INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
sustainable solutions for ending hunger and poverty
Supported by the CGIAR

ETHIOPIAN DEVELOPMENT RESEARCH INSTITUTE
EDRI

Ethiopian Strategy Support Program II (ESSP II)

An Introduction to Geographical Information Systems

Training Manual

Emily Schmidt, Helina Tilahun,
Mekamu Kedir, and Hailu Shiferaw

Ethiopia Strategy Support Program II (ESSP II)

Ethiopian Development Research Institute (EDRI)

ESSP II / EDRI Training Manual

December 2011

IFPRI-HEADQUARTERS
INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
2033 K Street, NW • Washington, DC 20006-1002 USA
Tel: +1-202-862-5600 • Skype: IFPRIhomeoffice
Fax: +1-202-467-4439 • E-mail: ifpri@cgiar.org

IFPRI-Addis Ababa
http://essp.ifpri.info
IFPRI c/o ILRI
P.O. Box 5689, Addis Ababa, Ethiopia
Tel: +251 11 6 17 25 55 Fax: +251 11 6 46 23 18
E-mail: ifpri-addis@cgiar.org
Contact: Bart Minten, Senior Research Fellow and Program Leader

ETHIOPIAN DEVELOPMENT RESEACH INSTITUTE
http://www.edri.org.et/
Blue Building • Addis Ababa Stadium
P.O. Box 2479 • Addis Ababa, Ethiopia
Tel: +251 11 5 50 60 66; +251 11 5 52 53 15
Fax: +251 11 5 50 55 88
Email: exe-director@edri.org.et

Contents

Lab 05: Understanding projections

Projecting Map Data using ArcMap

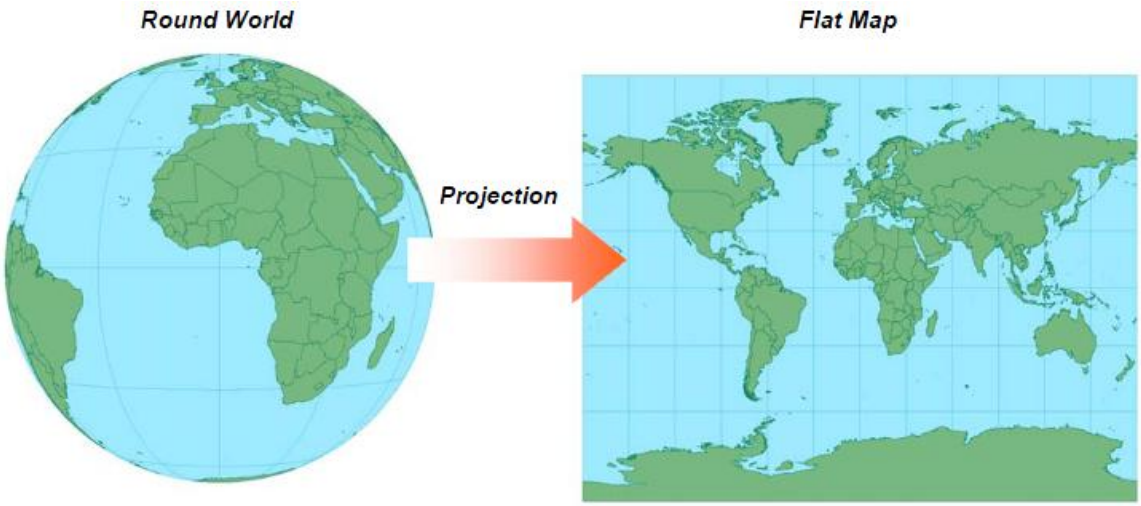
The location of any given place can be defined with reference to lines of latitude and longitude, which create an imaginary mesh over the world. Latitude and Longitude values belong to a spherical coordinate system – a system for defining locations and making measurements on a sphere, or something close to a sphere (a spheroid) like the earth.

The latitude - longitude value of a point depends on the assumptions you make about the earth's shape. The earth isn't perfectly round. It bulges at the equator and is flattened at the poles. Technically, this makes it an oblate spheroid.

Besides not being quite round to begin with, the surface of the earth has various bumps and indentations. Determining the exact shape of the earth is not a simple matter.

There are many different models and ArcMap recognizes almost three hundred different projections.

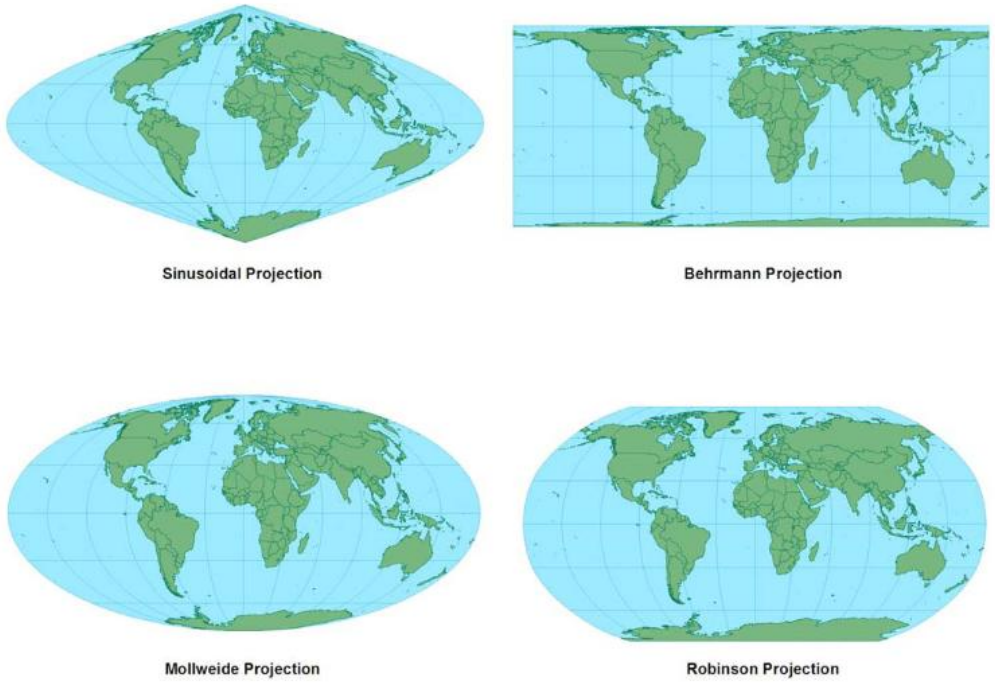
To make one map, one of these models of the earth (or some part of it must be represented on a flat surface. This is accomplished by a mathematical transformation called a ***map projection***.



Just as location on a sphere is defined by latitude and longitude, location on a map is defined by Cartesian coordinates, which assign values to points according to their positions on a horizontal x-axis and a vertical y-axis. As opposed to a spherical coordinate system, this is known as a planar coordinate system.

The exact location of a point on a map varies according to the map projection used.

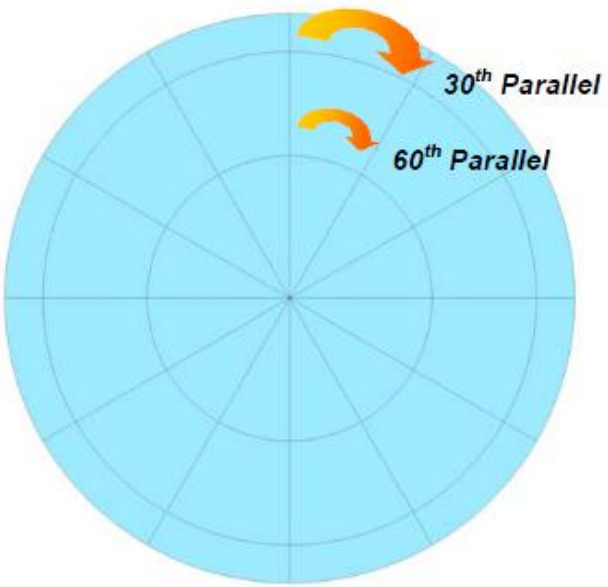
There are about 50 commonly used projections and many variations on each.



Four world projections. Many projections are made for individual continents, countries, parts of countries, or strips of land that may cross international boundaries.

Every spatial data set in a GIS stores geographic coordinates for its features. These coordinates make up its *geographic coordinate system* (GCS). A data set that has been projected also stores *Cartesian* coordinates for its features. These make up its *projected coordinate* system.

When you work with unprojected data (data that has only a geographic coordinate system), any measurements or calculations you make are only based on a sphere or spheroid. This is problematic because degrees of latitude do not have constant length. A degree of latitude at the 30th parallel (30 degrees north of the equator) is longer than a degree of latitude at the 60th parallel. Both represent 1/360th of a circle, but the circles have different circumferences.



View from the North Pole

Since degrees of latitude are not constant, they can't be used to make meaningful measurements of distance and area. This problem is overcome with map projections.On a flat surface, units of measurement (meters or feet, for example) are constant, which means that you can calculate meaningful area and distance measurements.There is another difficulty however. Since the world is a sphere, and maps are flat, you can't go from one to the other without changing the proportions of features on the surface. Map projections distort shape, area, distance and direction. Some projections preserve one of these properties at the expense of others, some compromise on all of them, and some preserve properties for one part of the world and not the others. The Mercator projection, for example, preserves direction, but distorts area. The sinusoidal projection preserves area but distorts shape.

Mercator Projection

Sinusoidal Projection

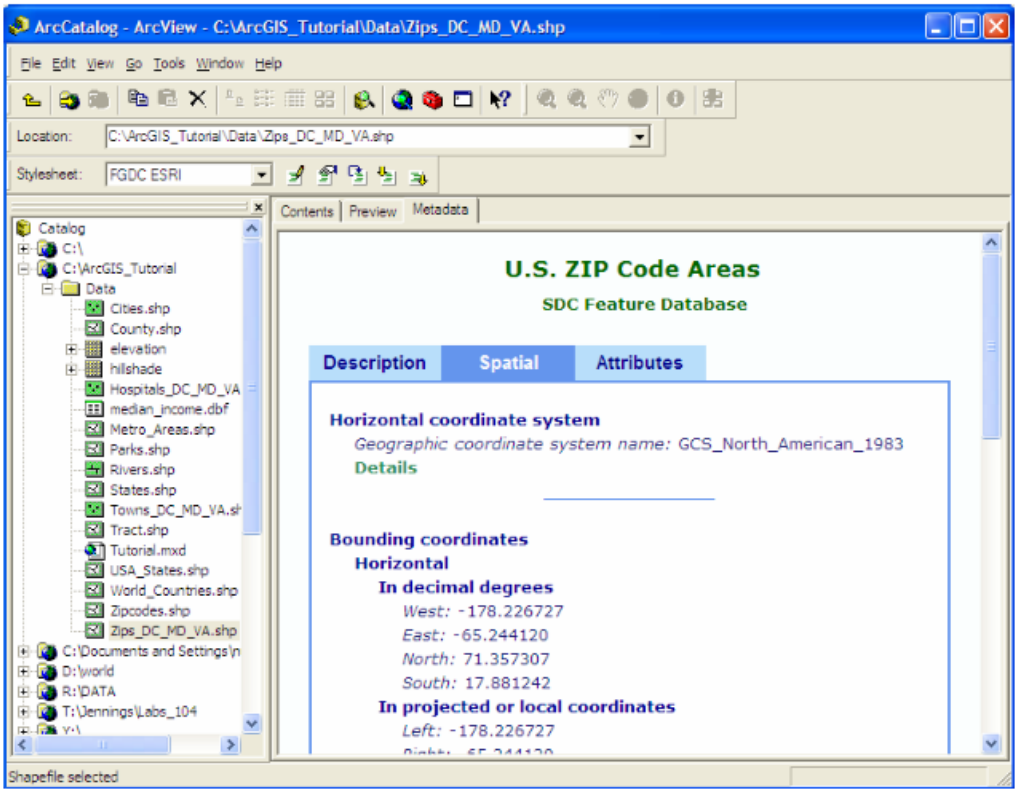
In the Mercator projection, Greenland looks larger than Brazil, although Brazil is four times its size. Because direction is preserved, Brazil correctly appears due south of Greenland. In the sinusoidal projection, the proportional sizes of Greenland and Brazil are correct. Their shapes however are distorted – Greenland is too narrow, and Brazil is too wide.

Your choice of map projection allows you to control the type of distortion in a map for your area of interest. If you are working with a fairly small area and using an appropriate projection, the effects of distortion are insignificant. If you are working with the whole world, there is bound to be significant distortion of some spatial property.

When you add a layer to a map, both its appearance, and, the results of measurements and calculations you make depend on its coordinate system. You can find a data set's coordinate system in its spatial metadata.

To access the metadata for your file (if indeed it exists) open ArcCatalog. ArcCatalog is a standalone application for managing geographic data. It is part of the ArcMap Suite, and is essentially an interactive browser for spatial information.

To view the metadata for a particular file, double-click on the ArcCatalog icon on your desktop. In the table of contents, navigate to your chosen file. Highlight this file, and click on the Metadata tab that sits over the view window. (You can also access a preview of your spatial data by choosing the preview tab).



When data sets that have the same coordinate system are added to a data frame, the features in each layer are correctly positioned with respect to each other. If you subsequently add a data set that has a different coordinate system, ArcMap changes it to match the others in a process called "on-the-fly" projection. This new, temporary projection is applied only within a particular data frame; the data set's native coordinate system (the one shown in its spatial metadata) does not change.

By default, layers are projected on the fly to the coordinate system of the first layer added to a data frame (even if the layer is later removed). The coordinate system is stored as a property of the data frame and can be changed. You can project all layers in a data frame to any coordinate system ArcMap supports.

To project a layer on the fly, ArcMap uses the information stored in its geographic coordinate system. On-the-fly projection works best when all layers in the map have the same GCS (in other words, when they all use the same model of the earth.)

