GEOG 4092/5092: Assignment 2
Vector encoding

Due in two weeks—September 17 (25 points)

Goal: Learn to encode vector geometry. Gain familiarity with Shapely, and Geopandas properties.

Data: Download the file *lab2.zip* from Canvas and unzip it into your home directory. The data folder contains three text files and two geotiff raster files. The text files contain coordinates delineating the boundaries of three districts (administrative units) in eastern Tanzania. The rasters are from the 2004 and 2009 GlobCover dataset, which is a global land cover classification with a nominal spatial resolution of 300 meters. The GlobCover layers have been reclassified to look solely at agricultural lands. Each is a Boolean raster where 1 represents land utilized for agriculture and 0 represents all other land cover classes.

Task: You will encode the text files into polygon features. Then you will calculate percentage of agricultural land in each polygon in 2004 and in 2009 using the GlobCover rasters.

Part I:

1) Read the text files and extract the x, y coordinates for each district.
2) Create a polygon from each file and create a geodataframe with the following two fields: 'num_coords' and 'district'. populate the corresponding value for the two fields you added. 'num_coords' should contain the number of vertices, or coordinate pairs, used to encode each polygon. 'districts' should be the suffix of each text file (i.e., '01', '05' or '06'). Hint: steps 1 and 2 should be done concurrently as you encode the polygons. TIP: What are the methods to create a geodataframe? Which data structure allows for data schema (pre-defined data structure such as column names)

Part II:

1) Calculate the number of agricultural pixels in each district (i.e., each polygon) in 2004 and in 2009.
2) In a final print statement, report the amount of agricultural land as a percentage of the total number of pixels in each district in 2004 and in 2009 (i.e., print these six values).

Go through the single steps very carefully and make use of the Geopandas and Shapely documentation to find out more about usage and the required. Make efficient use of help forums and blogs as well.

Hint: Which GIS process would you use to calculate the number of pixels within each district? Is there a python package that provide this functionality? If so, how do you install it?

By the start of next class in two weeks upload your script (*lastname_lab2.py*) to Canvas.

Grading: You will be evaluated on the following: your script runs without errors (5); Part I: correctly encodes the district boundaries (9) and adds and calculates the two fields (4); Part II: calculates the number of agricultural pixels in each polygon (2) and prints the correct percentages of agricultural land in each polygon for each GlobCover year (2); and the clarity, logic and efficiency of your script (3).

Optional Challenge: Determine the difference in area of each polygon between the vector and raster (based on the GlobCover resolution) representations of the areas.