

How Open Data Resources can be used to define a Composite Measure of Cultural Identity & Heritage for England and Wales

Yiran Wei¹, Guy Lansley¹ & Tim Rains²

¹University College London, ²J Sainsbury's plc

Project Background

Ethnicity has been a major subject in the realm of social research. Issues related to the definition, measurement and classification of ethnicity have attracted on-going debate in the literature. The multidimensional, subjective and complex nature of ethnicity has created barriers for researchers to measure and interpret the characteristics and distribution of various ethnic groups in the United Kingdom where the population has become more culturally diverse as a result of high levels of immigrants from all over the world. The purpose of this project is to identify the major spatial traits in ethnic identity in England and Wales using a range of cultural and heritage indicators that are available as open data sources. It will research ways in which these datasets can give a better understanding of overall cultural diversity at small area geography levels.

Data and Methods

A total of 134 variables were selected from 2011 Census for England and Wales at Output Area level, covering 7 dimensions of cultural identity including country of birth, ethnic group, religion, main language, proficiency in England, age of arrival in UK, and length of stay in UK. The core method applied to create the ethnicity classification is known as cluster analysis; specifically, the k-means clustering approach was adopted for this study. This technique constructs groups of the most similar areas based on the overall similarities and dissimilarities between the variables. Prior to the clustering analysis, Pearson correlation analysis and principal component analysis were applied to reduce the list of 134 variables as much as possible without losing the essential information they contain. Afterwards, a total of 60 variables were identified as the best suited for the cluster analysis.

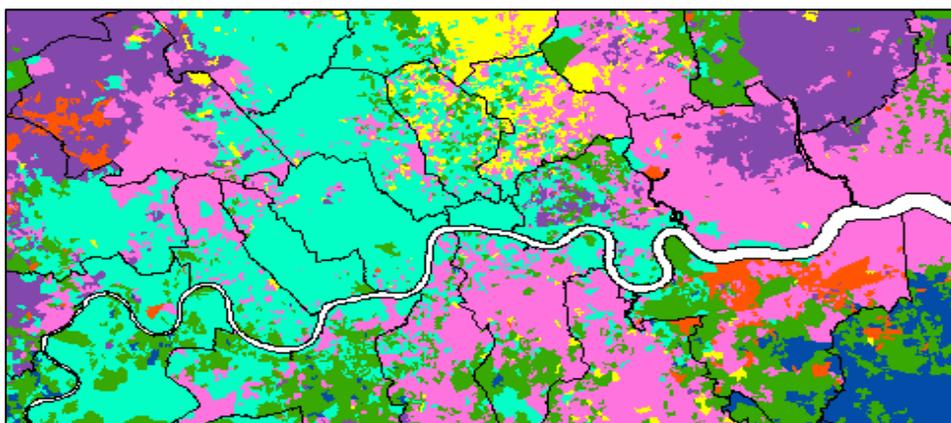
Key Findings

Every Output Areas in England and Wales were grouped into 7 clusters. Each of the groups is identified with

distinctive characteristics of ethnicity components and spatial distribution pattern, including Group A South Asian, Group B Turkish, Cypriot and Greek, Group C East Asian and Non-British White, Group D Nepalese, Group E Mixed Ethnic Group, Group F Black African and Caribbean, and Group G White British. In addition, a London-specific Output Area classification has also been created alongside the England and Wales classification using an identical methodology. Comparison between the two reveals some interesting patterns. Both outputs are similar in general, however the London-specific version further separates the South Asian group into two subgroups: one reflecting more Muslim African ethnicity whilst the other is far more Bangladeshi or South-East European. Furthermore, Nectar data on sales of ethnic food is provided by Sainsbury and used to benchmark the clustering results. Cross-referencing the cluster analysis with retail sale data has suggested the validity of the classification through positive association. For example, the consumption of halal meat in African and Caribbean neighbourhoods is found to be approximately 30 times higher than the White British group. Finally, the examination of distance from the cluster centre for measuring the uncertainty indicates that variables chosen for classification appear to have been reasonably sound, but it also highlights the issue of subjectivity regarding the variable selection process.

Value of the Research

Retailers that successfully understand and target the shifting cultural makeup of society will outperform those using antiquated research and customer stereotypes. This research provides valuable insight for Sainsbury's ethnic food sale planning. The classification results retained from this study can provide a sound reference for Sainsbury's ethnic food allocating planning. By knowing the spatial trait of different ethnic group, Sainsbury can then adjust and further improve their ethnic food product range to meet the need of their ethnic minority customers.



- A: South Asian
- B: Turkish, Cypriot & Greek
- C: East Asian & Non-British White
- D: Nepalese
- E: Mixed Ethnic
- F: African & Caribbean
- G: White British

2011 Cultural, Ethnic and Linguistic Output Area Classification for England and Wales – Central London