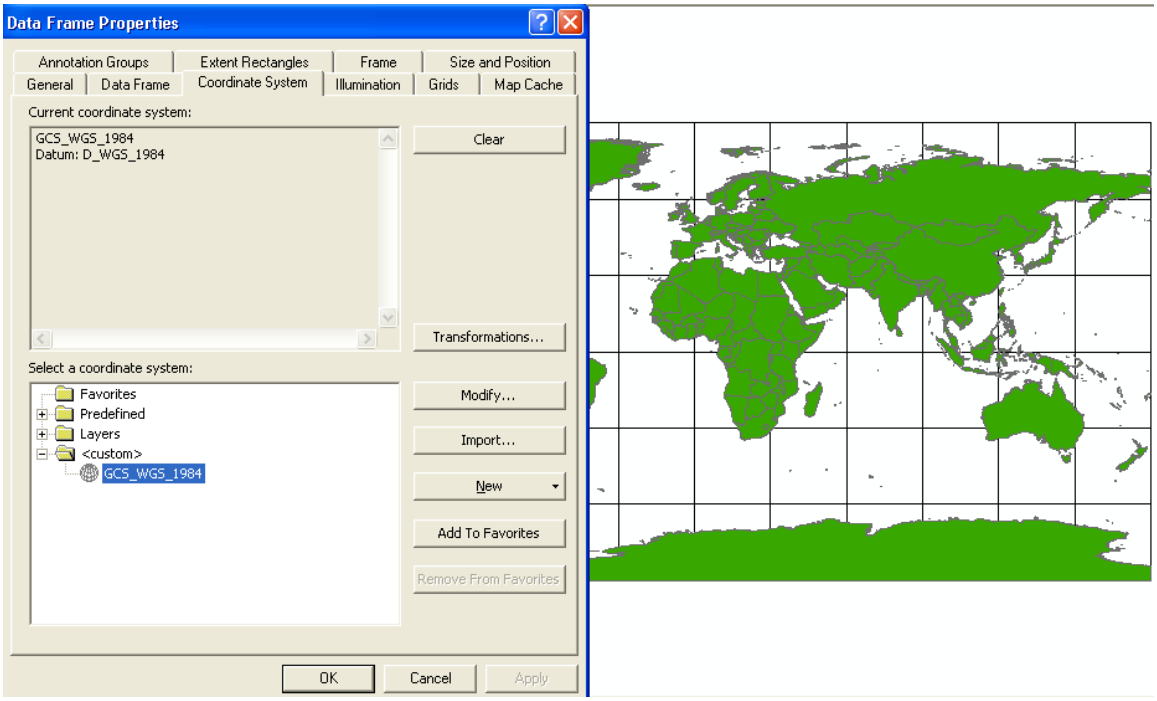
On-the-fly projections are less mathematically rigorous than permanent projections (which change the native coordinate system of the data set). If you plan to use data sets in an exacting analysis, you should project them permanently to the same coordinate system with the *ArcToolbox Projection Wizard.*

Go to the next page for the Lab05 exercise.

Lab 05 Exercise - Part 1: Changing Data Projections in ArcMap

1. Open ArcMap, and from your Lab05 folder, open the files, “World_Countries”, and “World30”. Change the Symbology of the “World30” layer so that it is a hollow fill, with black or grey outlines.
2. Right click on “Layers” in the table of contents and point to “Properties”. Select the tab “Coordinate Systems” (if not already selected by default).
3. Move the Properties window a little to the right so that the majority of your map is visible.
4. You will see, under the box “Current coordinate system” that these shapefiles presently have a geographic coordinate system called “GCS_WGS_1984”. This stands for “Geographic Coordinate System (GCS), World Geodetic Survey (WGS), 1984”

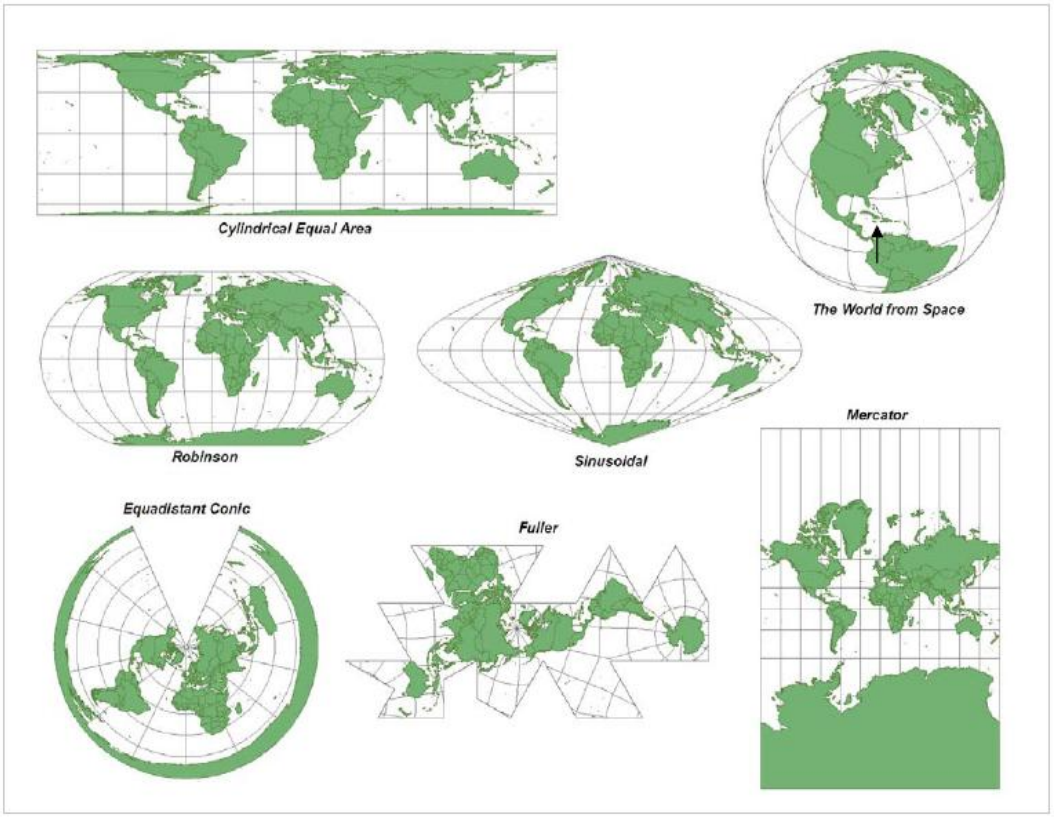
   This is a popular projection for the World, and is the reference system used by GPS (Global Positioning System) Units. However, there are alternate projections available (such as those introduced in the intro, and they have different purposes for different maps). Let us explore them.



5. In the folder tree, click on the folder “Predefined” (This is where all the alternate projections are stored). In the next list of folders, click on “Projected Coordinate systems”. The subsequent list is quite long, scroll to the end and click on “World”.

   These are the world level projections available in ArcMap.
6. Select “Cylindrical Equal Area”, then click Apply, and OK. The appearance of your world map should change dramatically. (You may get a pop-up box asking you if you are sure that you want to do this, Just click OK).
7. Move to the “Layout View” for your map. Resize your map so that it fits the whole page.
8. Return to “Layer Properties”, and click on the tab “Frame”. Remove the “Border” from around your map window. Then click OK. (This will make your final product a little less cluttered, as we plan to use several images).
9. At this point, save your map in your Lab05 folder as “Lab_05World.mxd”.

10. Next, go to the “File” dropdown menu. Choose “Export map”, and save your map, (as a .jpeg with resolution of 150), in your Lab05 folder as “C Equal Area”.
11. Return to the table of contents. Once again, right-click on “Layers” and scroll to “Layer Properties”. Select the tab “Coordinate System” and following the same route as Step 5; (Predefined/Projected…./World), select the projection “Equidistant Conic”. Click Apply, and OK.
12. Look at how different this view of the world is! Once again, export this map as a .jpeg with resolution of 150dpi. Choose an appropriate name, e.g. “Equid Conic”.
13. Using the instructions from the previous steps, produce 5 additional jpegs of the world using the following projections:
    a. Mercator
    b. Sinusoidal
    c. The World from Space
    d. Fuller
    e. Robinson
14. Using Microsoft word, or PowerPoint, inserts the jpegs onto one page/slide. Label each projection according. See the next page for a sample layout of all images.
15. Your finished product should look somewhat like the following graphic on the following:

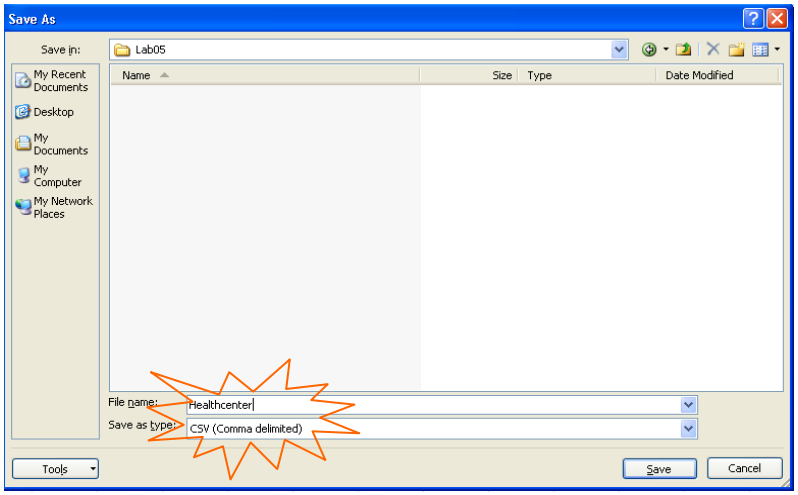


Lab 05 Exercise - Part 2: Bringing field data into the ArcGIS software program. Allowing visualization using GPS point data.

Your colleague has just returned from mission, and gave you a simple database file containing the X, Y (latitudinal & longitudinal) information of health centers located throughout Ethiopia. This information must be visualized, and integrated into the spatial data repository.

➔ **Close ArcGIS before you begin this exercise.**

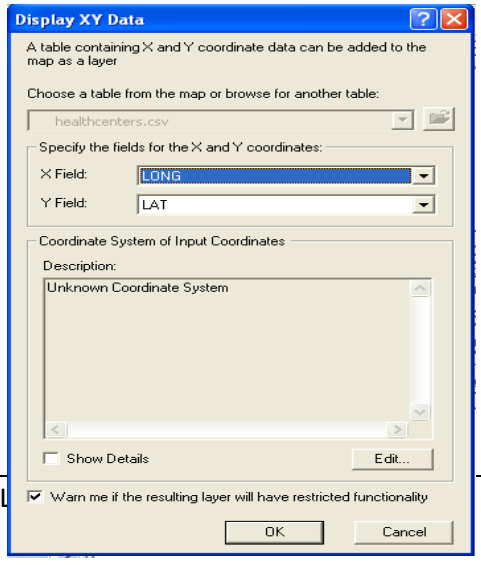
1. Open Excel, and go to your Lab 05 folder. Open Healthcenters.xls". You can see that there is a "LONG" (longitude) and "LAT" (latitude) field, along with other information that describes each health center.
2. In order to bring an Excel worksheet into ArcGIS, you must save it as a .dbf file, or a .csv file (make sure to choose ms-dos format). This is sometimes problematic because it occasionally truncates values during the saving process.
3. Resave your file as "Healthcenters.csv" into your Lab05 folder.



4. Close all worksheets in Excel.

   You need to have the worksheet that you intend to add to ArcGIS closed in excel in order to add correctly.

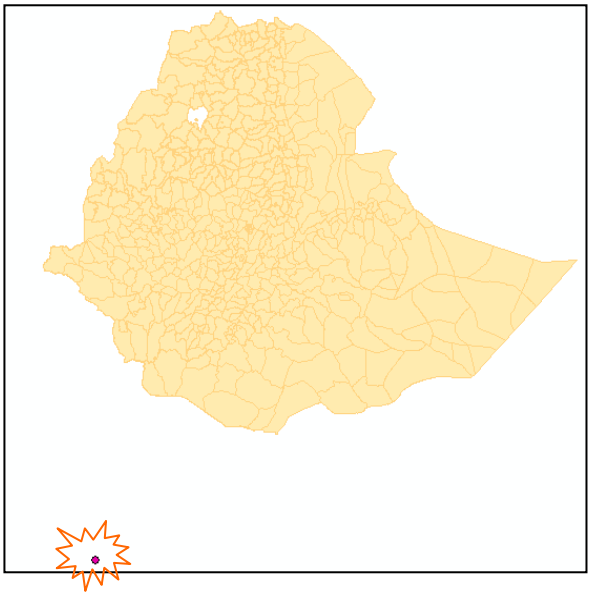
5. Open a new session in ArcGIS and add the file "Woreda" from your Lab05 folder and the database file: "Healthcenters.csv" that you saved in your Lab05 folder.
6. Right click on the Healthcenters.csv file and scroll to "Display XY data". Click on "Display XY data" and a window should pop up like the window to the right.
7. Make sure that your "X Field" displays LONG for the coordinates, and your "Y Field" displays LAT for your coordinates. Press OK



8. You may get a warning message stating that your "Table Does Not Have Object-ID Field". This is a

unique identifier that ArcGIS builds into all of its shapefiles. Press OK and ArcGIS will create this field for you.

9. Now, can you see your "HealthCenter" points? **Where are they?** Right click on the "HealthCenter" layer, and from the menu choose "Zoom to layer". The "HealthCenter" layer should now be visible, but not the "Woreda" layer.
10. Right click on the "Woreda" layer and choose "Zoom to Layer". What happens to the "HealthCenter" layer? Magic, it disappeared…or is it a projection problem??
11. Go to the main tool bar, and select the "Zoom to full Extent" button 
12. You should now see the entire "Woreda" layer, with one tiny dot to the Southwest of the country. If you use the regular zoom tool  , and zoom repeatedly into this dot or draw a small square with your zoom tool, you will realize that it is in fact the "HealthCenter" layer. As it is in a different projection, it is unable to locate and resize itself correctly in relation to the "Woreda" Layer.
13. If we take the assumption that this information was collected by GPS, then reverting to the default coordinate system used by GPS will correct this issue.



14. So let's try our hypothesis! First, Right click on the "Healthcenters.csv Events"; you select Remove.
15. Now, right click on the "healthcenters.csv" layer. Left click on the "Display XY Data".
16. As you can see, the coordinate system is unknown. Click on the "Edit" button. In the next window, click the "Select" button and choose the following path:

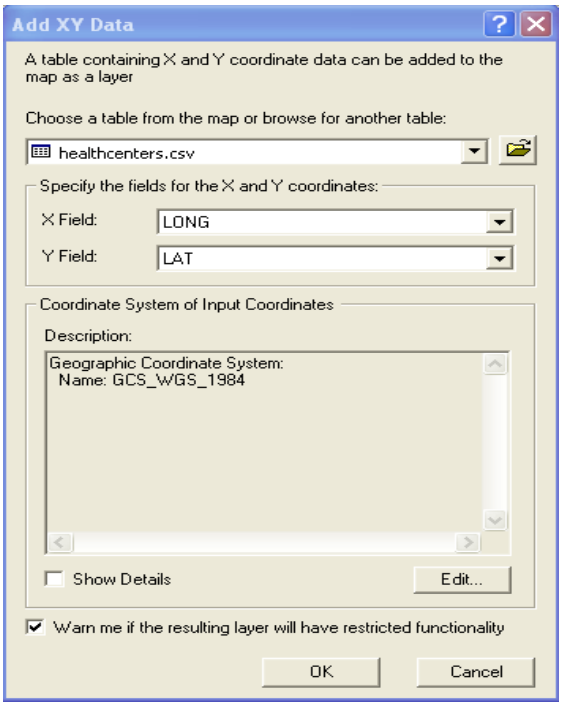
    Geographic Coordinate Systems

    World

    WGS 1984

17. Click Add.

18. Your “Add XY Data” window should now look like the graphic to the right. Click OK



19. Now your Healthcenters should be geographically contiguous with your “Woreda” layer.

20. Your “Healthcenters.csv Events” is currently only a cosmetic layer. We know this because it has the word “Events” following the name. It is not yet a shapefile.

