Tutorial 4 - Attribute data in ArcGIS

COPY the Lab4 archive to your server folder and unpack it.

The objectives of this tutorial include:

- Understand how ArcGIS stores and structures attribute data
- Learn to how to manipulate attribute data
- Learn how to "join" a data table to a data layer
- Learn how to construct a map with multiple dataframes
- Learn how to "relate" a data table to a data layer
- Learn how to download and use census data in ArcGIS

1. Viewing attribute data and changing the way the data is presented

Open ArcMap and add the QuakeHistory shapefile that you unpacked from Lab4.zip. The quake history shapefile contains historical earthquake information for the United States. Also note how the data are projected (North America Equidistant Conic). Add a map of the United States (states), Canada (province), Mexico (states), and the world (country) from the ESRI dataset and place the QuakeHistory map on top and the country layer on the bottom so that you can browse the data (Do you see that the U.S. map was "projected on-the-fly" because the projected QuakeHistory file was already there?).

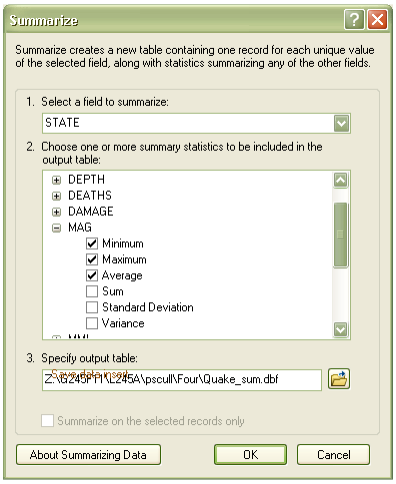
You likely just received an error message regarding coordinate systems. Is this the kind of error that should stop you in your tracks or can you ignore it?

Open the attribute table for QuakeHistory. You can change the way your data are organized by sorting on different variables. Right click on the 'mag' and select 'sort descending'. Which state or location had the highest magnitude earthquake according to these data? What type of cartographic symbol do you think would be best to portray the data? Why have there never been earthquakes in Southern Mexico or Alberta, yet so many in the US and some random Caribbean Island (hint: what island is it)?

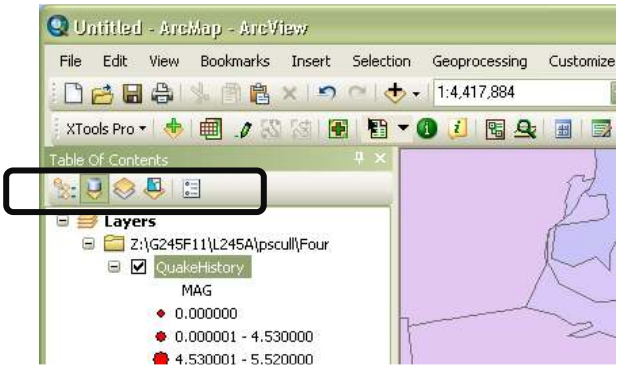
- Where is the location of the 13th largest quake? (And, isn't that odd?)

2. Summarizing attribute table information

Open the QuakeHistory attribute table, right click on the 'states' field (column) and select 'summarize'. This feature allows you to create a separate table from the QuakeHistory attribute information that summarizes different elements of QuakeHistory. Within the dialog box click next to 'MAG' (literally on the plus) and then select Average, Minimum and Maximum. Save the file as a .dbf file. What type of file is a .dbf? Finally, pay attention to the path of the output table to make sure it is written in your workspace (see below).



After you click 'ok' you will be prompted to add the table to your ArcMap view - select yes. Once this new table is created and added to the table of contents, 'open' it to view its contents (like you would any table associated with a data layer). Note that this new table looks identical to an attribute table for a shapefile. Once added, your table of contents will switch to the 'list by source view' as shown below (all the various view options are circled, mouse over each to identify them). This is because a table is inherently non-spatial and it cannot be clicked off and on the same in the same manner as a map. If you click over to the 'list by drawing order' view, you will see that it is not listed.



You should be able to answer the following questions of these data based on the summary you just performed and the general knowledge you have accumulated thus far this semester. If you cannot, please ask your lab instructor for help.

- Which state has the highest average magnitude earthquake?
- Which state experienced the most earthquakes according to these data?
- Which state experienced the largest earthquake on record (according to these data)?

