

International Journal of Population Data Science

Journal Website: www.ijpds.org



Developing a Geo-intelligent Housing Submarket Classification Framework for England and Wales

Ayisha Dubi¹, Evgeniya Lukinova¹, Nicholas Clare¹, and James Goulding¹

¹University of Nottingham, UK

Introduction & Background

Amid escalating housing affordability pressures, housing shortages, and widening spatial inequalities, with 1.33 million households on social housing waiting lists in England in 2024 (UK Government, 2025), there is growing demand for evidence-based housing policy (Rodríguez-Pose and Storper, 2020; Florida, Rodríguez-Pose and Storper, 2023). Yet housing markets are still commonly analysed using fixed administrative boundaries that often fail to reflect lived housing experiences.

Advances in urban analytics highlight the need for data-driven approaches to capture dynamic socio-spatial change (Batty, 2021). Housing markets are known to be internally segmented by dwelling, neighbourhood, and accessibility characteristics (MacLennan and Tu, 1996; Leishman, 2009).

Objectives & Approach

This research develops a framework to explore housing submarkets across England and Wales using digital footprints, spatial analytics, and housing data. It integrates HM Land Registry Price Paid Data (2018–2025), Energy Performance Certificate records (2008–2025), Indices of Multiple Deprivation, ONS geography, Census indicators, and OpenStreetMap amenities.

A pilot study in Doncaster and Rotherham linked property transactions to EPC records through postcode standardisation, address matching, and fuzzy record linkage, followed by aggregation to Lower Layer Super Output Areas (LSOAs). Exploratory spatial analysis was followed by correlation analysis and Principal Component Analysis to reduce dimensionality prior to K-means clustering for the identification of candidate housing submarkets. Subsequent phases will explore machine learning approaches to refine and predict submarket membership.

Relevance to Digital Footprints

This research uses housing transactions as behavioural digital footprints generated through property market activity. Future phases will incorporate social housing demand, homelessness

applications, and temporary accommodation placements. Future iterations will also integrate online property listing data to capture housing search behaviour.

Results

Exploratory spatial analysis and bivariate mapping revealed persistent socio-spatial variation in median house prices and neighbourhood deprivation. Correlation analysis identified meaningful relationships between housing value, floor area, energy efficiency, and neighbourhood context. Principal Component Analysis retained approximately 85% of total variance within four principal components, providing a reduced feature space for clustering. Preliminary clustering suggests the presence of four distinct candidate housing submarkets characterised by differences in dwelling type, energy efficiency, development age, and deprivation. Spatial mapping revealed that several candidate submarkets extend across Doncaster and Rotherham administrative boundaries, suggesting that functional housing markets evolve over time and are not fully captured by fixed geographic units.

Conclusions & Implications

This research demonstrates the potential of integrating digital footprints, open spatial data, and spatial analytics to identify and characterise functional housing submarkets. The proposed framework provides a foundation for subsequent predictive modelling of submarket dynamics and offers a transferable approach for supporting evidence-based housing policy, regeneration planning, homelessness prevention, and future analysis of housing inequality across England and Wales.

Keywords

housing submarkets, digital footprints, spatial analytics, housing inequality