21. To create a permanent shapefile from this cosmetic layer, right click on the “Healthcenters.csv Events”, scroll down to “Data” and select the “Export Data” option.

22. Leave all the initial options as default, but make sure to save the final file to Lab05 folder, calling the file “HealthCenter.shp”

23. A pop – up window will ask you if you would like to “Add to map”. Select OK and the new shapefile should automatically add to the dataframe.

24. Look at these data, where are Healthcenters missing, why? Add some other geographic data from your Lab01 – Lab04 folder (rivers, roads) to see if you can hypothesize. Just looking at your map, can you find a place where a hospital should be built?

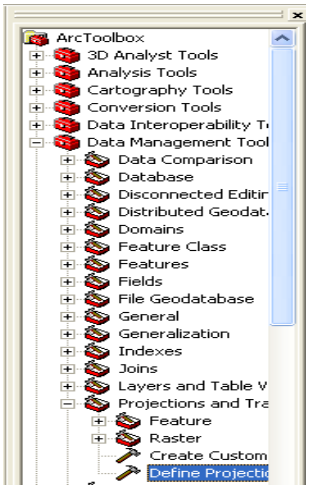
25. Create a map with the Healthcenters layer, and other key data that you think may affect where these Healthcenters are located. (hint: are they located in large cities?, are they near roads?)

Lab 06: Exporting external database information from ArcMap

Exercise overview

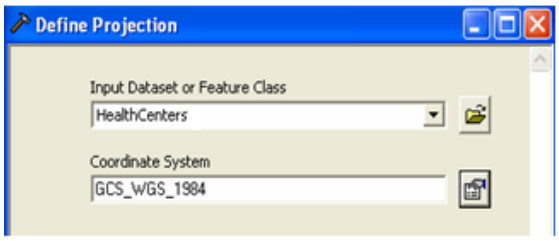
You have been asked to simplify the healthcenters data You will perform what is known as a **Spatial Join** to generate this information, then you will export this information to excel so that the project lead can generate an excel graph to demonstrate these numbers. A spatial join, links/combines the attributes of two layers, based on the location of each layer's features. Just like a table join, a spatial join appends the attributes of one layer to another. You can then use the additional information to query your data in new ways.

1. Open a new ArcMap session and add your new "HealthCenters" layer that you created in the previous exercise.
2. Before you begin this exercise, you may need to add the spatial toolboxes to your datascreen. If the list of toolboxes (see left) is not on your screen click on the  button in the center of the toolbar.
3. Now, let's make sure that your "HealthCenters" layer has a defined projection by going to the toolbox "Define projection" (see graphic Right).



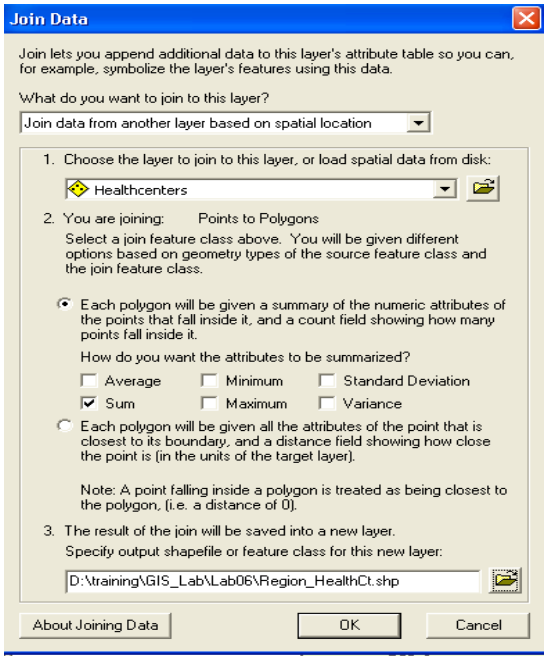
4. Define the projection in the "Select a Coordinate System" box, choose the following path:

   Predefined
   Geographic Coordinate Systems
   World
   WGS 1984



5. Press OK. Now that we have defined all of our projections we can do a Spatial Join!
6. Add your "Regions" and "Lakes" layer from your Lab01 folder.
7. Right click on the "Regions" layer. Go to "Joins and Relates" > "Join"

8. In the first drop down menu (see right), change the option to “Join data from another layer based on spatial location”. This choice will change the look of the wizard layout.
9. Use the graphic to the right to select the correct options.
10. For Option 2: make sure the first radio button is selected (it should be by default), and that the “Sum” check box is ticked.
11. Save the new file to your Lab06 folder, and call it “Region_HealthCt”.
12. Click OK.



13. The join process may take several seconds, as ArcMap must count the number of health centers located in each region, and report that data in the attribute table of your new file “Region_HealthCt”

➔ As you can see, there are many other options within the Spatial Join tool. It can be used to determine how close an individual point or polygon is to another point or polygon in a different layer, and report its distance. You will use this function in a later exercise.

14. “Region_HealthCt” should automatically add to your table of contents, if not, go to the “Add data” button, and from your Lab06 folder, add the new “Region_HealthCt” layer. Re-organize your layers so you can see the “Lakes”
15. Open the attribute table of “Region_HealthCt”, and scroll right until you see “Count_”.
16. The field that you are interested in is the “Count_” field. Right-click on the field heading, and choose the “Sort by Descending” option. You will see that the maximum number of health centers for any one Region is 26.

Attributes of Region_HealthCt

FID	Shape	FID_1	OBJECTID	REGION_R2I	REGION_R6I	Region_Nam	Area_KM2	Shape_Le	Shape_Ar	Count_
4	Polygon	4	174	04	040000	Oromia	324494.2066	57.970976	26.58502	26
2	Polygon	2	29	03	030000	Amhara	152588.038228	30.791905	12.650807	15
9	Polygon	9	313	07	070000	SNNP	112899.958774	24.483418	9.220553	15
0	Polygon	0	1	01	010000	Tigray	50206.413993	16.44823	4.198869	12
5	Polygon	5	183	05	050000	Somali	315619.705388	34.498025	25.686615	6
1	Polygon	1	7	02	020000	Afar	95623.124127	16.117192	7.939469	3
3	Polygon	3	76	06	060000	Benishangul Gumuz	50003.469499	19.570954	4.118421	2
6	Polygon	6	223	15	150000	Dire Dawa	1055.561004	1.696146	0.086776	1
7	Polygon	7	246	13	130000	Harari	371.65333	0.763938	0.030517	1
8	Polygon	8	278	14	140000	Addis Ababa	550.687794	1.031316	0.045314	1
10	Polygon	10	358	12	120000	Gambella	25701.344422	8.279851	2.094177	1

Record: 0 Show: All Selected Records (0 out of 11 Selected) Options

17. Close the Attribute table for now, and let’s create a thematic map.

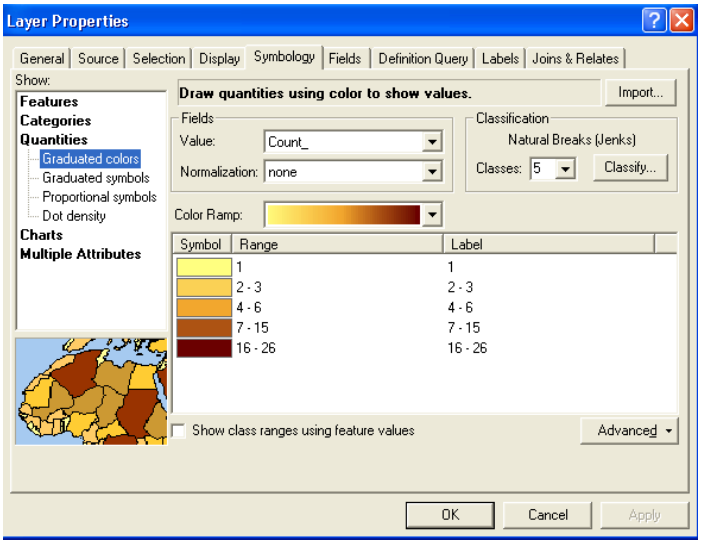
18. Double click on the “Region_HealthCt” layer. This will take you to the “Layer Properties” dialog. Click on the “Symbology” tab.

    In the “Show” menu to the left, Click on Quantities > Graduated colors.

19. In the “Value field”, click the down arrow and scroll to “Count_”.

20. In the “Color Ramp” field click on the down arrow and choose a color scheme that you like.

21. Click OK when complete.



22. Before you progress too far in your map production, let’s once again save your map as an ArcMap Document file (.mxd) so you won’t loose your work to date. Go to “File/Save As”, navigate to your “Lab06” folder, and name the Project file: “Lab06_YOURNAME”.

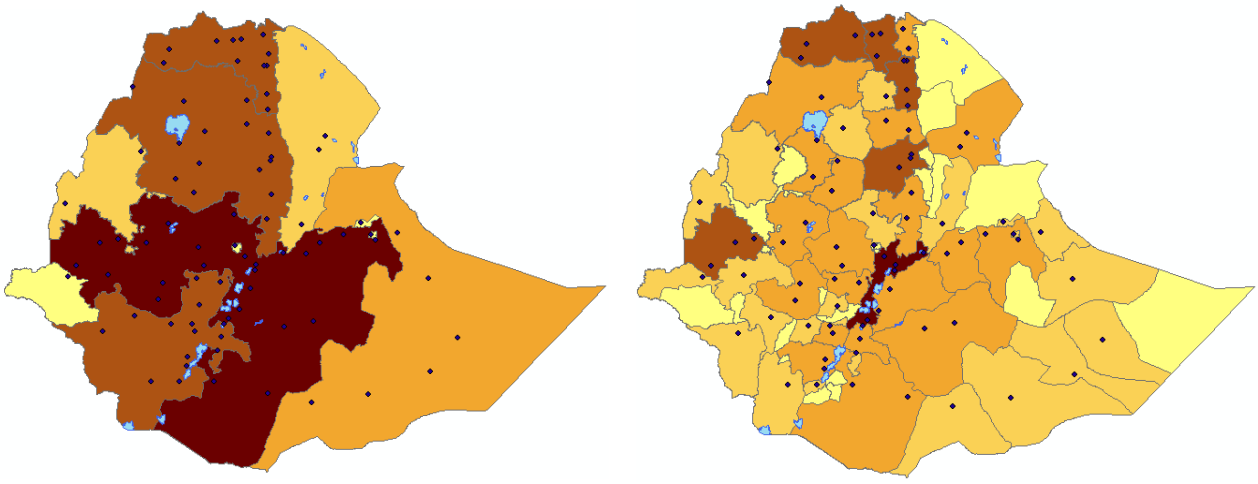
23. Now make sure that your “HealthCenters” shapefile is above your “Region_HealthCt” file in the table of contents so you are able to see both layers.

24. Given that the Regions are geographically quite large in Ethiopia, would it be better to look at statistics on a more disaggregated level. You can do the same spatial join on your Zone layer to understand health center placement at a finer level.

25. Create a map with one of the joins that you did in order to count number of hospitals for each administrative unit

Think about how you would aggregate similar data such as road data – you could do a spatial join, and average road lengths by Woreda in order to arrive at a road density figure.

Here is the difference between the Region and Zone level statistics!



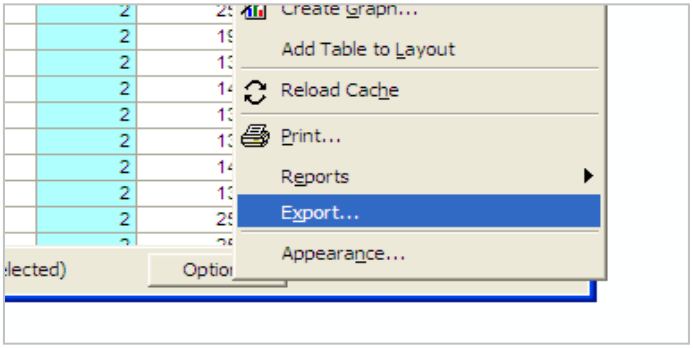
Region Level

Zone Level

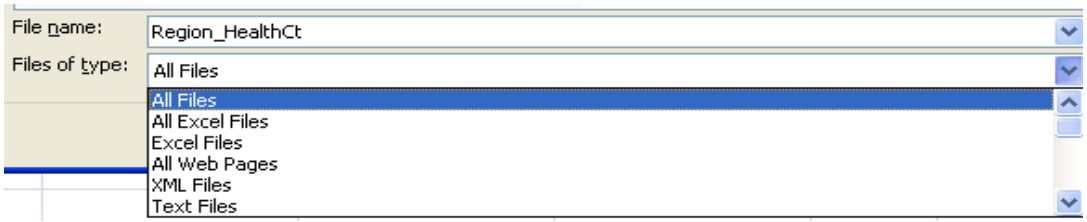
26. Now you will export this data to excel for graphing purposes.

27. Open the “Region_HealthCt” attribute table again. On the bottom right hand corner of the attribute table, click on the “Options” button. From the menu, choose the “Export” option.

28. Export your table as “Region_HealthCt” to your Lab06 folder. Although the file carries a .dbf extension, you will be able to open it in Excel, and save it to as regular .xls file from which you will create your graphs.



29. Go to Microsoft Excel, and from your Lab06 folder open the “Region_HealthCt.dbf” file. Note you have to change the “Files of Type” dropdown to “All Files” in order to see those files with a .dbf extension (see below).



30. Now you can save your table as an .xls file (excel) and graph it if you would like.

Lab 07: Data integration and thematic mapping

Exercise overview

You have received detailed Household Census information from the Central Statistics Agency of Ethiopia. The data is in Microsoft Excel format, and you need to integrate this information into ArcMap to create a thematic map.

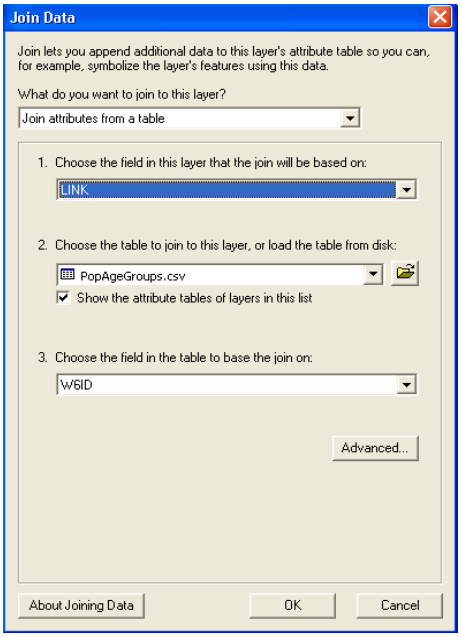
To achieve this you must conduct a **Table Join**. A table join appends attributes of a non-spatial table, to the attributes of a map table. (Non-Spatial means "without geography", i.e.: without map attached). In order for this join to be successful there must be a way to match records in one table with appropriate records in another. This is done with an attribute common to both, such as a name or ID code.

