

Homework 3: Spatial data management – JOE BREW (UFID: 0402-8902)
PHC 6937: Spatial Epidemiology

Part 1 Working with attribute table

Background: The Florida Department of Health plans to create a county-level map of cancer mortality rates. They ask you for assistance. They provide you a shapefile of U.S counties and a excel file with the county level crude and age adjusted cancer mortality rate in Florida. You are asked to create a new shapefile of counties in Florida by subsetting the U.S county map, then to join the excel data with the created shapefile and finally to export the joined file to create another new shapefile that is ready for disease mapping. By the way, they also want to learn what the five top counties with the highest cancer mortality rate are. To complete the work, please perform the following steps:

1. Add the shapefile of “usa_county” into ArcMap; Open the attribute table and select all records in FL (Please refer to “selecting records in a table” in the handout); Then export the selected record to create a new feature (i.e. Shapefile) of “CTY_FL”.

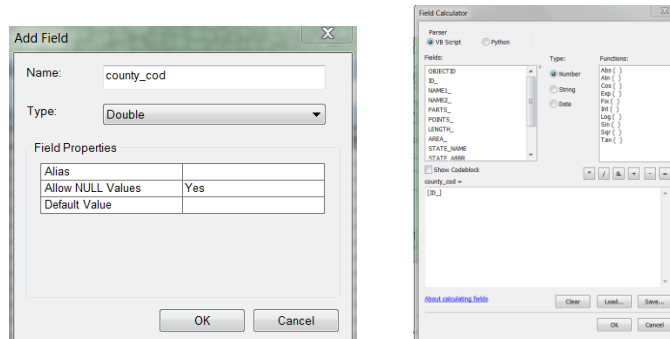
The screenshot displays the ArcMap interface with three main windows open:

- Select by Attributes:** This dialog is used to filter records based on attribute values. The 'Method' is set to 'Create a new selection'. The 'Where' clause is 'STATE_NAME = 'Florida''. The 'Output table' is 'C:\Users\Brew\Documents\ArcGIS\Default.gdb\CTY_FL'.
- Table:** This window shows the attribute table for the 'usa_county' shapefile. The columns are: OBJECTID, Shape, ID, NAME1, NAME2, PARTS, POINTS, LENGTH, AREA, and STATE_NAME. The records for Florida are highlighted in blue.
- Export Data:** This dialog is used to export the selected records to a new shapefile. The 'Export' option is 'Selected records'. The 'Output table' is 'C:\Users\Brew\Documents\ArcGIS\Default.gdb\CTY_FL'.

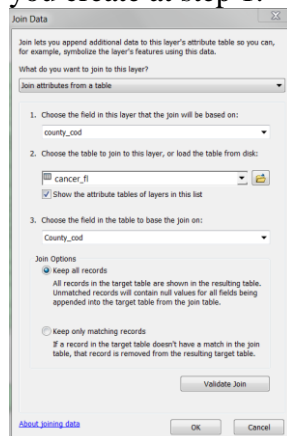
OBJECTID	Shape	ID	NAME1	NAME2	PARTS	POINTS	LENGTH	AREA	STATE_NAME
288	Polygon	10003	10003	NEW CASTLE	1	15	102.0094	450.6849	Delaware
289	Polygon	10005	10005	SUSSEX	1	17	129.8009	973.1685	Delaware
290	Polygon	12001	12001	ALACHUA	1	24	134.1004	967.2103	Florida
291	Polygon	12003	12003	BAKER	1	14	107.5154	597.0373	Florida
292	Polygon	12005	12005	BAY	1	15	143.3631	899.5897	Florida
293	Polygon	12007	12007	BRADFORD	1	12	86.15348	309.1252	Florida
294	Polygon	12009	12009	BREVARD	1	20	184.2497	1316.796	Florida
295	Polygon	12011	12011	BROWARD	1	8	146.9438	1198.039	Florida
296	Polygon	12013	12013	CALHOUN	1	14	105.3266	575.9456	Florida
297	Polygon	12015	12015	CHARLOTTE	1	9	130.5776	817.1231	Florida
298	Polygon	12017	12017	CITRUS	1	19	111.2361	675.355	Florida
299	Polygon	12019	12019	CLAY	1	13	112.5144	654.2513	Florida
300	Polygon	12021	12021	COLLIER	1	16	211.6061	2120.581	Florida
301	Polygon	12023	12023	COLUMBIA	1	24	143.7574	798.7055	Florida
302	Polygon	12025	12025	DADE	2	36	224.1625	2037.327	Florida
303	Polygon	12027	12027	DE SOTO	1	6	103.0399	642.4259	Florida

