Politecnico di Milano
Scuola di Ingegneria Industriale e dell'Informazione
Applied Statistics
July 24, 2024

Problem n.1

The shapefile `LSOA` contains detailed information gathered during the Brexit voting process at the Lower Layer Super Output Area (LSOA) level. Each LSOA includes between 400 and 1200 households, with a typically resident population ranging from 1000 to 3000 people, and is univocally identified by the variable `lad16cd`. In the following, let us focus on `Pct_Lev` (the percentage of people who voted for Brexit).

a) Define a **Queen** contiguity-based spatial weight with an order of contiguity equal to 1.

- Report the *mean* and *median* number of neighbors, as well as the *sparsity percentage*.
- *How many* LSOAs are neighborless?

b) Based on the weight defined in a), visualize the LISA cluster Map.

- Report the slope of the line displayed in Moran's scatterplot.
- *How many* and *which* LSOAs are identified as outliers for a p-value ≤ 0.05 when testing spatial autocorrelation? Report the number and their value in terms of the `lad16nm` variable.
- Among the LSOAs characterized by a positive spatial autocorrelation, report the `lad16nm` variable of the LSOA with the highest $\sum_j w_{ij}(z_j - \overline{z})$
- What socio-economic conclusions can you draw from the results displayed in the LISA cluster map?

c) Define now a **Rook** contiguity-based spatial weight with an order of contiguity equal to 1.

- Report the *mean* and *median* number of neighbors, as well as the *sparsity percentage*.
- *How many* LSOAs are neighborless?

d) Based on the weight defined in c), visualize the LISA cluster Map.

- Report the slope of the line displayed in Moran's scatterplot.
- Comment on the differences obtained in the LISA cluster map with the spatial weights in a) and c), explaining the reason for the differences.

Upload your solution https://forms.office.com/e/MQuAJbXaPg