Data integration

1. Open a new session of ArcMap. From your Lab07 folder, add the layer "Woreda_pop".
2. Right click on the "Woreda_pop" file to open the "Attribute Table". This table contains some basic demographic information, including population count and population density.
3. Open Microsoft Excel. Go to File > Open > navigate to your Exercise07 folder and open the file: PopAgeGroups.xls
4. You will see that this table has more population variables disaggregated by age group.
5. In order to import this information correctly into ArcMap, a number of formatting rules must be observed:
    - ArcMap will NOT accept field names longer than 11 digits. If field names are longer than 11 digits, they will be automatically truncated.
    - ArcMap will NOT accept spaces in field names; this will usually result in an error opening the file in the ArcMap environment. (Use an underscore _ instead of a space).
    - Numerical fields must be designated as such, and text fields must be designated as such. (Field's carrying general formatting are open to interpretation in the switch from a regular .xls file to a .csv or a .dbf file. This is an excel issue, not a GIS issue)
    - Occasionally, data is prone to truncation in the switch from a regular .xls file to a .csv or a .dbf file (again, this is an excel issue, not a GIS issue), therefore it is advisable to over-widen the columns to allow for slight truncation.
6. Make sure that the table is formatted correctly with only 11 characters for each variable name. When finished, remember that **you must resave the table to a .csv file** in order to import it into ArcGIS (make sure none of your fields were truncated during the save).
7. Once you have saved the file in .csv format into your Lab07 folder, **close excel and return back to ArcGIS.**
8. Now add your database file "PopAgeGroups.csv" from
9. You will notice that the original excel table "AgePopGroup.xls" is also visible in the "Add Data" window. Newer versions of ArcMap have incorporated the ability to read regular .xls files, and their individual worksheet components. Unfortunately, beyond the read capability, we have found this new function to be slightly inconsistent, and basic data management operations (such as table joins) seem to work best using .dbf and/or .csv files.
10. Right-click on your "AgePopGroups.csv" table to open it. Make sure that all field names carried through correctly, and that all data appears in working order. If it looks good, close the attribute table and move on!
11. The next step is to join the "AgePopGroups.csv" table to the "Woreda_pop" layer so that you can utilize its spatial properties to visualize the population/poverty info.

➔ It is always preferable to use codes, rather than place names to conduct joins. Place names can vary in spelling and accent (which contribute to the unique nature of a particular name), and these may not always transfer from one software to another.

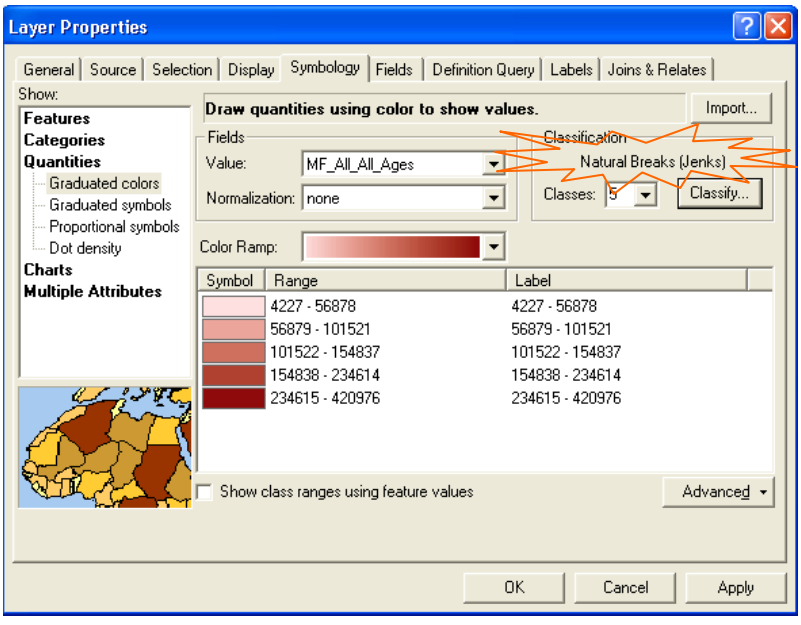
12. In this case, the joining variable is called “LINK”.



13. Right click on the “Woreda_pop” layer, Scroll to “Join & Relates” > “Join”. Make the following dropdown selections, and click OK.
14. When complete, open the attribute table of the “Woreda_pop” layer to make sure the join was successful. As you will see, some of the fields will say <null>, this is okay for this specific join because data were not collected for these specific Woreda.
15. Before continuing, save your map to your Lab07 folder as “Lab07_YOURNAME”.
16. Double click on “Woreda_pop”. The “Layer Properties” Window should now pop-up. If the tab “Symbology” is not selected, then select that tab.
17. In the box labeled “Show”, select the option “Quantities”, and click on the sub-option “Graduated Colors”.
18. In the drop down menu “Value:” **scroll to the end of the list**, and choose the field named “MF_All_All_Ages”: This stands for Male and Female all ages.

    **The variable labels follow a pattern, and are disaggregated by ages. Note that each group of variables are labeled as such:

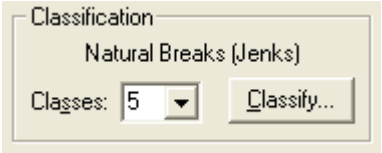
    - MF_All_All_Ages: Male and Female all ages
    - M_All_All_Ages: Male all ages
    - F_All_All_Ages: Female all ages



19. Once this field has been selected, choose a color ramp that you like by clicking on the down arrow next to the color ramp. Now click, Apply, and OK.
20. Look at how the colors are distribued. The classification

brackets chosen by ArcMap are based on their default statistical classification; “Natural Breaks”.

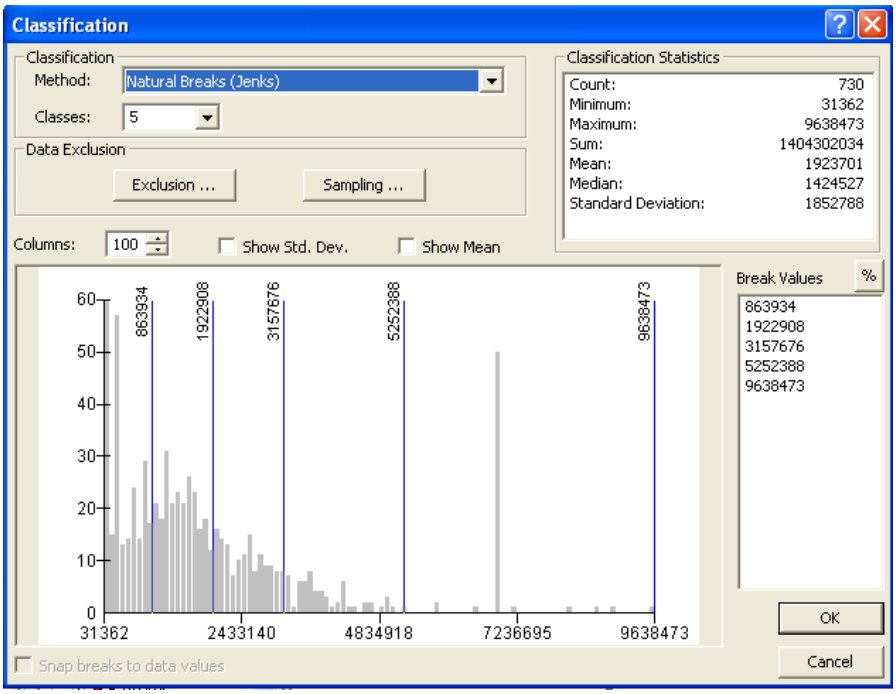
21. Reopen the “Layer Properties” dialog for the “Woreda_pop” layer. Return to the symbology tab. Under the Classification menu (top right-hand corner), Click the “Classify” button.



22. In the “Classificaton Wizard” you will see a histogram illustrating the data distrubution along the number line.

23. In the “Method” drop down list, you will see several classification alernatives to the “Natural Breaks” system. You also have the opportunity to change the number of classes that you use. (See graphic below)

24. Experiment with the different classification schemes, and look at how they alter the classification breaks (blue lines) on the histogram data. By clicking OK on both wizards, you will see the effect of your class scheme changes on the map itself.

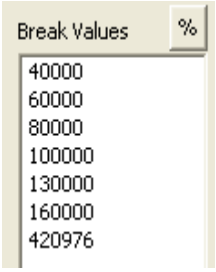


25. For something as simple as population count, “Natural breaks” is not a bad starting point. To make the interval ranges a little more “user friendly” it is advisable to begin with “Natural Breaks”, then switch to “Quantiles”, and modestly round up/down the break values of each category.

26. Obviously, this decision will be determined by the nature of your data, and a classification method such as “Standard Deviation” may be more appropriate in certain cases.

27. Return to the “Classification Wizard” screen. Choose first the “Quantiles” scheme, next change the number of classes to 7.

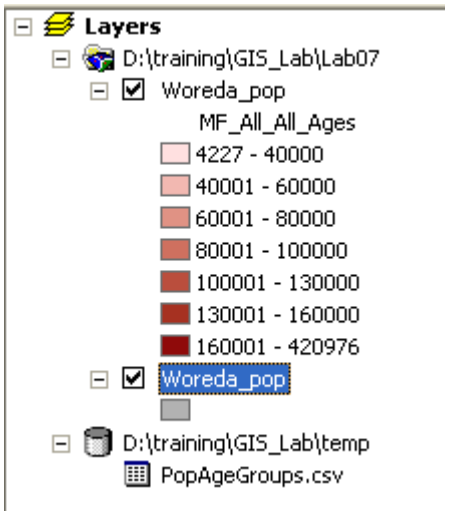
28. In the “Break Values” box to the right hand side of the wizard, set the break values to the following numbers, by simply typing over the existing values.

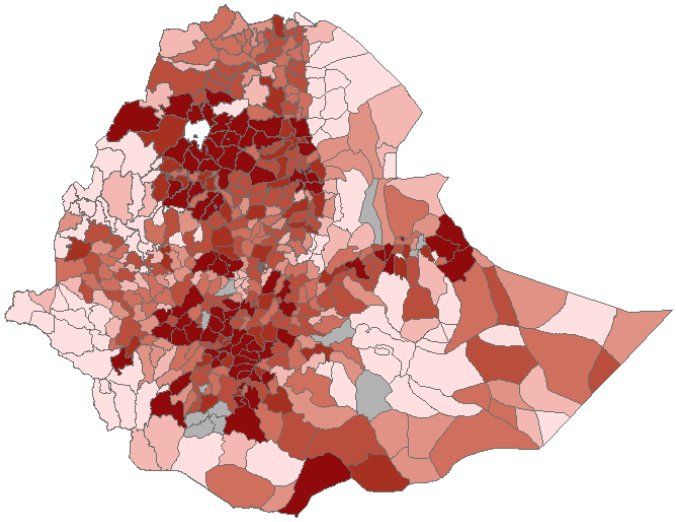


29. When done, click OK. Now you are back at the “Symbology” window. The “Label” side of the menu will reflect the changes that you make to the “Range” side, but you may also use text in your labels.

30. There are a number of Woredas with “Null” or “0” values. (You should always account for these in your mapping). Given that we have only done a Table Join, and we haven’t exported our data as a new shapefile, our values reflect “Null”, we will show these by adding another layer.

31. Add the “Woreda_pop” layer again. Organize your “Table of Contents” so they look like the graphic below. Also, choose the underlying Woreda_pop layer to be a grey color to reflect “Null” data.

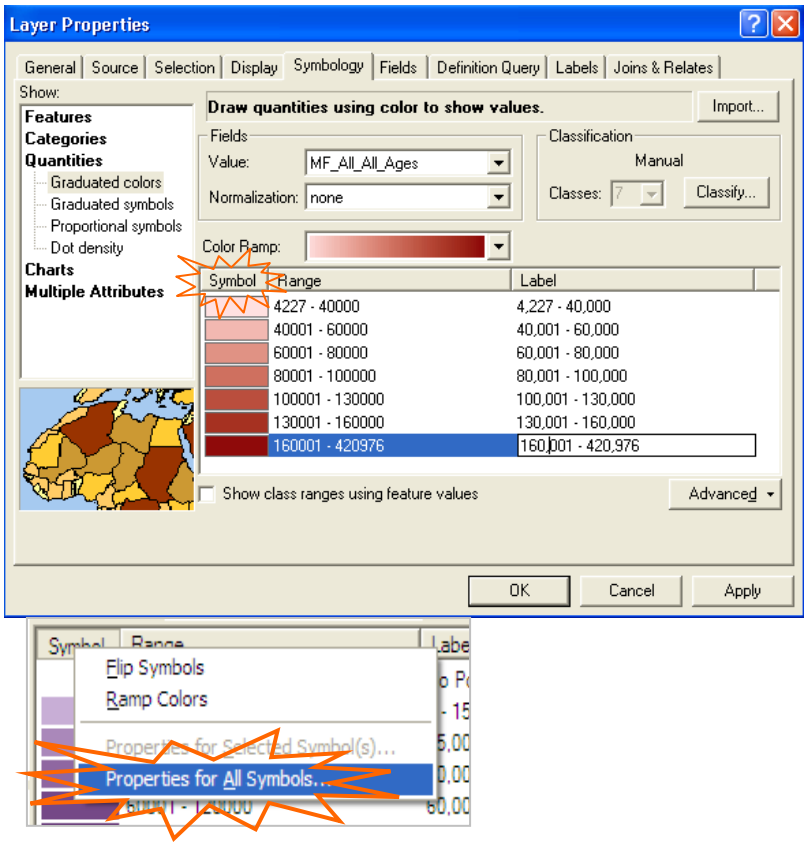




32. Now return back to your “Woreda_pop” layer with the table join. Double click on the layer and return to the Symbology tab.

33. Modify the label options with some additional text, and commas by clicking on the value under the Label column. (This will determine the look of your legend, and reads better than the default categories).

