to twelve clusters.

To join these classification schemes to the boundary file, right-click on the layer or table you want to join *to* in the table of contents – in this case the county layer – and select Join from the Joins and Relates drop-bar, which opens the Join Data Wizard. In it, first,

choose COUNTY as the field on which to match the two data files; secondly, choose the table you wish to append to the primary layer (in the current example there is only one option); and finally, also choose COUNTY as the field on which to match the two data files in the secondary table. Keep in mind that in principle field names do not matter as you could choose to join COUNTY to STATE – only that attributes are unique and identical (which would obviously not be the case in this particular instance[2]). The resulting attribute table should look something like this:

caliroadcounties.COUNTY	caliroadcounties.FIRST_STAT	clusters.OID	clusters.COUNTY	clusters.CLU12_1	clusters.CLU11_1	clusters.CLU10
1	6	0	1	1	1	
3	6	1	3	2	2	
5	6	2	5	3	3	
7	6	3	7	1	1	
9	6	4	9	3	3	
11	6	5	11	4	4	
13	6	6	13	1	1	
15	6	7	15	5	5	
17	6	8	17	3	3	
19	6	9	19	6	6	
21	6	10	21	4	4	
23	6	11	23	7	7	
25	6	12	25	8	8	
27	6	13	27	7	7	
29	6	14	29	6	6	
31	6	15	31	6	6	
33	6	16	33	1	1	
35	6	17	35	9	9	
37	6	18	37	8	8	

Record: 1   Show: All Selected   Records (0 out of 58 Selected.)   Options

You will find that all the required information is now accessible in one table marked by the title of the original file and a field name. However, the join is not permanent and may be removed by right-clicking the county layer and selecting Remove Join(s) from the Joins and Relates drop-bar. To make the appendage permanent (which you need not do at this point) you would right-click the county layer, select Export Data… from the Data drop-bar, and save the layer under a new file name.

3) Now it has become possible to color each county according to the desired attribute, specifically one or other of the proposed regional groupings. Right-click the county layer and select Properties…, or just double-click the county layer, to access the layer properties, and bring the Symbology tab to the front. Next, click on Categories in the panel to the left. You will see that currently all values have been assigned the same color (the color the counties have on your screen at present). We want to give them a number of different colors equivalent to the desired number of clusters, let's say four. Thus, select the field name xxx.CLU4_1 in the Value Field drop-bar at center and press Add All

---

[2] However, ArcGIS would try to do it anyways, so be cautious and aware of sudden null values! As well, if unsure about the quality of your link you may try to make the join both ways.

Values below the main panel. This action results in four random, but distinct colors being drawn from the active color ramp and assigned to the four existing values, that is, potential clusters. For instance it could look like this:
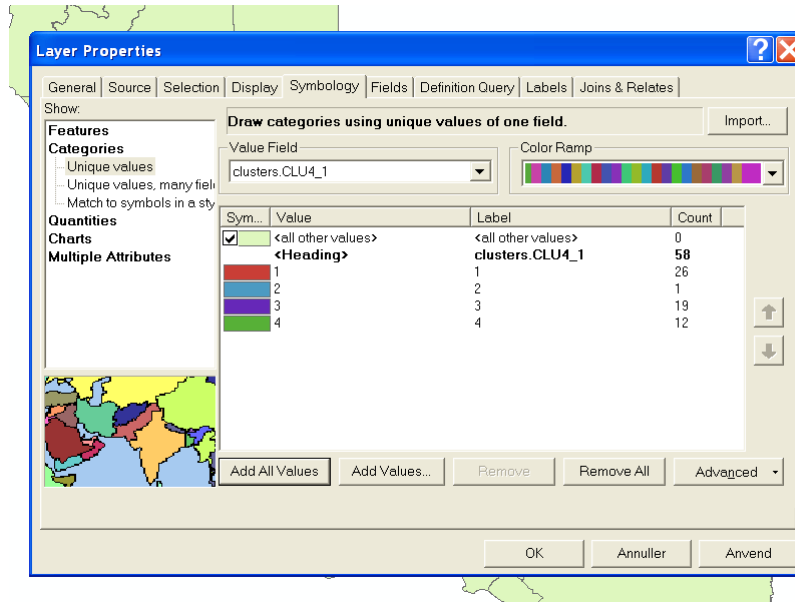
It is possible to manually choose any colors as well by double-clicking on the separate color fields in the main panel. Once satisfied, pressing OK then provides the corresponding display of Californian counties:
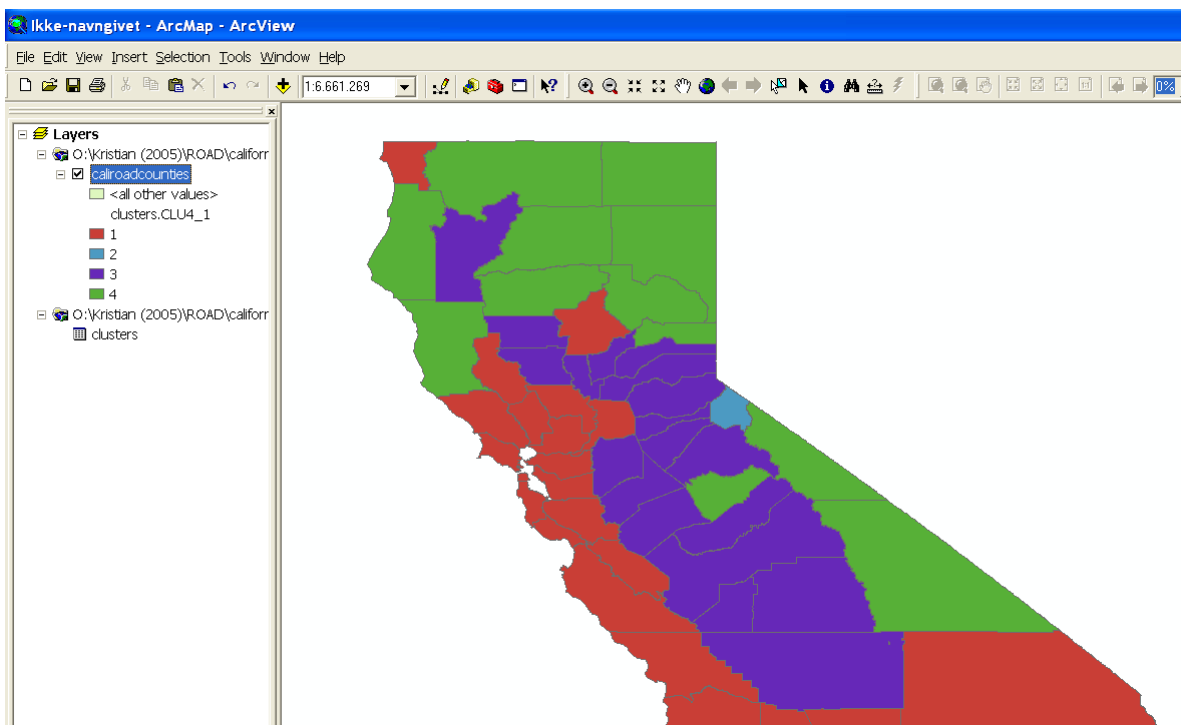
Toggle off the <all other values> option by use of the check box in the layer properties symbology tab if you do not want it to show in the table of contents. Click any of the colored fields in the table of contents to directly change a displeasing color. Note that neither of these color changes will affect the data files either, only the current graphic representation of the data, which may be saved to a map file through the File drop-bar from the menu bar at the top or by pressing the usual Save icon (depicting a disk).