TIP: When the table of contents is in 'list by source' mode you cannot re-arrange the order of maps. This can be easy to forget.

3. Adding a table to your map

Tables can be added to ArcMap in the same manner as individual maps. While they do not appear in the data viewer since they are not necessarily spatial data, they are accessible in the table of contents and can be manipulated in various ways. Once joined to a spatial data set using a unique identifier the contents of a table can be mapped similar to any other attribute data.

Open a new ArcMap document and add the SD_Counties map from your workspace (this is one of the files you downloaded from the server). By now you can probably tell that the map is not properly projected (which should bother your cartographic instinct).

Project the dataframe with the following system parameters: State Plane, NAD 1983 (US feet), South Dakota (South), FIPS 4002. Do you see where that projection is located? Don't confuse it with similar other ones.

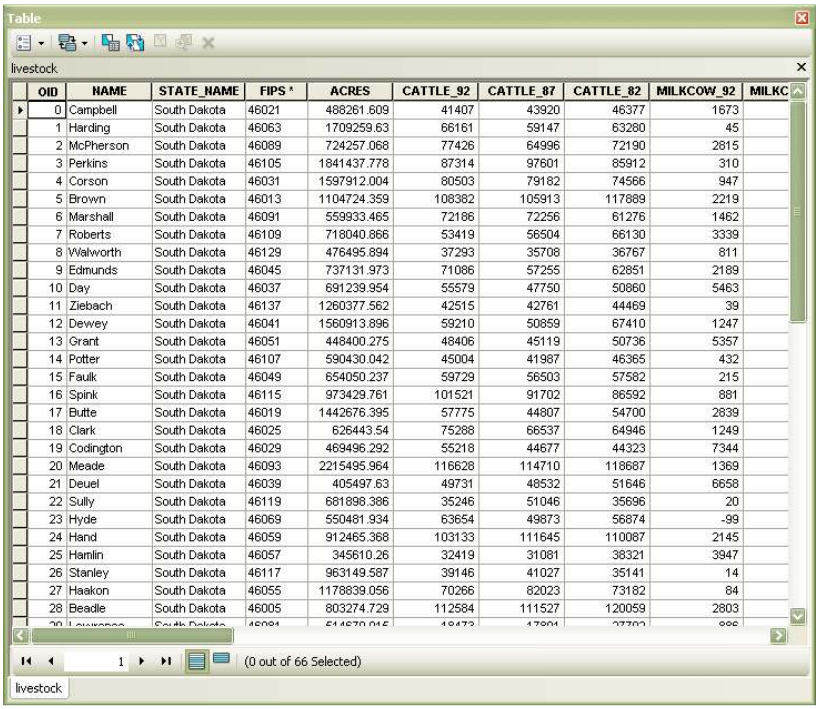
What is the difference between this projection transformation and the use of the "project" tool in the toolbox? If you don't understand ask for help.

Click on the add data button and add the file 'livestock.dbf'. These are tabulated livestock data for each county in South Dakota. Once added, your table of contents will switch to the 'list by source' mode as noted before, if it is not already.

Tip: Tables can be removed by right clicking on them and selecting 'remove.'

4. Opening a table

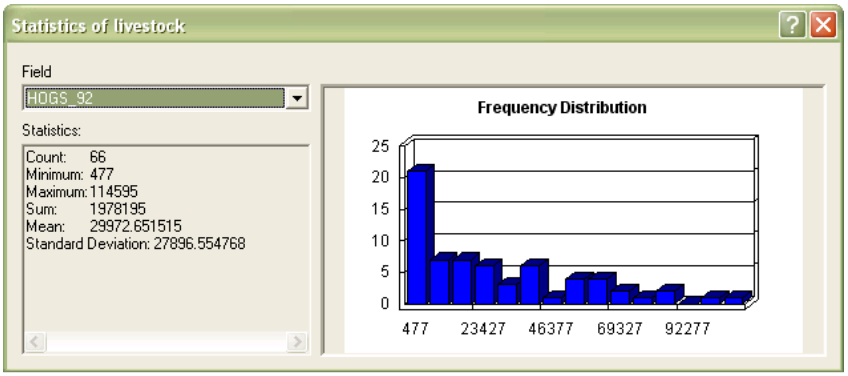
Once added to your .mxd file (this is the map that you are building in the ArcMap view) you can open a table by right clicking on it in the table of contents and selecting 'open'. The table should look similar to the one shown below and should also appear identical to an attribute table for a map.