34. For a softer more subtle style, you will remove the boundaries from between the individual woredas. Click on the word “Symbol” above the colored category symbols, and in the pop-up menu, choose “Properties for all symbols”. (see right)



35. In the “Symbol Selector” dialog, change the outline color to “No color”.

36. Click OK.

37. To distinguish the boundaries of the higher order administrative units, add the “Regions” layer from the Lab01 folder, and symbolize as hollow with an appropriate outline thickness. Remember how to do this? (hint: double click on colored box symbol below the layer name in the Table of Contents)

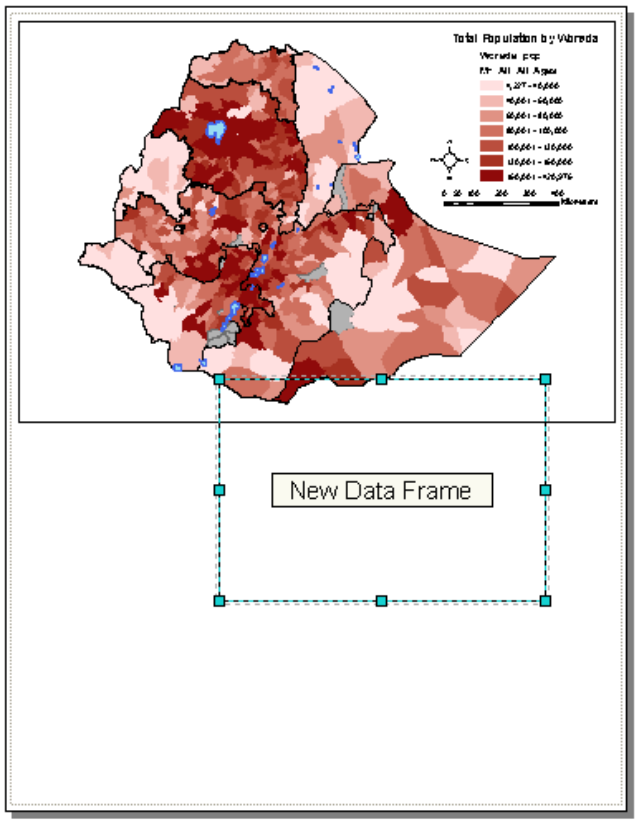
38. Switch to the layout view, by clicking on the layout symbol in the bottom left and corner of the map window. 

39. Using the “Insert” menu, add a legend, and other appropriate cartographic elements to your map (North Arrow, and simple scale bar)

➔ **Save your map if you haven’t recently!**

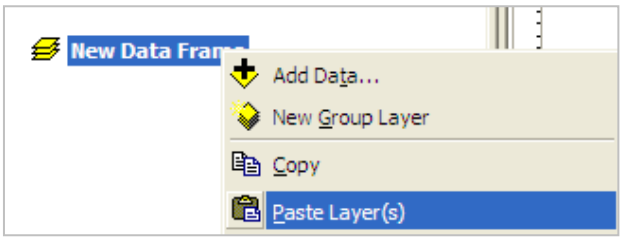
40. In the layout window, resize the map, so that it is roughly one-half the size of the page by clicking once on the map and then dragging one of the corner sizing squares to diminish the size of the map. (You will also have to resize the legend and other elements). Move them to the top right hand corner of your layout.

41. From the “Insert” drop down menu, choose the “New Data Frame” option. This will add a new empty map window to your layout. (see right)



42. Now to add data to your new map window. In the table of contents, right-click on the “Woreda_pop” layer in your first Data Frame that you have been working with up until now and scroll to “Copy”.

43. Now right-click on the “New Data Frame” listing, and scroll to “Paste”. Repeat the process to add the “Region” layer to your second map window.

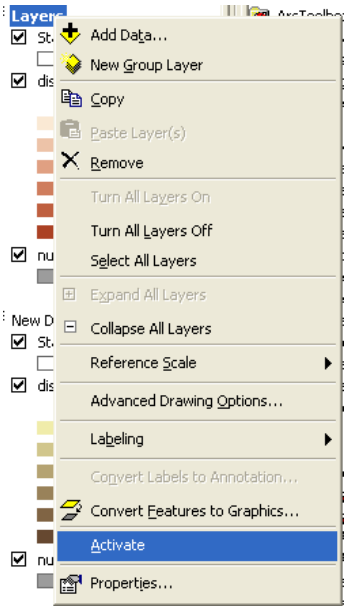


44. Right now, each window looks identical. To help distinguish between the windows, double click on the “Woreda_pop” layer for your new map (leave the first map as is), and return to the “Symbology” tab of “Layer Properties”. In the “Show” box click once on the word “Quantities”, and then “Graduated Colors”. In your “Value” drop down list, choose any of the variables you are

interested in the drop down list. You can choose a specific age group, or if you want, to aggregate population at a specific age group, you can add a field and sum columns. I chose “MF_All_30-34” to map “total population of ages between 30-34”, but it may be interesting to look at male versus female population.

45. Once you have chosen your variable, press OK to map that variable and see how it looks spatially across the Woredas of Ethiopia. How does it vary from your other map?

46. In a multi-map set-up such as this. Only one map may be considered “Active” at any given time. This means that you can only work on the elements of one map at a time. To “Activate” a map, you can simply click on it in the layout view, or, you can right-click on the title of its corresponding set of layers in the table of contents, and scroll to “Activate”.

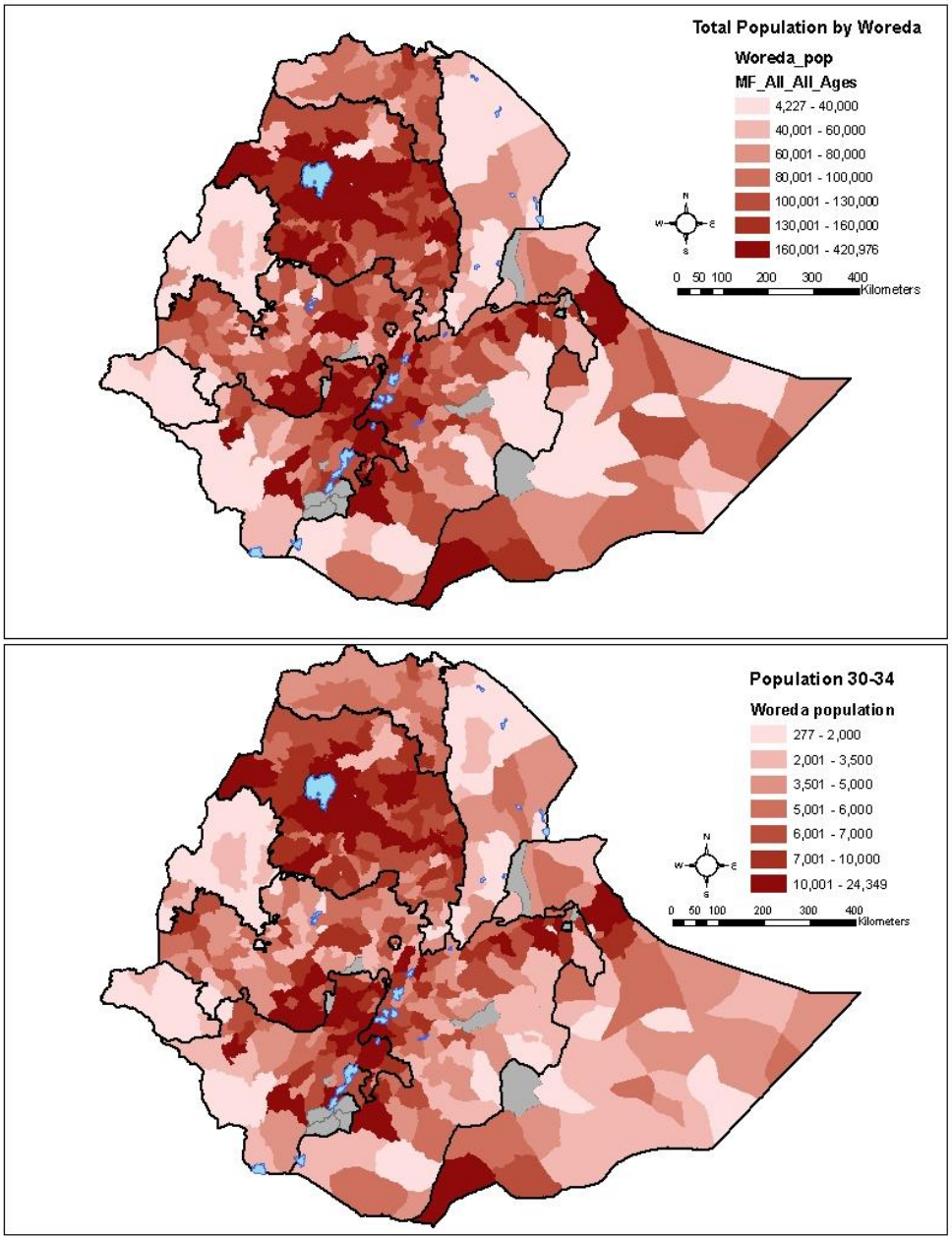


47. On switching to the “Data View” (remember, the little globe symbol),  you will notice that only the active map is shown.

48. Test this by switching to the “Data View” and activating (from the table of contents) each map in turn.

49. After understanding varying “Data View” interfaces, put the final touches on the maps that you chose to make and export them as a .jpeg by going to the “File” button in the upper, left corner. Scroll to export and save as .jpeg in your Lab07 folder!

Your layout window should now look something like this (of course yours may be mapping different variables):



Lab 08: Spatial proximity analysis

Exercise overview

A transportation economist would like to know where gaps exist in transportation infrastructure in order to create a project that facilitates goods to Agricultural Coops. You have been asked to determine distance relationships of each Agricultural Coop to a major road.

Spatial Join – Measuring distance

1. Open a new ArcMap session.
2. From your Lab01 folder, add the “Regions”, “Lakes” and the “Cities” layers. From your Lab02 folder add the “Roads” layer. From your Lab08 folder add the “AgCoop” layer.
3. For the moment, turn off the “Cities” layer. First you will determine the distance of Agriculture Cooperatives to a major road.
4. The roads layer contains various types of roads by surface type (SURFACE_TY). We are mainly interested in the primary roads, so we will create a new shapefile by “Selecting by Attribute”. The abbreviations in the roads dataset are as follows:

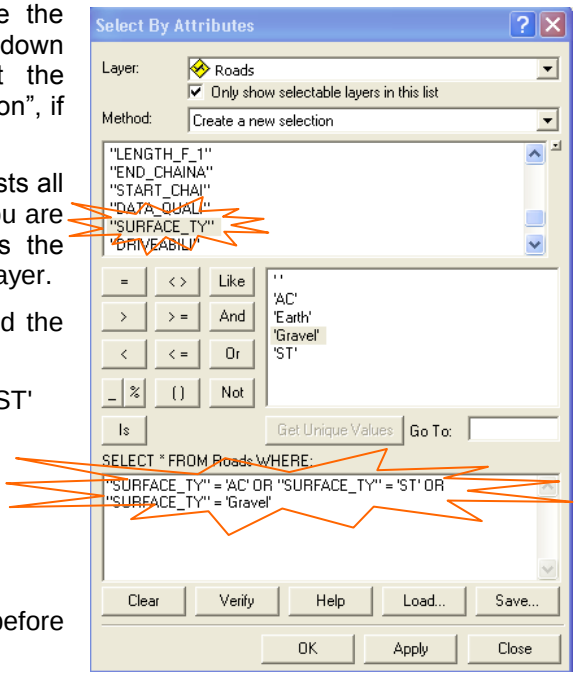
Abbreviations in Roads Dataset:	**SURFACE_TY**
AC	Asphalted Concrete (Standard road)
ST	Surface treatment
Earth	Earth roads
Gravel	Gravel roads
-	Unclassified roads

   For this exercise, we will assume that primary roads are the ‘AC’,‘ST’, or ‘Gravel’ classified roads.

5. From the “Selection” drop down menu, choose the “Select by Attribute” option. In the “Layer” dropdown menu, choose the “road” layer. By default the “Method” option should be “Create a new selection”, if not, choose that.
6. The scroll menu under the “Method” drop down, lists all the fields in the Road attribute table. The field you are interested in is “SURFACE_TY”. This indicates the relevant road type for each road segment in the layer.
7. Using the field names, and query operators, build the following query:

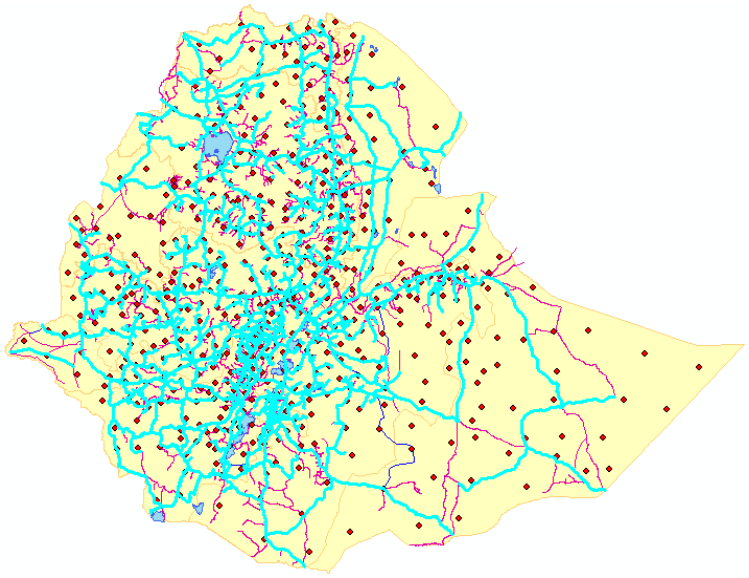
   SURFACE_TY" = 'AC' OR "SURFACE_TY" = 'ST'
   OR "SURFACE_TY" = 'Gravel'

   Your interface should look like the graphic right.

