Assignment 6 - Basic Spatial Analysis

Assignment 6: Project Data Preparation and Basic Spatial Analysis

Format: Word document with graphics. Place on Student Work Zone section of the wiki by midnight of due date

Assignment Goals

The objective of this assignment is to help you prepare your data for your project, to get to know your data sets, and to practice using some of the basic processing and analysis tools with your project data.

This assignment assumes you are focusing in on a particular area for your project and that you know where to find data sets for it. The assignment requires the use of at least four GIS data sets, including at least one polygon data set and one line data set. By GIS data sets, we mean geographic data (shape files, geodatabases, raster-based data). So use data for your project area if you can. If you do not yet have data for your project, you can use existing data for a place, region, or country. Here are some suggestions:

- **US Projects**
  - Work with a state and/or city GIS data clearinghouse - see the Tufts guide to GIS Data for US Cities and Metropolitan Regions for ideas
- **International projects:**
  - You could use data downloaded for a country from the Humanitarian Response: Common and Fundamental Operational Datasets Registry (go to the country/region tab to download data by country). You could also include the Global Landcover 2000 product; Grided Population of the World data set; if available; and/or data we have on the M: drive (e.g., ESRIDataMaps10 – world…) or the Tufts Geodata Portal

What to Turn In

Turn in a paper that documents the following - apply questions 2 and 3 to at least four data sets:

1. Create a folder structure to effectively organize data for your project - list it out in text or diagram format or provide a screen shot
2. Decide what would be an appropriate coordinate system for your project, if necessary, project your data sets into that coordinate system. For documentation, provide the projection information (name, projection type, linear units, and create a screen print of the Source tab’s coordinate system section (including linear units) for each of your four data sets). Tips:
   a. Use the DEFINE PROJECTION tool to project data into a new desired coordinate system
   b. Only use the DEFINE PROJECTION tool if you have a data set that is missing a spatial reference and you know what coordinate system it is in
   c. Also note that there is a DEFINE PROJECTION tool for raster data. Click on this graphic for guidance.
   d. IMPORTANT: Also set your Data Frame to your desired coordinate system before doing analysis!!
3. Clip and/or select as necessary to reduce your data sets to just your project area of interest - create shape files out of these. For documentation, export a graphic of your map showing the data sets reduced to your area of interest. Here is an example of clipped data for Uganda. Tips:
   a. The CLIP tool creates a new clipped shape file when you use it
   b. If you use the SELECT tools, you then need to right-click on the data layer and use the DATA EXPORT function to create a shape file from the selected features
   c. Note there is a CLIP tool for features (vector) and a CLIP tool for raster data - use the Search Tool to find the right tool (you can access the Search tool by clicking on the Windows menu and choosing Search). Click on this graphic for guidance.
4. Calculate areas for your polygon data set(s) in desired units (e.g., square km, hectares, acres) and length for your line data set(s). Make sure your attribute table column name for the area has the unit in its name, e.g., Sq Mile or Sq feet or Sq Km or Acres. To document this, create a screen print of the relevant attribute tables with your new area/length fields. Tip: create new attribute table fields for the area/length attributes, then use the Calculate Geometry function - your data must be projected to do this!!
5. Use at least 4 of the following tools to tell us something about your project area - document your answers by telling us what question(s) you asked, what you did to find the answer (you can use screen prints) and what the answer is. Also include a brief explanation of why each answer might be wrong, and if you encountered technical difficulties, give a brief explanation of those.
   a. Select by attribute
   b. Select by location
   c. Viewing statistics in a table
   d. Summarize data in an attribute table
   e. Spatial join (typically used to aggregate point data up to a polygon overlay - for example, creating a count of school points for each district polygon; see About joining features by their location)
   f. Field Calculator (see Adding a Field to a Table and using the Field Calculator)
   g. Any of the Overlay tools
   h. Any of the Proximity tools
   i. Zonal Statistics as Table (this is a Spatial Analyst tool, so you need to enable the Spatial Analyst extension - in ArcMap, click on Customize - Extensions)

NOTE 1: You have to use at least four of the above tools. You could ask one question that requires using all four tools (e.g., finding a suitable site for some purpose) or you could ask four separate unrelated questions using each tool. Or some combination thereof. Up to you! Use this as a way to explore your data sets. We provide some example questions below, but we encourage you to come up with your own.

NOTE 2: You will likely run into problems like missing values or incomplete data sets or other data issues which mean your answers are not correct. You don't have to fix these issues for this assignment! Just give the answer you come up with using the tools, then explain why it is or might be wrong. For your actual project, you'll try to fix these problems.

Useful Resources:
Exercises we have done in class that can help you with this assignment:

- Haiti Basic Queries exercise (the select tools, statistics, summarize)
- Proximity exercise - Nuclear Power Plants in Southern New England (all the proximity tools plus spatial join and zonal statistics as table function)
- Trouble-shooting coordinate systems - exercise (using the Project tool, and Define Projection if you need it)

These tip sheets from our Tufts GIS Tips and Tutorials may be of help:

- Data Management Tips
- Trouble-shooting coordinate systems - exercise
- Creating a smaller data set from a larger one - VECTOR data
- Creating a smaller data set from a larger one - RASTER data
- Select by attribute (see the tip sheet for Using the Selection Tools)
- Select by location (see the tip sheet for Using the Selection Tools)
- View statistics for selected features (see the tip sheet for Viewing Statistics and Summarizing Data)
- Summarize by an attribute field value (see the tip sheet for Viewing Statistics and Summarizing Data)
- Adding a Field to a Table and using the Field Calculator
- Calculating Area, Perimeter, and Length

Examples of questions you might answer - you only need to use four tools so at most you only need to answer four questions!

Questions you could answer about neighborhoods/provinces/districts/tracts or whatever your level of analysis is:

- What is the average population density for your study area? Is there a wide range?
- Where are the most densely populated areas? The least densely populated?
- How many schools are in your study area? Which areas have the most or least schools?
- How many miles of roads or other infrastructure are there in your study area? Which areas are the best or least provisioned for this infrastructure?
- How many services of other types or businesses are there in each of these areas? Which areas have the most or least services?
- Can you provide a count of each type of service in an area?
- What's the area of each type of land use or land cover in your project area?
- Where are the most high value properties? The lowest valued properties? Can you express this in value per square foot? Do areas close to services have higher-valued properties?
- Where are some suitable sites for a project of some sort? (e.g., affordable housing, urban agriculture, infrastructure investment)
- Where are vulnerable areas for some natural or human-induced hazard? (e.g., close to hazardous sites, close to shoreline, close to conflict zones, away from potential help, lacking resources)

For national or regional analysis:

- How many districts contain major cities (over a certain population threshold)? What's the urban population of a district or province compared to the population not living in cities?
- How many people live near primary roads? What proportion is that of the total population? Can you tell us anything about those people?
- How many people live close to the international border? If you have census data, can you tell us anything about those people?
- How many miles of each type of road are there?
- How many schools are farther than 10 kilometers from a major town? What proportion is that of all schools?
- Similarly, you could provide information about other point data sets – perhaps different kinds of infrastructure or services.
- If you have conflict data (e.g., from ACLED), provide some statistics about conflicts by province, or by proximity to borders or major population centers.