2. The dataset of “cancer_fl.dbf” have the crude and age adjusted mortality rate in all counties in FL.

The data of map (ID_) and table (county_cod) have different variable formats: “ID_” is a string and “county_cod” is a double. To solve the problem, you need to open the attribute table of the map after you select Florida. Then add a new field to the table with a format of “double” and use the “field calculator” to assign the values from “ID” to this created field.



Please join the cancer mortality data with the new shapefile of “CTY_FL” that you create at step 1.



Then, use the joined information to create a new shapefile with the cancer mortality information attached. (Note: please zip the Shapefile and submit it together with this word document for evaluation)

Name	Date modified	Type	Size
question2joebrew	9/21/2014 4:45 PM	File folder	
HW3.gdb	9/21/2014 4:45 PM	File folder	
question2joebrew	9/21/2014 4:47 PM	Compressed (zipp...	
question2joebrew.shx	9/21/2014 4:45 PM	SHX File	
question2joebrew.shp.CHD01MLTB89PB...	9/21/2014 4:45 PM	LOCK File	
question2joebrew.shp	9/21/2014 4:45 PM	SHP File	
question2joebrew	9/21/2014 4:45 PM	SBX File	
question2joebrew.sbn	9/21/2014 4:45 PM	SBN File	
question2joebrew.prj	9/21/2014 4:45 PM	PRJ File	
question2joebrew.dbf	9/21/2014 4:45 PM	DBF File	
question2joebrew.cpg	9/21/2014 4:45 PM	CPG File	

3. Using the shapefile created in step 2, please answer the following questions (Hint: Selection by attribute):

a. Please find five top counties with the highest crude cancer mortality rate in Florida and list the information of county name and their rates below.

NAME2	PARTS	POINTS	LENGTH	AREA	STATE_NAME	STATE_ABBR	Shape_Le_1	Shape_Area	county_cod	OID	County	County_c_1	Deaths	Population	Crude_Rate	Age_Adjust
CITRUS	1	19	111 2361	875 356	Florida	FL	1 751859	0 1614	12017	8	Citrus County, FL	12017	233	638950	429.2	210.4
UNION	1	13	67 76862	245 6321	Florida	FL	1 083301	0 05944	12125	62	Union County, FL	12125	233	58487	398.4	472.1
CHARLOTTE	1	9	130 5776	817 1231	Florida	FL	2 058926	0 191916	12015	7	Charlotte County, FL	12015	2271	614482	369.6	175
HERNANDO	1	12	106 293	499 7113	Florida	FL	1 685571	0 119212	12053	25	Hernando County, FL	12053	2361	637366	368.9	266
INDIAN RIVER	1	10	161 6185	528 2437	Florida	FL	1 580785	0 124946	12061	29	Indian River County, FL	12061	1863	611263	362.4	191.4

Citrus	429.2
Union	398.4
Charlotte	369.6
Hernando	368.9
Indian River	362.4

b. Please find five top counties with the highest age-adjusted cancer mortality rates in Florida and also list the information of county name and their rates;

County	County_c_1	Deaths	Population	Crude_Rate	Age_Adjust
Union County, FL	12125	233	58487	398.4	472.1
Madison County, FL	12079	222	75477	294.1	253.1
Washington County, FL	12133	261	88776	294	248.1
Hamilton County, FL	12047	128	56052	228.4	245.2
Levy County, FL	12075	503	151405	332.2	242.3

Union	472.1
Madison	253.1
Washington	248.1
Hamilton	245.2
Levy	242.3

- c. Please check if the selected five counties with highest crude and age-adjusted rates are same or different. If they are different, please explain why it is.

With the exception of Union County (which has a top-5 crude *and* age-adjusted mortality rate), the counties at the top of the two lists are *different*. The reason for this is that cancer mortality is very subject to the age of a population – older populations, even if healthy, have high mortality from cancer, and younger populations, even if unhealthy, have lower mortality.

