

## Social Energy Responsibility: Identifying Vulnerable Energy Customers Through a K-Means Clustering Approach

Ffion Carney<sup>1</sup>, Alex Singleton<sup>1</sup> and Ben McKeown<sup>2</sup>  
<sup>1</sup>University of Liverpool, <sup>2</sup>E.ON UK

### Project Background

In order to reduce domestic energy consumption and increase domestic energy efficiency, the government implemented the Energy Company Obligation (ECO); a scheme to obligate large energy suppliers to deliver energy efficiency measures to domestic premises. One of the obligations of this scheme is focused on improving the ability of low income and vulnerable households to heat their homes. In order to achieve this obligation, it is vital that energy companies are able to determine what constitutes as a 'vulnerable household' and are also able to identify which areas should be targeted. This study therefore aimed to identify areas that contain a high proportion of vulnerable households and should be targeted as part of the ECO, by taking into account demographic and property characteristics alongside average annual energy consumption data.

### Data and Methods

The main dataset used in this study was E.ON's in-house customer data, which consisted of actual and modelled annual electricity and gas consumption data alongside several demographic and property characteristics for over 3.6 million E.ON customers. The dataset was aggregated to LSOA level to allow for additional data from the 2011 Census to be included, primarily consisting of data on housing tenure, residence type and average property size that were not included in the main E.ON dataset. These variables were selected as they had all been shown to have a strong relationship with energy consumption, efficiency and vulnerability. The final dataset was then used to undertake a k-means clustering analysis, with the aim of identifying the cluster that contained LSOAs with the highest proportion of vulnerable households.

### Key Findings

The k-means clustering algorithm was found to be an effective method of segmenting the dataset into seven distinct clusters. Comparing the defining characteristics of each of these clusters, alongside their electricity and gas consumption, allowed for the cluster containing LSOAs with the highest proportion of vulnerable households to be identified. Cluster 6, named 'Fuel Poor Private Renters', was identified as the

most vulnerable cluster, primarily due to the high proportion of low income households along with the higher than expected average energy consumption when compared to income. This cluster also contained a significant proportion of solid walled properties and households suffering from fuel poverty.

In addition to identifying the cluster containing LSOAs with a high proportion of vulnerable households, this study also ranked the remaining clusters in terms of vulnerability based on their defining characteristics and average energy consumption, as seen in Figure 1.

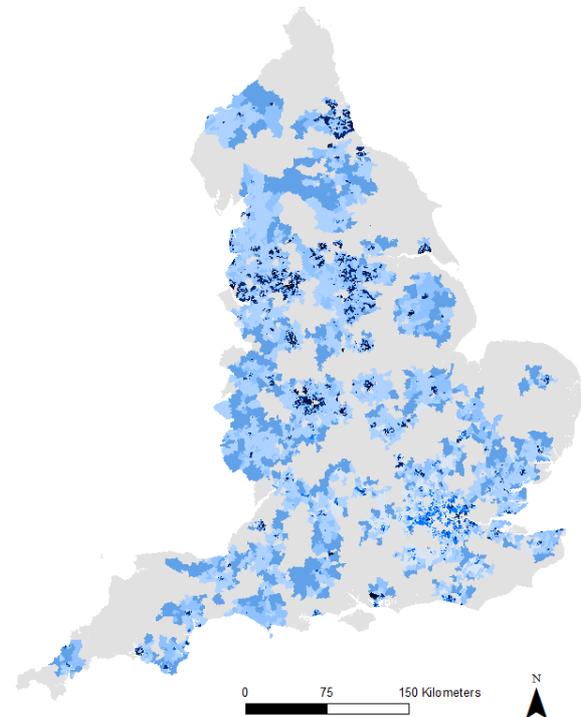


Figure 1. Cluster vulnerability ranking

### Value of the Research

This study provides a method for identifying areas that contain a high proportion of vulnerable E.ON energy customers and should therefore be targeted as part of the ECO. The vulnerability ranking produced in this study could also assist in any future targeting of energy efficiency measures, allowing for areas with a higher proportion of vulnerable households to be prioritised. This understanding of vulnerability is vital for the effective targeting of vulnerable households in order to ensure the successful implementation of energy efficiency measures and carbon policies in the future.