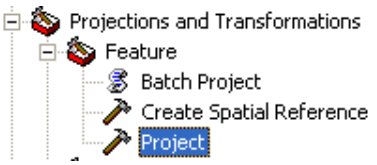


8. You may wish to check the syntax of the query before you finish. Then, Click OK.

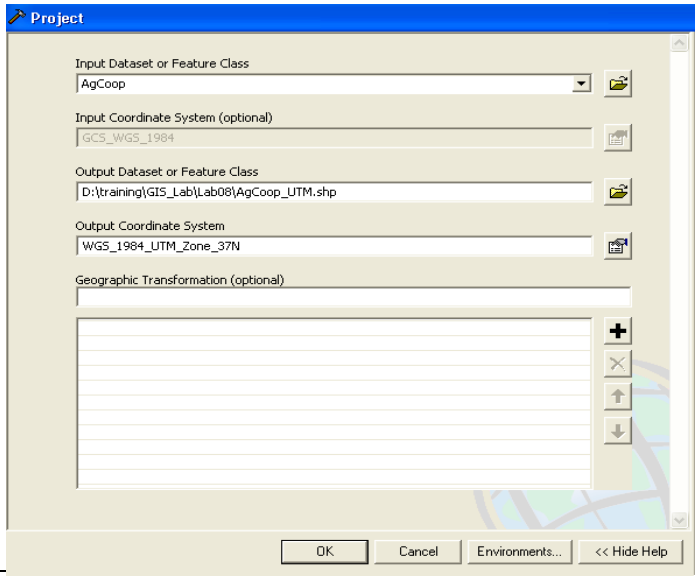
9. The primary roads in Ethiopia should now be highlighted in your map (see below)



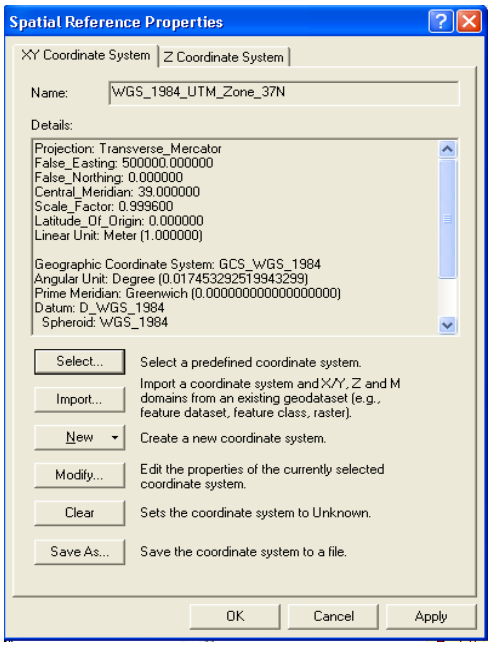
10. To export this selection as a new file, right-click on the "road" layer, scroll down to "Data" > "Export Data".
11. Save the output to your Lab08 folder as "PrimaryRd.shp"
12. The new layer should automatically add to your map.
13. Now let's take a moment to Save your map file. Save this map as Lab08_YOURNAME.mxd
14. Now that you have identified the geographical extent of your study area, you may remove the original "road" layer from the table of contents, by right clicking on it, and selecting "Remove".
15. Now we will be able to complete a "Spatial Join" in order to measure the distance of each Ag Coop to a primary road.
16. But, remember, that first we need to project our data to a projection that will accurately measure distance in Ethiopia. A good projection to use for Ethiopia is projection UTM 37N. We will need to project our Roads layer and our AgCoop layer.
17. Under your toolboxes, click on the box "Data Management and Tools" and scroll to Projections and Transformations > Feature > Project



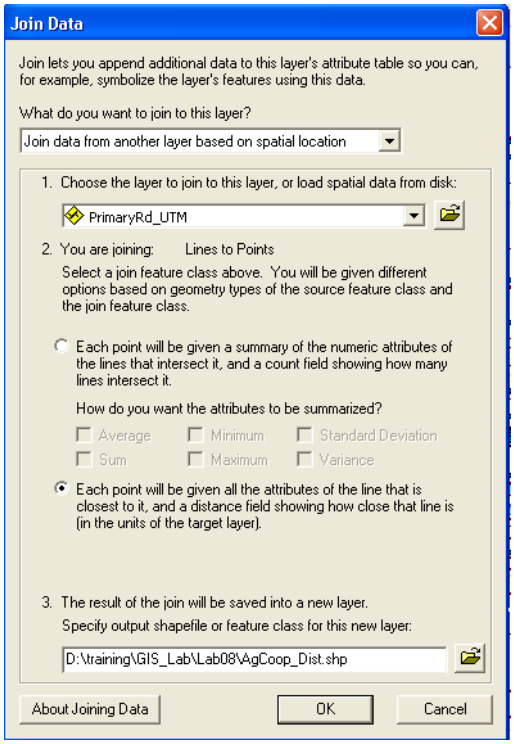
18. Double click on the "Project" toolbox to open the Project window. You will need to fill out this window like the graphic to the right. The Input Dataset and Input Coordinate System will already be filled out for you.
19. In order to fill out the Output Coordinate System, left click on the box to the right and a Spatial Reference Properties window will open, navigate the following path:
    Select
    Projected Coordinate Systems
    UTM
    WGS 1984
    WGS 1984 UTM Zone 37N.prj



20. Press Add after you have found your Zone.
21. The Spatial Reference Properties window will show again displaying the projection that you have chosen (see right). Press OK if it is the correct projection.



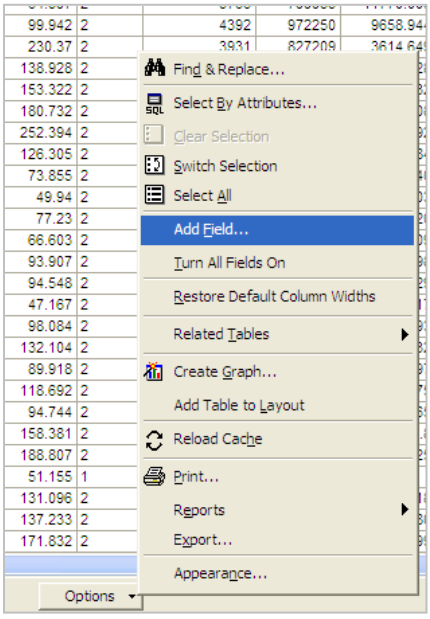
22. Under Output Dataset or Feature Class in the Project window, make sure that you Save your new projected shapefile into your Lab08 folder and save as "AgCoop_UTM.shp" (see graphic to make sure you have filled out each category correctly and then press OK.
23. It will take a moment for the layer to reproject, and then the projected shapfile should add to your table of contents.
24. Now, do the same steps to project your "PrimaryRd" shapefile.
25. When you have both your "PrimaryRd" and your "AgCoop" shapefiles projected into the UTM 37N projection, we can move onto measuring distance!
26. Right-click on the "AgCoop_UTM" layer. Navigate to "Join & Relates" > "Join". In the first dropdown list, change the option to "Join data from another layer based on spatial location".
27. For the Criteria 1 dropdown, choose the "PrimaryRd_UTM" layer.
28. For the Criteria 2 option, choose the second radio button – "each point will be given all the attributes …/…and a distance field showing how close that line is (in the distance units of the target layer)"
29. Unlike a regular table join, a spatial join creates a new file. Save the output file to your Lab08 folder as "AgCoop_dist".
30. Check that your wizard looks like the one to the right, and click OK.



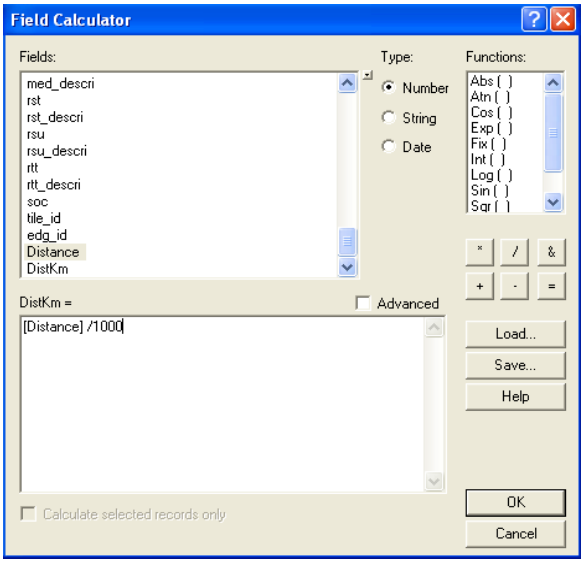
31. The new file will be automatically added to your map window.
32. Right click on the new "AgCoop_dist" file to open the attribute table.
33. Scroll to the far right of the table and you will see a "Distance" field. The units of this field are in meters, as the units of the target file (the AgCoop_UTM layer) was in meters.

34. How do we know this measurement is in meters? To check the units of any of your layers, go to "Layer Properties" for that layer, click on the "Source" tab, and scroll to the bottom of the "Data Source" information. The layer units will always be determined by the layer's projection, which in this case is UTM 37N. UTM projections are always measured in meters. Return to the attribute table of your new "AgCoop_dist" layer.

35. You can create a new field, and recalculate the distances in kilometers, by dividing by 1000. Go to the "Options" tab at the bottom right hand corner of the attribute table, and from the pop-up menu choose the option "Add Field".



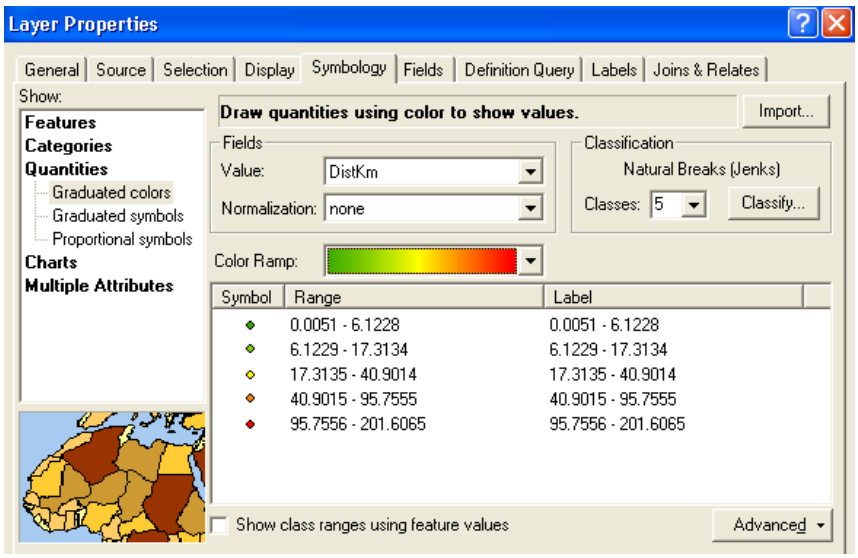
36. In the "Add Field" dialog, name the new field DistKm.
37. For "Type" select "Double"
38. For the "Precision", type 15 (This relates to the length of the allowed cell content), and for "Scale" type 4 (This relates to the number of decimal places allowed from the above 15 spaces).
39. Then click OK.
40. You will notice that a new field has been added to the far right / end of the "AgCoop_dist" table.
41. To populate this field, right-click on the field name "DistKm", and from the pop-up menu, choose "Field Calculator".
42. You may get a warning message to inform you that you are about to perform a calculation outside an edit session. Just click Yes/OK, and the Field Calculator Wizard will appear.
43. You are going to use the values from the "Distance" data fields to calculate values for the "DistKm" field. Using the field names from the scroll list (You will have to scroll down to see the "Distance" field), and the mathematical operators, build the following simple equation (you will need to double click on the field name to add them):
44. When complete, click OK. The "DistKm" field should now be populated with the km distance values for each Ag Coop.



45. You can from 0 to 204 kilometers from a primary road!
46. As a last step, we are going to use our "DistKm" attribute information to color code your mines by distance to the nearest major road.

47. Exit out of the attribute table and double click on your AgCoop_dist layer. Go to the “Symbology” tab.

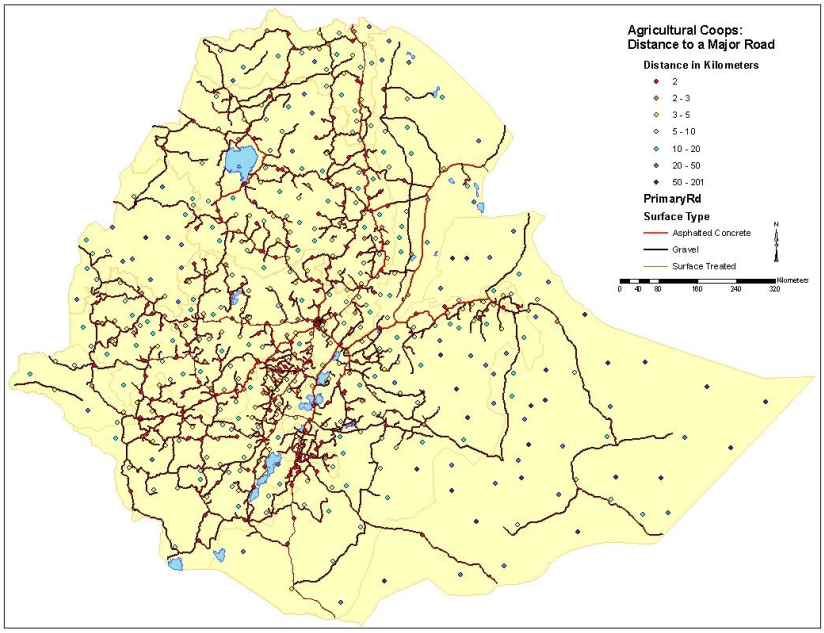
48. You have worked with Symbology before, but in case you need a hint, see below.



49. Now you are ready to put the final touches on your map. Make sure that your “Regions” polygon are an appropriate color, and make sure that your “Primary Roads” are clearly visible.

50. Save your Map in your Lab08_YOURNAME Folder again!

You map should look similar to the below map, of course, you may choose different colors to represent your attributes.



Why might the projection that you used UTM 37N not be the perfect projection for this type of measurement?

What would be a better projection?

Explore the different options in the projection dropdown menu to see if there are projections specifically for Ethiopia or the region.

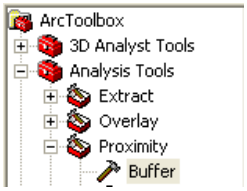
Lab 09: Buffering for distance calculations and area definition

The Buffer tool is used to identify or define an area within a specified distance around a feature. For example, you may create a buffer to define an area around a river to identify land that should not be developed, or, you may want to create a buffer to select features within a specified distance of a feature (Rural Access Indicator – 2 km from a road).

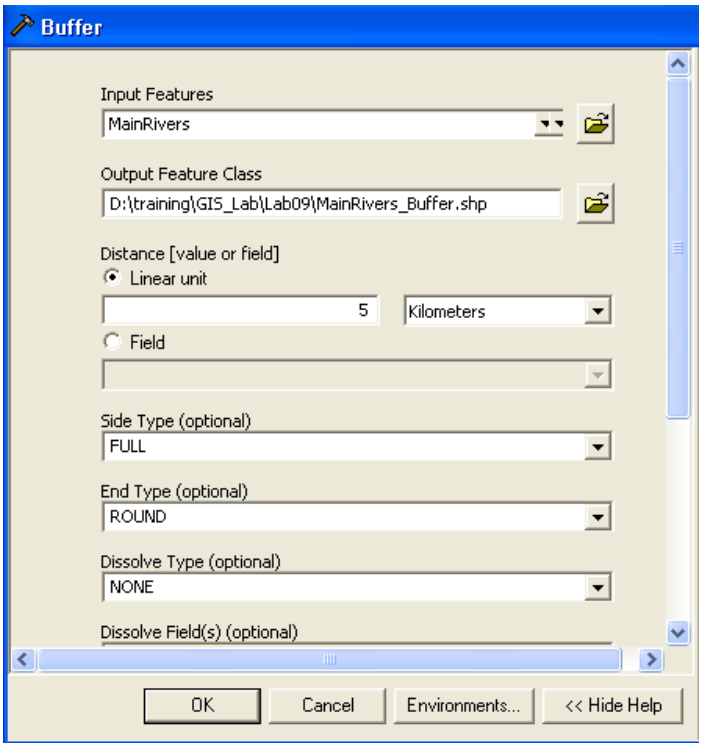
Exercise overview

In this section we will create exclusion zones around the rivers and streams in Ethiopia. A national agency wants to evaluate areas where natural water areas should be cleaned to preserve freshwater and increase water fowl. and would like to know how many cities are within 5 km of a stream or river, and what the total population of these cities is in order to analyze affected population and model economic impact.