OID	NAME	STATE_NAME	FIPS *	ACRES	CATTLE_92	CATTLE_87	CATTLE_82	MILKCOW_92	MILKC
0	Campbell	South Dakota	46021	488261.609	41407	43920	46377	1673	
1	Harding	South Dakota	46063	1709259.63	66161	59147	63280	45	
2	McPherson	South Dakota	46089	724257.068	77426	64996	72190	2815	
3	Perkins	South Dakota	46105	1841437.778	87314	97601	85912	310	
4	Corson	South Dakota	46031	1597912.004	80503	79182	74566	947	
5	Brown	South Dakota	46013	1104724.359	108382	105913	117889	2219	
6	Marshall	South Dakota	46091	559933.465	72186	72256	61276	1462	
7	Roberts	South Dakota	46109	718040.866	53419	56504	66130	3339	
8	Walworth	South Dakota	46129	476495.894	37293	35708	36767	811	
9	Edmunds	South Dakota	46045	737131.973	71086	57255	62851	2189	
10	Day	South Dakota	46037	691239.954	55579	47750	50860	5463	
11	Ziebach	South Dakota	46137	1260377.562	42515	42761	44469	39	
12	Dewey	South Dakota	46041	1560913.896	59210	50859	67410	1247	
13	Grant	South Dakota	46051	448400.275	48406	45119	50736	5357	
14	Potter	South Dakota	46107	590430.042	45004	41987	46365	432	
15	Faulk	South Dakota	46049	654050.237	59729	56503	57582	215	
16	Spink	South Dakota	46115	973429.761	101521	91702	86592	881	
17	Butte	South Dakota	46019	1442676.395	57775	44807	54700	2839	
18	Clark	South Dakota	46025	626443.54	75288	66537	64946	1249	
19	Codington	South Dakota	46029	469496.292	55218	44677	44323	7344	
20	Meade	South Dakota	46093	2215495.964	116628	114710	118687	1369	
21	Deuel	South Dakota	46039	405497.63	49731	48532	51646	6658	
22	Sully	South Dakota	46119	681898.386	35246	51046	35696	20	
23	Hyde	South Dakota	46069	550481.934	63654	49873	56874	-99	
24	Hand	South Dakota	46059	912465.368	103133	111645	110087	2145	
25	Hamlin	South Dakota	46057	345610.26	32419	31081	38321	3947	
26	Stanley	South Dakota	46117	963149.587	39146	41027	35141	14	
27	Haakon	South Dakota	46055	1178839.056	70266	82023	73182	84	
28	Beadle	South Dakota	46005	803274.729	112584	111527	120059	2803	

5. Getting statistics on tables

You can view statistics about each field (column) in order to explore the data. In the opened table window scroll to the right and right click on the 'Hogs_92' attribute and select 'Statistics'. A window should appear that looks similar to the one below. Once the statistics dialog box is open you can view the statistics for other attributes using the pulldown field menu. This same operation can be performed on attribute tables; again, the two are functionally identical.



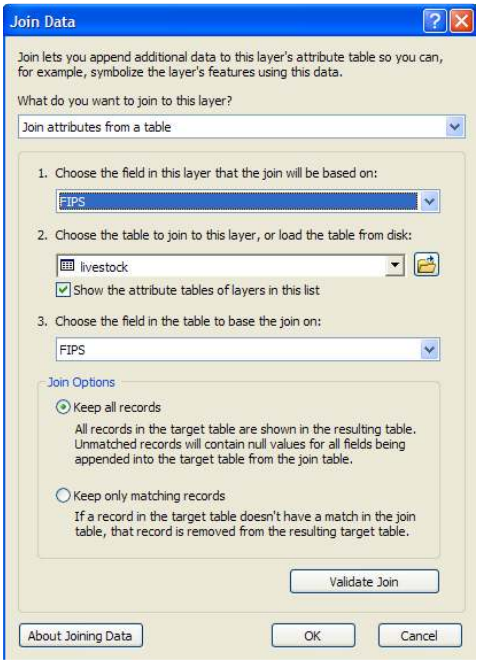
6. Joining tables

Even though the livestock data is tabulated by county it is not a map layer (i.e. spatial data set). Therefore you cannot map the data. However, if you join them to a map layer of South Dakota counties, then you could map them. The way you do this in ArcMAP is to 'join' the data table to the shapefile attribute table.