Scaled attributes are displayed in similar fashion by clicking on Quantities rather than Categories in the layer properties symbology tab, choosing an appropriate color ramp, and the proper number of intervals (different classification options are offered such as natural breaks, equal interval, defined interval, quantile, and standard deviation). You may also choose to normalize the desired attribute values by any other attribute.

4) Switching to the layout view by pressing the Layout View icon (depicting a sheet of paper) right below the main panel you are able to see- and adjust how the choropleth map will look in print. Since you would typically want to incorporate such a map in a text document of some sort, a couple of considerations should be observed. By far the best way to incorporate a map from ArcGIS is as a picture file, which is accomplished by first selecting Export Map… through the File drop-bar in ArcGIS (choose your own preferred graphic file type, but probably the most widely used formats are JPEG, BMP, and GIF) and then selecting From File… through the Insert/ Picture drop-bars in Word or any analogous motions for inserting pictures in other text processing programs. That being the case, free margins should be minimized in the layout view to maximize map space in the picture (alternatively, if using Word, you may crop the picture afterwards through the Picture toolbar, but the result will not be as nice). Further, any numbers and text added to the layout should be a little larger than the employed point size of your text document as numbers and text will decrease somewhat when the picture is inserted within the set margins of the document.

To add a title or some text explanation and emphasis select Title or Text through the Insert drop-bar from the menu bar at the top. Once inserted, double-click on the text box to access the text properties from where pressing the Change Symbol button will allow

you to choose font and point size (note the conventional map hierarchies of text in the scroll panel). Moreover, pressing the Properties button from there presents you with additional usable options, particularly the choice of tilting your text by typing in a value in the Angle field. To add a ledger of the applied symbols, i.e. colors, shown in the table of contents, select Legend through the Insert drop-bar. This will open the Legend Wizard (it's a bit tricky, but play with it, and it will do wonders; use Preview often!). You may replace the value of each color with a name by either clicking twice (not double-clicking) on each value in the table of contents in turn, or by double-clicking on any value in the table of contents and rewriting them all in the Label section of the main panel. Finally, you may set guides to align text and data frames by clicking anywhere on the rulers and dragging them to their desired position.

Exercises:

- Redo the map of Californian counties to show 3 and 5 potential clusters.
- Create a layout comparing three different classification schemes to show the effect of additional regions on the overall coherence of the proposed political regions. Try to maintain the colors of the initial regions as best possible (TIP: you will need to establish three data frames; further you may add the same data to several data frames as these frames are merely graphic representations of the data).
- Finally, add a reference map to the layout by adding three other layers besides the county layer – representing urban areas, primary metropolitan- and metropolitan statistical areas – to be found at O:\Kristian (2005)\ROAD\californien\census tiger\1990 census[3] (TIP: the ordering of the layers may be determined by dragging the layers up and down in the table of contents; label the urban areas by right-clicking on the layer in the table of contents and selecting Label Features from the drop-bar).

Kristian (2005)

---

[3] Persons of an inquisitive mind may read the precise definitions of these categories in
O:\Kristian (2005)\ROAD\census tiger metadata.doc.