1. Open a new ArcMap session. Click on the new map icon  .
2. Add the “Cities” and “Regions” from your Lab01 folder and the “Main_Rivers” from your Lab09 folder.
3. Since you are starting a new Exercise, save your mapfile as: Lab09_YOURNAME.mxd
4. Activate ArcToolbox by clicking on the toolbox icon. Choose Analysis Tools > Proximity > Buffer.

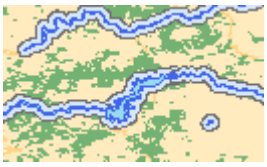


5. Populate your buffer wizard following the example to the right, making sure to save your output to your “Lab09 folder” as MainRivers_buffer”. (This should be the default name).
6. Click OK.

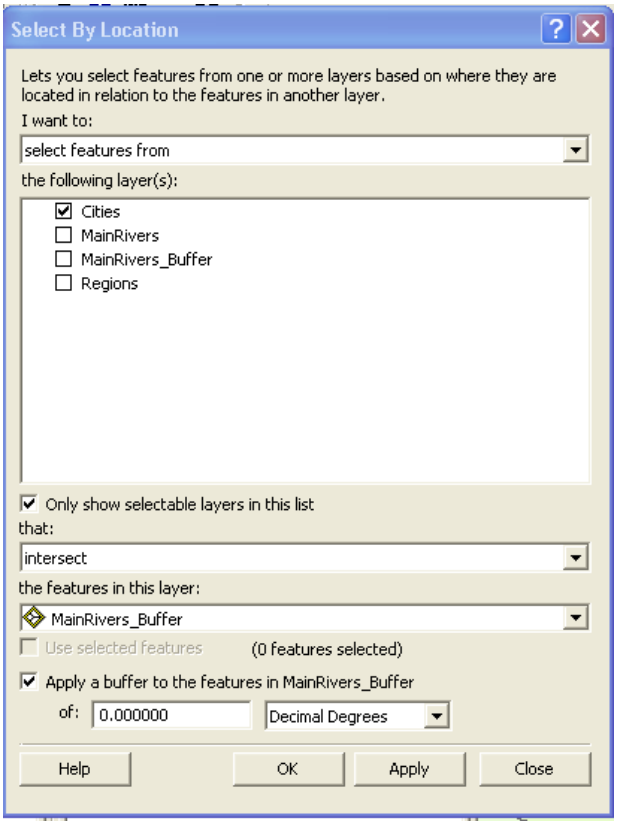


7. The output should automatically add to the data frame.
8. Now you have determined the 5km buffer zone.

9. If you display your “Rivers” layer on top of the buffer layer and you zoom to a specific area, you will see how the buffer has created a 5 kilometer area around the river (see right).



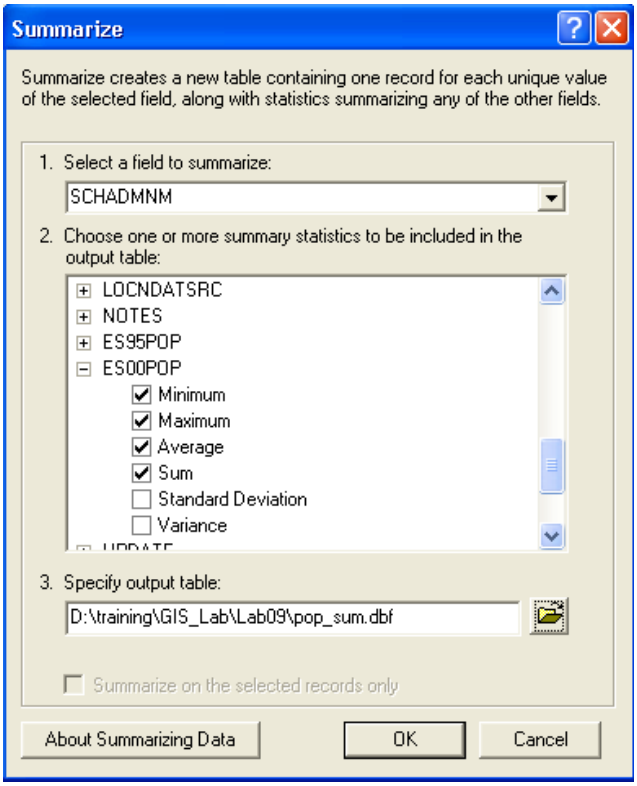
10. Now you will now isolate the communities that lie within this 5 km of this buffer zone.
11. Go to the “Selection” dropdown menu, and choose the “Select by Location” option.
12. Using what you learned from the previous exercise, select features from “Cities” that intersect “MainRivers_buffer”
13. Click OK.



14. The appropriate cities should now be highlighted in your map.
15. To export these as a separate file, right click on the “Cities” layer, scroll to “Data” > “Export Data”, and save to your Lab09 folder as “Cities_5km”
16. Now turn off all layers except the “Regions”, River_buffer”, and “Cities_5km”.
17. We will now compute some simple statistics to answer the questions proposed at the beginning of the exercise.
18. First let’s see how much population per Region will be affected.

19. Open your “Cities_5km” attribute table. Right click on the column title “SCHADMNM” and click on “Summarize”. Fill out the summarize window as you see to the right.

20. Save your table to your Lab09 folder as pop_sum. When finished, click OK.



21. Now right click on your table that you added to your screen and opent the attribute table and look at the various columns that were created.

22. The Count_SCHADMNM column tells you how many cities per Region are within 5 km of a river. The remaining 4 columns gives you the minimum, maximum, average, and sum of populations in the cities that are within 5 km of a major river by Region. (see below)

Attributes of pop_sum

OID	SCHADMNM	Count_SCHADMNM	Minimum_ES00POP	Maximum_ES00POP	Average_ES00POP	Sum_ES00POP
0	ADDIS ABABA	1	119897	119897	119897	119897
1	AFAR	4	3340	22578	12886.75	51547
2	AMHARA	31	373	67729	10279.4516	318663
3	GAMBELA	2	1376	24990	13183	26366
4	HARIAI	1	104511	104511	104511	104511
5	OROMIYA	65	506	31808	7848.3231	510141
6	SOMALI	2	5189	90030	47609.5	95219
7	SOUTHERN	39	136	94647	13129.9744	512069
8	TIGRAY	10	1156	24841	8488.4	84884

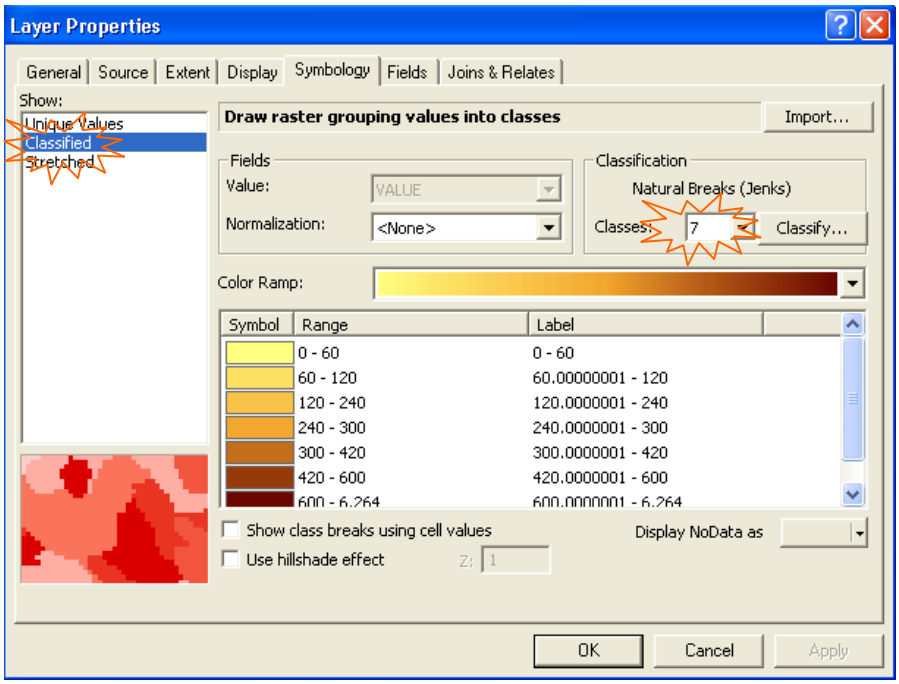
Record: 1 Show: All Selected Records (0 out of 9 Selected) Options

23. Save your project again. Now you can export your table, and relax for a moment before starting the last exercise.

Lab 10: Travel time raster and market access visualization

Understanding distance to key markets is important in understanding possible marketing and import / export opportunities. There are several ways to measure distance. In the previous exercise, we measured distance by kilometers to a primary road. In this exercise, we will measure distance by travel time. For some analyses, this may be a better measurement because you take into consideration road types and conditions, as well as travel speeds.

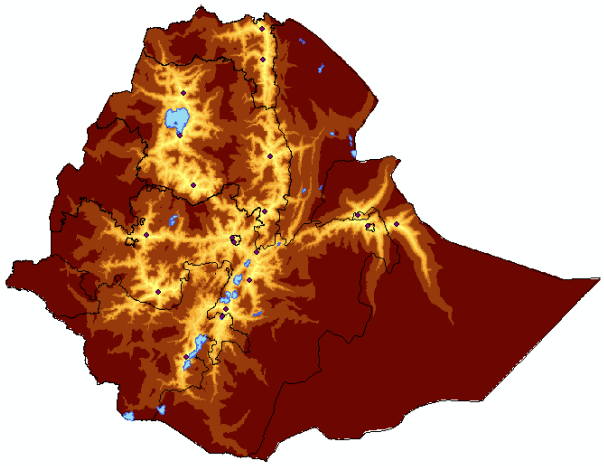
1. Open a New ArcGIS session.
2. You may need to enable an extension in order to complete the following analysis, so go to the toolbar and click on “Tools”, scroll to “Extensions” and enable the “Spatial Analyst” extension.
3. From your Lab01 folder, add your “Regions” shapefile and make it hollow with a black outline (double click on the colored square below the layer name, this will take you to color options).
4. Add the shapefile “tt50k” to your data screen from your Lab10 folder.
5. In order to better understand this shapefile, you should “classify” the values in the “Symbology” tab within the “Layer Properties” window. See example to the right.
6. Click OK when your “Layer Properties” window looks the same as the one to the right.



7. Now, on your dataframe you are looking at a travel time raster dataset that was built using a Cost Distance function.
8. The Cost Distance function calculates accessibility to a given target (in this case, cities of 50,000 people or more.). In order to extract these data, we built a raster dataset at a 1 kilometer resolution where each pixel records the time in minutes to travel from a specific pixel to a city of 50 thousand or more.
9. There are several components in this model: A friction grid layer is produced in order to assess speeds over a geographic surface. The Cost Distance function is then run on the Friction grid. Incorporated in the friction grid is:
    1. Primary, secondary, and tertiary roads (each with varying travel speeds)
    2. Rivers and water bodies
    3. International borders
    4. Elevation

10. As you can see, the Southeastern regions of Ethiopia have very low accessibility. This is due to the fact that there are limited transportation networks in these areas, but also because there are very few cities of 50,000 people or more in Somali region.

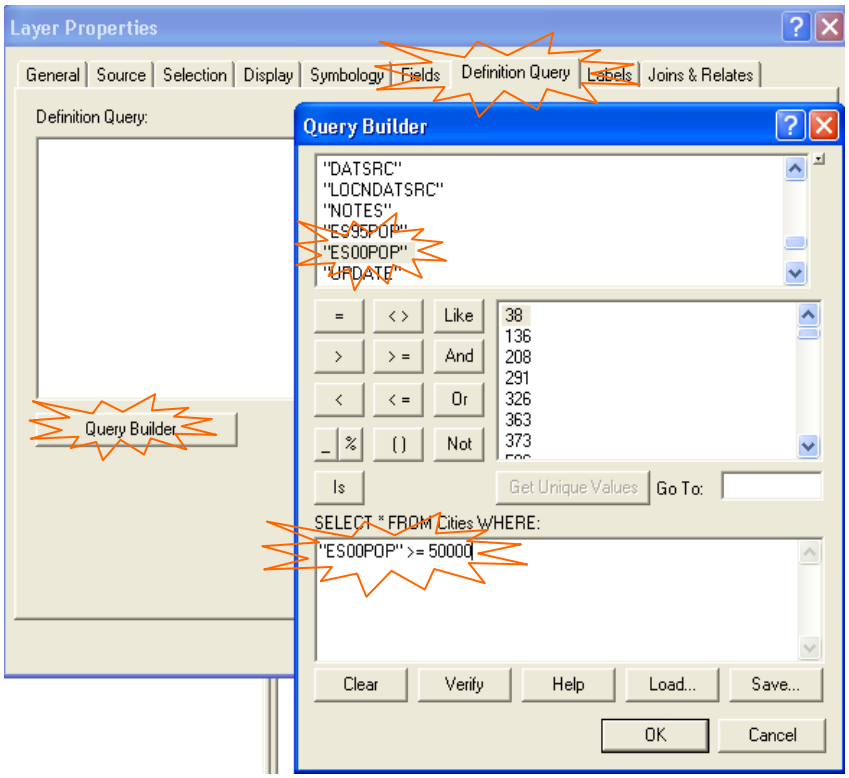
11. Now we will use this raster to do some visualization and analysis.



12. Add your “Cities” layer from your Lab01 folder. Perform a selection query that will show only the cities of 50,000 or more.

13. In order to perform the selection query, double click on the “Cities” layer and click on the “Definition Query” tab in your “Layer Properties” window.