If you open the attribute table of SD_counties and livestock.dbf you'll notice they both contain a similar field, 'Fips'. This is the attribute you will use to connect the two.

Right click on SD_Counties in the table of contents and select 'joins and relates' and then 'join'. This will open up a window similar to the one below. In this dialog box you need to specify that you want to 'Join attributes from a table' at the top. Then, you need to choose the item from the SD_Counties attribute table that will link the data [choose 'fips'], the table you wish to join [select 'livestock'], and field in the data table that will link the data [select 'fips']. Select the default join options at the bottom. Your window should look similar to the one below.

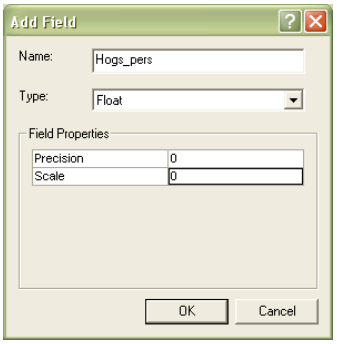
Once joined, open up the SD_Counties attribute table. You should see the livestock data appended to the far right of the table. Now you should be able to symbolize based on the livestock data. Where were most of the 1992 hogs raised in the state? What about Cattle? Are they raised in the same area?



Tip: Keeping all records is generally the best join option at the bottom of the Join Data dialog box, but sometimes it is useful to drop records for which you have no data. You can then export the joined data, creating a permanent version of the shapefile that only contains records for which you have data.

7. Adding fields to a table – Field types

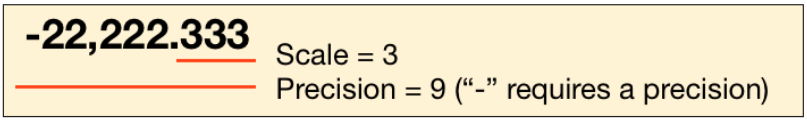
Let's say we want to know how many hogs per person were being raised in each county according to the newly joined data. While this data doesn't literally exist within the dataset we can easily compute it from what we have. To get started we'll need to add a new field to our attribute table. Open the attribute table and click on the pulldown button in the top left of the table; then, select 'add field'. In the dialog box that opens (see below) name your field 'hogs_pers' and select float as the type. There are many different data field types and you should choose the one that is most appropriate for your data. More information regarding field types can be found using the help menu. Choose float.



Field types in ArcGIS include:

Data type	Range	Size (bytes)	Example
Short integer	-32,768 to 32,767	2	Land use code
Long integer	-2,147,483,648 to 2,147,483,647	4	State population
Float	approximately □-3.4E38 to 1.2E38	4	Area, percentage
Double	approximately□-2.2E308 to 1.8E308	8	

You can define the size of the number (precision) and the number of digits after the decimal point (scale). See below.



You should now see your new field appended to the shapefile attribute table. By default ArcGIS adds the field to the end of the table you add it to. You should also see that each county has a value of zero because we haven't performed the calculation yet.

8. Editing and calculating fields

Now we need to edit the field values by dividing the total number of hogs by the total population. To accomplish this right click on the 'Hogs_pers' column header and select 'field calculator'. The field calculator is used to perform row by row calculations

utilizing existing attribute information. Using exclusively your mouse calculate Hogs_pers equal to Hogs92 divided by Pop1990. You window should look like the one below. Again, SD_Counties and Livestock prefixes are used to distinguish between the attributes that come from each of the two joined tables.

