ADF COMPANION DOCUMENT D: Instructions for Calculating the DIS Score

These instructions assume that you have downloaded the data from the UKAN website and have it open in appropriate software. The file is synthetic data but the data structure is that which might typically be found in a census, survey or administrative file.

We are using an eight variable key, which represents information that somebody might plausibly know about a neighbour. Nothing should be read into the details of the results (the data is not real) – the exercises simply serve to demonstrate the technique that you can then use with your own data.

We are using a file where the sampling faction is 10%.

D.1 Instructions for Excel

1. Sort the file by the following columns (checking the 'my data has headers' box is checked): sex, age, ethnic, accomtype, tenure, marstatus, ncars, cenheat. For each column, sort from smallest to largest.
2. Enter the word 'ccount' into cell N1
3. Enter 1 in cell N2
4. Enter the following formula into cell N3
   =IF(AND(A3=A2,B3=B2,C3=C2,D3=D2,E3=E2,F3=F2,J3=J2,M3=M2),N2+1,1)
5. Fill down from N3 to N210745
6. Select and copy column N
7. Right click 'Paste' and pick the values option (ensuring the values are associated with the correct row as you carry out further sorting and calculations)
8. Repeat the sort you did at stage 1, but adding ccount to the end of the list sorted from largest to smallest.
9. Enter the word 'csize' into cell O1
10. Enter the following formula into cell O2:
=N2
11. Enter the following formula into cell O3:

=IF(N3<N2,O2,N3)

12. Fill down from O3 to O210745
13. Switch to the output page tab
14. In cell B2 type the formula

=COUNTIF(BarsettonSample!O:O,1)

15. In cell B3 type the formula

=COUNTIF(BarsettonSample!O:O,2)

16. Enter the sample fraction 0.1 into cell B4
17. Enter the following formula into Cell B5

=B2*B4/(B2*B4+B3*(1-B4))