

Spatial Analysis and Data Exploration in History and Archaeology, Spring 2021

LDA-H313

FINAL ASSIGNMENT EXERCISE

The final grade of the course consists of two parts, both graded 1-5: the course Learning Report (60% of the final grade) and Final Exercise assignment (40% of the final grade). The deadline for submission of the work via Moodle is Friday 21st May at 23.55.

In the Final Assignment Exercise, you are asked to conduct a series of short analyses on the datasets provided. You are asked to present the results (maps, graphs) and to briefly discuss them, also referring the choices you have made during the analysis. The length of the submitted results file should be 3-5 pages. The text should be submitted as a pdf file. Include your surname and 'SADE_2021' in the file name.

Because pictures/tables form part of the assignment there is no set word limit. You should, however, have a minimum of 500 words of total discussion (roughly one and a half pages of total text at font size 11, margins no wider than 2.6 inches, spacing 1.5).

The assignment has the following goals:

- Demonstrating that the student is able to apply methods and workflows learned to real-world data.
- Demonstrating that the student is able to read and interpret the results.

There is not necessarily a single "correct" answer to the task questions. The aim is to demonstrate you understand the technique you use, you understand the how characteristics of the data can influence analysis, and that you can justify your interpretation.

There are five tasks, each contributing one point on a pass/fail basis towards the final assignment grade of 1 to 5.

You will find four datasets included:

Dataset 0: Polygon shapefile of the historical border of the county of Kent in south-eastern England. You will use it in tasks 1 and 2. You can also use the England DEM from weeks 2 and 3 to help topographically contextualise the point data.

Dataset 1: Point shapefile containing the locations and taxable value of medieval tax 'vills' (usually villages) in the Domesday assessment of 1086 for Kent. Due to lack of accuracy in the original recording process, the points are in a grid pattern with 1 km accuracy.

Dataset 2: Point shapefile of objects from the Portable Antiquities Scheme database that have been found in Kent and dated after AD 800. Due to site protection concerns, the findspot coordinates are only available with 1 km accuracy, therefore also in a grid pattern.

(Except for a handful of exceptions that were moved so they are not outside the county boundaries.)

Dataset 3: Same as above, except given as a CSV file. (ArchSeries does not work with spatial points data frames.)

Dataset 4: CSV file containing a set of Romano-British pottery fragments. Rows are ceramics pieces and the columns indicate what oxides (chemicals) they contain.

TASKS

Include the plots/maps that you discuss.

- 1) Using Dataset 1, plot medieval settlements in the county of Kent, and run K, L and Pair Correlation Function analysis with 99 runs of Monte Carlo simulation on the data. Note that some regions had not yet been reclaimed from wilderness (forest/marshes) in the eleventh century. Include the graphs and a map plot of the point distribution. What does it suggest about the data? [Note: in the simulation function arguments Kest=K-function, Lest=L-function and pcf=PCF]
- 2) Using Dataset 2, plot the distribution of medieval object findspots in the county of Kent. Perform kernel density analysis on the findspots, using a few different search radii. Also perform a relative risk surface analysis, using the object type "seal matrix". Include the kden and relative risk surface maps. What do they show, why would you use this method to examine distributions, and how does the choice of bandwidth value reveal or hide interesting patterns? What scale would you use?
- 3) Perform aoristic analysis on Dataset 3, both on the whole dataset and on just the seal matrices. Include the graphs for aoristic weighing (bar charts) and for aoristic analysis with MC simulation (your choice of line, polygon "blocks colour" or boxplot charts), both for the overall and the seal matrices analyses. What overall temporal patterns does this reveal, and what do they indicate about both historical processes and modern recording processes.
- 4) Perform PCA on dataset 4. The ceramics vessel pieces were collected from sites that are located in a small number of different regions. Assuming that pieces that come from the same region have share characteristics in terms of chemical composition, what patterns does the analysis indicate? Include biplots for the individuals and for the variables (either as separate plots or as a combined plot).
- 5) Perform cluster analysis on dataset 4. Which method did you choose and why, how many groupings would you divide the ceramics pieces into, and how do you think it corresponds with the PCA in tasks 4? Include at least one relevant plot.