So, it should be no surprise that the *crude* rates largely reflect how old a population is, whereas the age-adjusted rates largely reflect a different indicator – how *unhealthy* the population is.

Part 2: Prepare spatial data

Data:

1. FL_Hospitals: This dataset contains 2011 Hospital Facility Information for the State of Florida. It is a combination of hospital facility addresses from seven different sources. The data contains selected fields denoting the name, physical address, and other facility information for hospitals located in Florida.
2. Alachua_boundary: A polygon is to show the boundary of Alachua County, FL.
3. NED01: This dataset contains the U.S. Geological Survey National Elevation Dataset (NED) for Alachua County. This is a Grid data.
4. Alachua_CT_POP: A polygon includes the population information at the census tract level in Alachua County. (The variable of Population number is “pop”).

Task 1. From the feature class of “FL_Hospitals”, please select all hospitals within Alachua County. (Hint: Using Alachua_boundary and “Intersect” tool).

List the ID (FID FL Hospitals) and the names of the selected hospitals below for evaluation.

FID_FL_Hospitals	OBJECTID	STATUS	SCORE	SIDE	FDEM_ID	NAME
27	27	M	0	*		SELECT SPECIALTY HOSPITAL - GAINESVILLE
28	28	M	0	*	10224863	MALCOM RANDALL VA MEDICAL CENTER
29	29	M	0	*	10224866	SHANDS AT THE UNIVERSITY OF FLORIDA
31	31	M	0	*	10225061	NORTH FLORIDA REGIONAL MEDICAL CENTER
33	33	M	0	*	10225073	SHANDS REHAB HOSPITAL
32	32	M	0	*		SHANDS VISTA BEHAVIORAL HEALTH

27	Select Specialty Hospital
28	Malcom Randall
29	Shands at UF
31	NFRMC
33	Shands Rehab

32	Shands Vista
----	--------------

Task 2. Find the elevation information for the selected hospitals in Alachua County using the data of “NED01”. **List the ID (FID FL Hospitals) of the hospitals and their elevations for evaluation.**

FIDEM_ID	NAME	DESCRIPT	FLAG	UPDATE_DAY	FGDLAQDATE	AUTOID	FID_Alachua_boundary	Dissolve_id	RASTERVALU
	SELECT SPECIALTY HOSPITAL - GAINESVILLE	SELECT SPECIALTY HOSPITAL - GAINESVILLE		5/10/2011	4/3/2009	27		1	26
10224863	MALCOM RANDALL VA MEDICAL CENTER	MALCOM RANDALL VA MEDICAL CENTER		5/10/2011	4/3/2009	28		1	28
10224866	SHANDS AT THE UNIVERSITY OF FLORIDA	SHANDS AT THE UNIVERSITY OF FLORIDA		5/10/2011	4/3/2009	29		1	27
10225061	NORTH FLORIDA REGIONAL MEDICAL CENTER	NORTH FLORIDA REGIONAL MEDICAL CENTER		5/10/2011	4/3/2009	31		1	35
10225073	SHANDS REHAB HOSPITAL	SHANDS REHAB HOSPITAL		5/10/2011	4/3/2009	33		1	43
	SHANDS VISTA BEHAVIORAL HEALTH	SHANDS VISTA BEHAVIORAL HEALTH		5/10/2011	4/3/2009	32		1	48

ID	Elevation
27	26
28	28
29	27
31	35
33	43
32	48

Task 3. Find the serving population within 5km distance of each hospital using census tract population in Alachua. **List the ID (FID FL Hospitals) of the hospitals and the number of the serving population within 5 km distance for evaluation.**

(Hint: Creating a buffer of 5 km for each hospital and then spatial join the census population using Spatial Join tool. Figure 1 shows how to assess the tool. In addition, in spatial join tool, right click the variables and you can select the statistics as shown in the second figure)

FID_FL_Hospitals	DESCRIPT	pop
27	SELECT SPECIALTY HOSPITAL - GAINESVILLE	7011
28	MALCOM RANDALL VA MEDICAL CENTER	7011
29	SHANDS AT THE UNIVERSITY OF FLORIDA	7011
31	NORTH FLORIDA REGIONAL MEDICAL CENTER	3208
33	SHANDS REHAB HOSPITAL	3208
32	SHANDS VISTA BEHAVIORAL HEALTH	3208