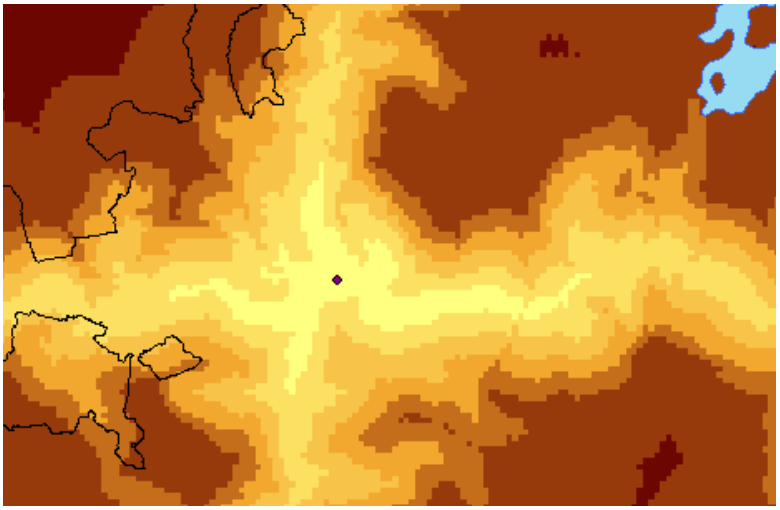
14. Click on the “Query Builder” tab in the lower left corner of the “Definition Query” window and build the query as seen in the graphic right:



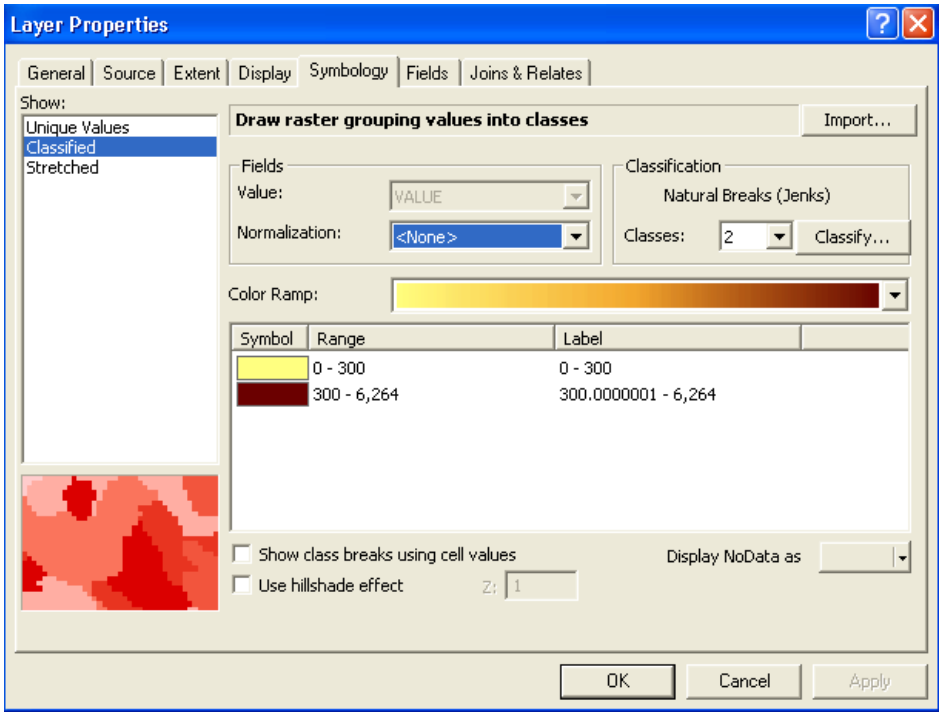
15. When finished, click OK several times until you return to your main data frame. You should now see only the “Cities” greater than or equal to 50,000. By doing a definition query, you are not changing your dataset, but merely isolating and displaying the features you are interested in. (If you want to see all of your “Cities” again, enter into the “Query Builder” and delete the ‘ES00POP’>=50,000 query. Then press OK and you will see all of your data again)

16. For now, let’s keep just the 50,000 population “Cities” in our dataframe.

17. Zoom into one of the cities, do you see how it is the target of the Travel time grid? You can see that at the city center is the target input to the Travel time grid.



18. So, let's change our "Symbology" again to see what areas are within 5 hours travel time to a major city of 50,000 people or more. Do you remember how to do this? (Make your "Symbology" tab look like the below)



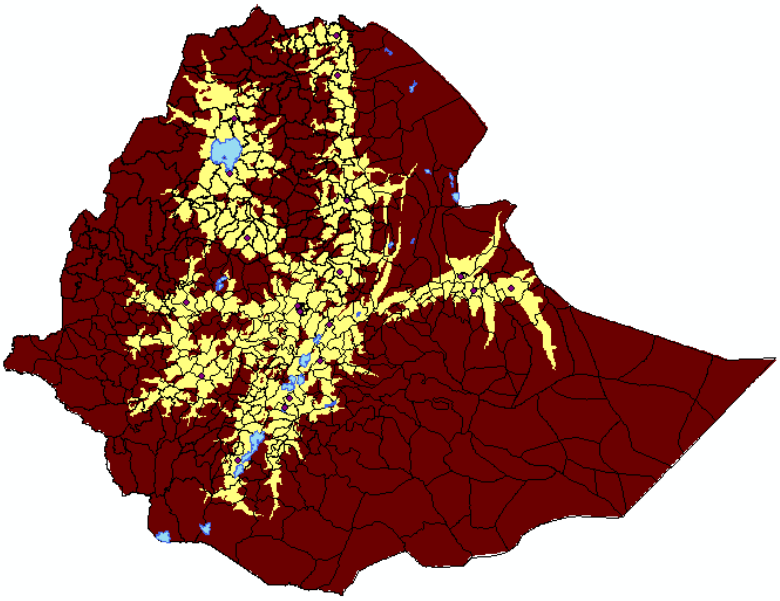
19. Why did we choose 300 as a break off point? Well, because this grid is calculated in minute travel time, so there are 300 minutes in 5 hours. Now we can see which areas within in the country have access to a 50k city within 5 hours.

20. If you would like, you can add your "HealthCenters" or "AgCoop" layer and see how well-connected "HealthCenters" are from major cities.

21. Now we will run a "Zonal Statistics" command in order to average travel time to nearest city of 50,000 population by Woreda.

22. Add your “Woreda” layer from your Lab03 folder to the data screen and make the color of the layer hollow with a black outline (double click on the colored square below the layer name, this will take you to color options).

23. Now you can see by Woreda the travel time grid and which Woredas are connected to a major city within 5 hours travel time.



24. Now go to the Zonal Statistics tool under the Spatial Analyst > Zonal toolbox.

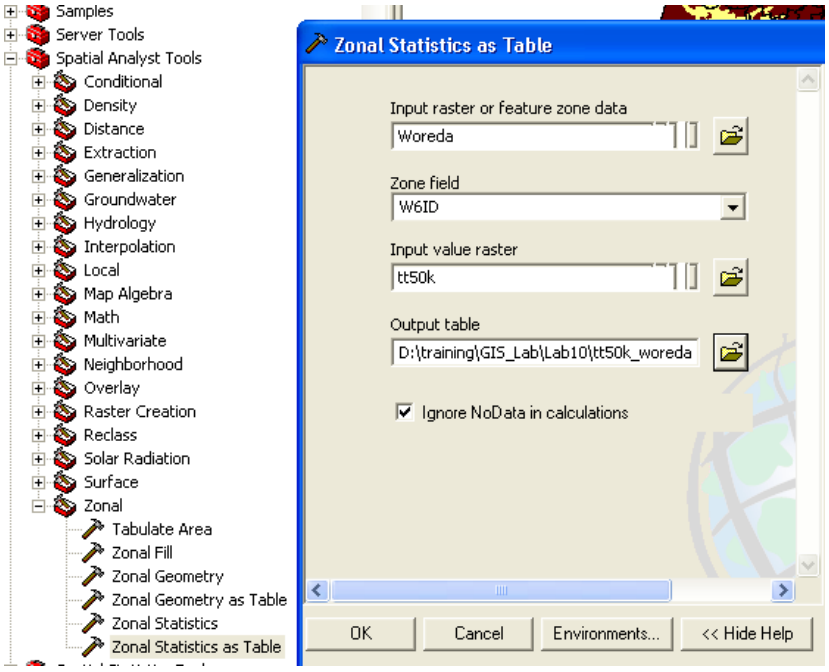
25. Open the “Zonal Statistics as a Table” tool.

26. Fill in the window as you see to the right. We want to average travel time by State.

27. Make sure to save your table in your Lab10 folder under your name. Save as: tt50k_woreda

28. When finished, press OK.

29. Zonal Statistics will run for awhile.

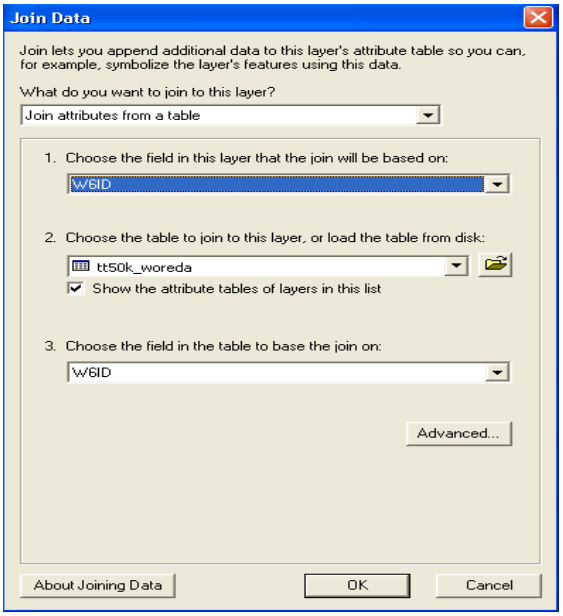


30. Add your zonal statistics table from your Lab10 folder to the data frame and then open the Attribute table. You will see that the travel times to the nearest city of 50k or more have been averaged by Woreda. The Minimum, Maximum, Range, Standard Deviation (STD) of travel time by Woreda has been calculated as well.

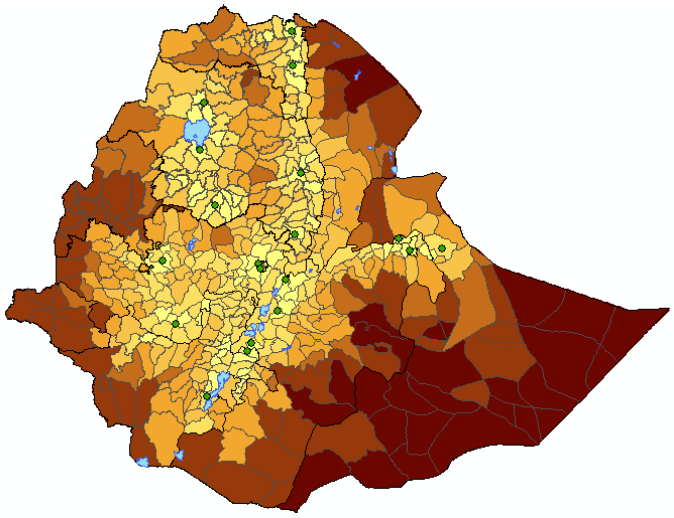
Attributes of tt50k_woreda

Rowid	W6ID *	ZONE-CODE	COUNT	AREA	MIN	MAX	RANGE	MEAN	STD	SUM	VARIETY	MAJORITY	MINORITY	MEDIAN
1	010102	1	3819	3819000100	351	1044	693	623.87695	139.86115	2382586	628	500	351	600
2	010103	2	1825	1824999900	275	862	587	515.0603	127.37447	939985	506	592	275	517
3	010302	3	876	876000000	255	1288	1033	628.42236	235.53993	550498	534	550	268	578
4	010202	4	1292	1292000000	45	560	515	235.85527	120.85897	304725	428	197	45	220
5	010301	5	610	610000000	8	356	348	122.43771	76.489388	74687	250	98	8	110
6	010201	6	1213	1212999900	122	711	589	336.98105	119.88858	408758	445	250	122	328

31. You can merge these data to your “Woreda” shapefile as you did in Lab07 (Part 1, Step 11.) See right for a hint



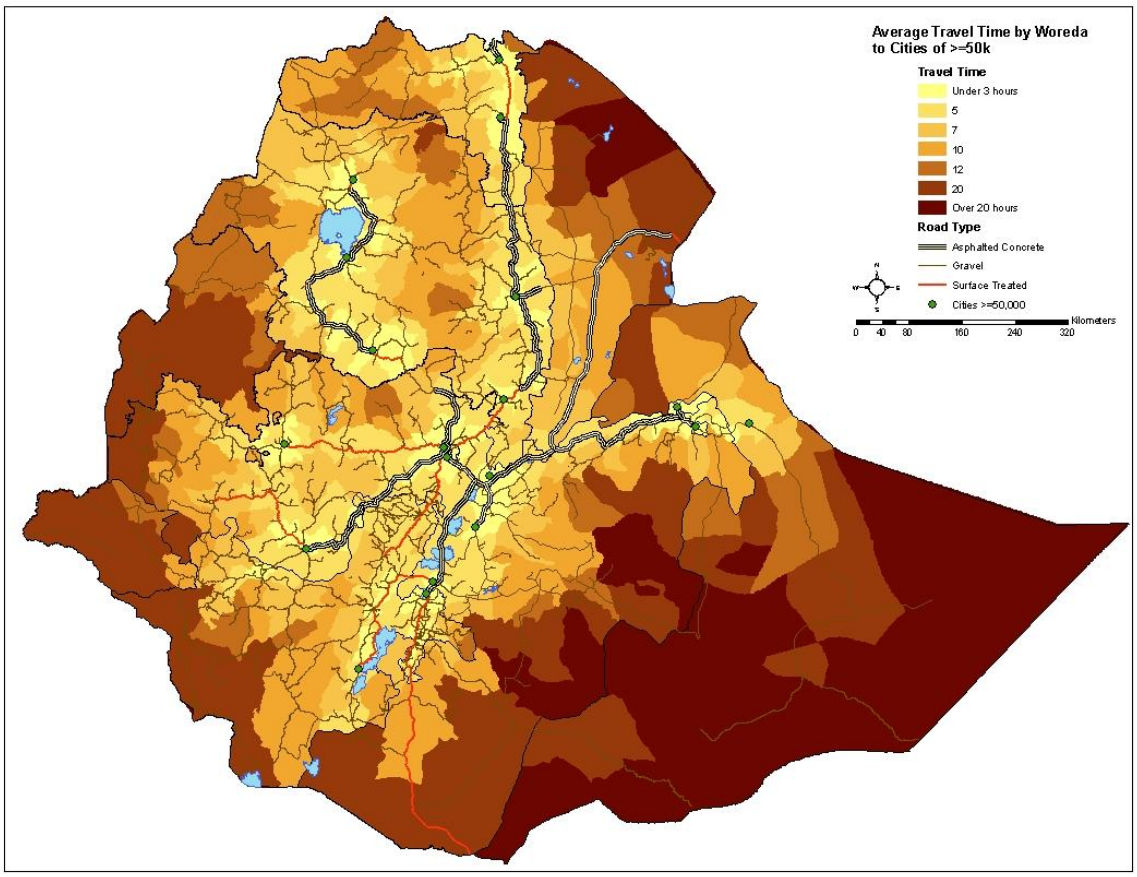
32. Then you can symbolize the data by color to view Woreda differences of market access. See map right.



33. You can do the same type of table merge with the “Regions” or “Zones” layer and create statistics on a more aggregated level.

34. Do the same analysis using your Zone shapefile and compare the two maps spatially to see if you gain any further knowledge.

35. How you choose to display and analyze data spatially will change your results considerably depending on the richness of your data source. Try to think through what challenge you may have spatially, and how you can overcome those challenges by the millions of interesting geospatial datasets that exist, and that you can now create!


Average Travel Time by Woreda
to Cities of >=50k
Travel Time
Under 3 hours
5
7
10
12
20
Over 20 hours
Road Type
Asphalted Concrete
Gravel
Surface Treated
Cities >=50,000
N
W
E
S
0
40
80
160
240
320
Kilometers