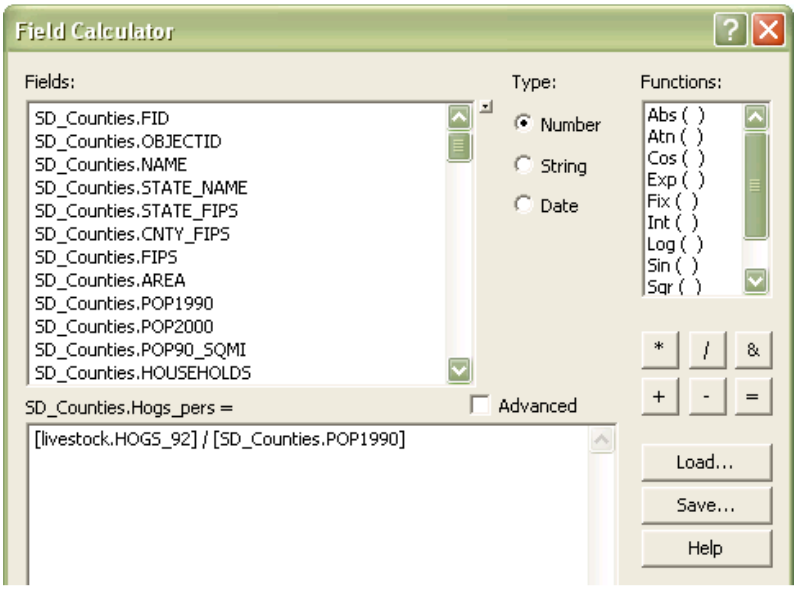


Figure 8

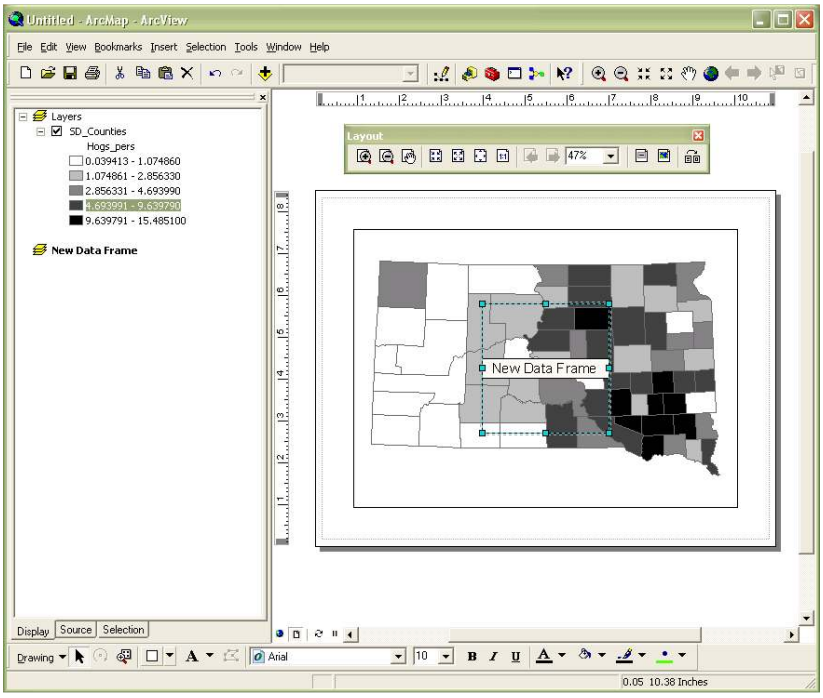
Once you select 'ok' you should now be able to symbolize based on hogs per person. Which county has the greatest number of hogs per person?

9. Multiple data frames

You can create multiple data frames within the same map in order to allow the map viewer to compare two geographic patterns side by side, or to simply provide an inset map to orient the map viewer. Let's say that we want to provide an inset map that shows the location of South Dakota within the contiguous 48 states.

Switch over to layout view and select Insert -> 'Data Frame' from the GUI. This will add a second dataframe to your map, which you will see both on your map as well as in the table of contents, as shown below.

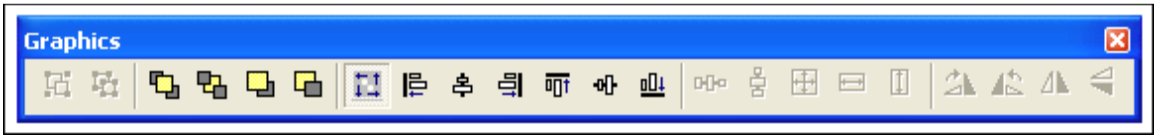
The dataframe that appears in bold typeface in the table of contents is "active", meaning it is the one you're working on. If you switch over to dataview you will not see anything because you haven't added any data yet. To go back to your previous dataframe, right-click on the dataframe name and click on "activate."



Add the states map to the new dataframe and symbolize it appropriately to highlight South Dakota from the other 48 states.

