Assignment Two: The MAUP and Multilevel Modelling

Contents

1 Demonstrating the MAUP

1.1 Background

Areal units in zoning systems amalgamate into objects that constitute the basic units for the observation and analysis of spatial phenomena (Openshaw 2015). Yet, no gold standard for guiding the spatial aggregation process exists, with the validity of zonal objects subject to the arbitrary and modifiable decision-making of quantitative geographers. Problematically, the analysis of socioeconomic data involving areal units is encumbered by the modifiable areal unit problem (MAUP): "the sensitivity of analytical results to the definition of units for which data are collected." According to the literature, the MAUP constrains the reliability of analyses for aggregated spatial data, as findings have shown varying results with the scale of aggregation and configuration of zoning systems (Avery and Clark 2015).

In practice, the MAUP is condensed into two issues of scale and zoning sensitivity which this paper will attempt to demonstrate in Section 1.2. The first issue, described as the *scale problem*, is the variation in findings when data for zonal units are progressively aggregated. This has been demonstrated empirically by Avery and Clark (2015) who found that whilst correlation coefficients did not increase monotonically with aggregation[1], a general increase in data aggregation corresponds to an increase in correlation coefficients.

The second issue, the *zoning problem*, pertains to the variation in findings when alternative combinations of zonal units are analysed with the scale or number of units held constant (Openshaw 2015). Zoning sensitivity in multivariate analysis has been demonstrated empirically in Fotheringham and Wong (2015) who simulated the aggregation of 871 block groups into 218 zones in 150 different iterations. They highlight the severity of the zoning problem by demonstrating the possibility of concluding with one iteration of zones no association between the percentage of blue-collar workers and mean family income, with another iteration finding a unit increase in blue-collar worker percentages as reducing mean family income by $20,000. In all, ignoring scale and zoning sensitivity in model calibration can lead to inferential conclusions that a researcher's areal data is applicable to the constituents who form the zones under study - the ecological fallacy problem (Openshaw 2015).

[1] At higher levels of aggregation, there is smaller adjacency of zonal units meaning groupings are more heterogenous leading to lower correlation coefficients.

1.2 MAUP analysis

1.2.1 Data

To demonstrate the MAUP issue, we analyse the correlation between property price data derived from Land Registry (**???**)

1.2.2 MAUP findings

** FIND WAY TO CONSTRUCT RANDOM AGGREGATION PROCESS FROM LSOA **

** CORRELATION MAPS **

** REGRESSION PLOTS **

2 Multilevel Modelling

2.1 Interpretation

3 Bibliography

Avery and Clark. 2015. "R: A Language and Environment for Statistical Computing." Journal Article. http://www.R-project.org.

Fotheringham and Wong. 2015. "R: A Language and Environment for Statistical Computing." Journal Article. http://www.R-project.org.

Openshaw. 2015. "R: A Language and Environment for Statistical Computing." Journal Article. http://www.R-project.org.