You will now need to move back over to page layout view and organize the two dataframes appropriately. By default the added dataframe sits right in the middle of the page, likely in the middle of your other dataframe, and therefore on top of your map of the SD_Counties.

Remember to take advantage of the "Align/Distribute" functions to line up maps and map elements precisely. You may want to turn on the "Graphics" toolbar (View -> Toolbars -> Graphics) to access these functions easily.



Hint: if you want to add maps from one dataframe to another you can copy and paste them like you would text. Right click on the map, select copy and then right click on the empty dataframe and select paste. (You can also do this in the table of contents area.)

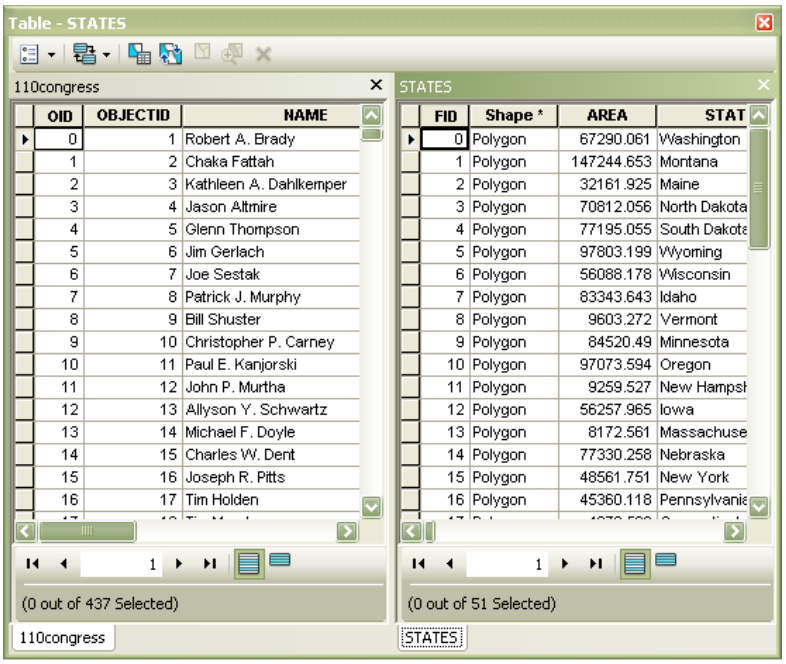
Help menu question to consider – what is a "float" item?

10. Relating a table to a layer

Let's say we want to compare the number of congresspeople from two different regions. The data provided include more than one congressperson per state. Thus, this is a situation where a relate is needed.

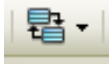
A relate is a way of creating a temporary relationship between two tables that have at least one field in common. It is used when a join would be unwieldy or the relationship is more complicated than a simple one-to-one.

Open up a new map document. Add c:\esri\esridata\USA\states.shp. Reproject your dataframe appropriately. Add 110congress.dbf from your Lab4 server folder. Open both tables and look at them side-by-side. (To do this, click-and-hold on the bottom tab of one. As you begin to move it, you will see 4-directional arrows. Move the table until you "catch" onto one of the arrows (the shadowing will indicate where the table will dock – see below).)



- What field do they have in common that would allow them to be relate-able?
- Why wouldn't a join be appropriate in this instance?

Similar to a join, you'll want to right click on the states.shp in the TOC, select joins and relates, and then "relate". In the dialog box that opens relate 110congress to states. (Note that the active table is the one with the darkened header.)

A. Select by attributes the sub-region of "Mtn." (How many states have been selected?) Go to the "related tables" icon  on the tables window.

Click on the related table (in this case, 110congress). You should see a subset of the 110congress selected.

- How many congresspeople are from the Mountain sub-region?

Click on the 110congress table. Go to "Select by Attribute." Make sure you change the type of selection you are performing from the default "Create a new selection" to "Select from current selection."

- How many congresspeople from the Mountain sub-region are Republicans?

Clear all selections using the "Clear Selection" button. 

(Do this on both tables. You might need to use the "Show All Records" button  to see the records.)

Follow the same procedure as step A above, but select by attributes for the sub-region of "Mid Atl."

- How many states have been selected?
- How many congresspeople are from the Mid-Atlantic sub-region?
- How many congresspeople from the Mid-Atlantic sub-region are Democrats?