

Quantifying Street Attractiveness and Retail Performance: Statistical Associations Between Street-View Imagery and Retail Data in the United Kingdom

Can perceived street qualities around a retail centre be associated with its retail profile?

Itzel Alejandra Soto Crisanto | Supervised by Dr Seppo Virtanen (University of Leeds); Jan Magnuszewski, and Letitia Bailey (CACI Ltd)
MATH5872M: Dissertation in Data Science and Analytics, University of Leeds, School of Mathematics

Background and Motivation

It is widely assumed that how a street looks shapes where people shop. Yet whether the perceived qualities captured in street-level photographs are statistically associated with the commercial and demographic profile of surrounding retail centres has received little direct quantitative attention in the UK.

This project addresses that gap without assuming in advance which visual quality or retail measure should matter most. It combines 109,962 UK street-level images from the Global Streetscapes dataset — each carrying six model-derived perception scores — with expenditure and catchment data for 3,532 retail centres supplied by CACI Ltd. The two datasets were never designed to be joined. The question is whether, statistically, they share a common structure.

Data and Methods

Street imagery: 109,962 UK street-level images from the Global Streetscapes dataset (Hou et al., 2024), each carrying six model-derived perception scores (0–10): Beautiful, Boring, Depressing, Lively, Safe, and Wealthy.

Retail data: CACI Ltd Retail Footprint — 20 measures of retail and leisure spend, demographic mix, and visitor type across 3,532 UK retail centres. Both datasets are modelled representations, not direct observations.

Spatial linkage: a buffer circle (500–1600 m) was drawn around each image and matched to any retail centre whose boundary it touches, tested at six radii.

Method: Canonical Correlation Analysis (CCA) — examines all six perception scores and all twenty retail measures simultaneously, without pre-committing to any single perception–retail pair.

Validation: permutation testing ($p < 0.002$), cross-validated log-likelihood, consistency across all six buffer radii, sample representativeness check, and blind visual inspection of what the scores respond to in the images.

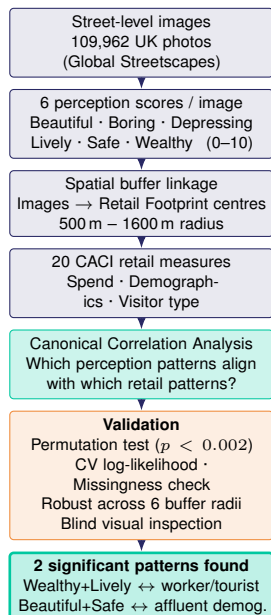
Key Findings

- A statistically significant but modest relationship exists:** two canonical pairs link street perception scores to retail performance, stable across all six buffer radii ($p < 0.002$). The first pair accounts for 10.2% of shared variance, the second for 3.7% — the relationship is real but small in magnitude.
- Wealthy and Lively drive the first pattern ($\rho = 0.319$, shared variance 10.2%):** retail centres whose surrounding streets score high on Wealthy (0.91) and Lively (0.72) tend to be the same centres where worker and tourist spending are large relative to residential spending, and the Retail Footprints of those centres tend to draw more affluent visitor groups, such as Luxury Lifestyles. Beautiful, Boring, and Depressing play minor roles. The component distinguishes centres by the composition of their retail profile: more visitor and worker-oriented versus more residential and middle-income.
- Beautiful and Safe drive the second ($\rho = 0.193$, shared variance 3.7%):** retail centres with more Beautiful (0.82) and Safe (0.49) streets tend to have lower values across all twenty retail measures — a “more versus less of everything” dimension rather than a distinction between types of activity. However, blind visual inspection found that ten of the fifteen highest-scoring images were parks and open green spaces, not commercial streets. Beautiful here reflects open, green settings such as parks rather than attractive commercial streets.
- Two components are statistically supported:** cross-validation confirms a third adds no out-of-sample predictive value, and remaining pairs were either indistinguishable from chance or too unstable across radii to interpret reliably.
- Coverage limits the conclusion:** only 78–143 of 3,532 UK retail centres had sufficient image coverage, concentrated in City Centres and Major Town Centres. Findings do not generalise to smaller, rural, or less-photographed centres.

Value of the Research

This project provides a solid first step towards understanding whether image-based perceptions of streets and the commercial profiles of nearby retail centres share a common structure. The findings show that they do — modestly, systematically, and in a way that is statistically supported. Before perception scores are used in retail or planning analysis, their visual meaning must be examined rather than assumed from their label. The limitations documented here define exactly what a follow-up study would need to address.

Study Pipeline



CCA Pair 1: Perception vs Retail

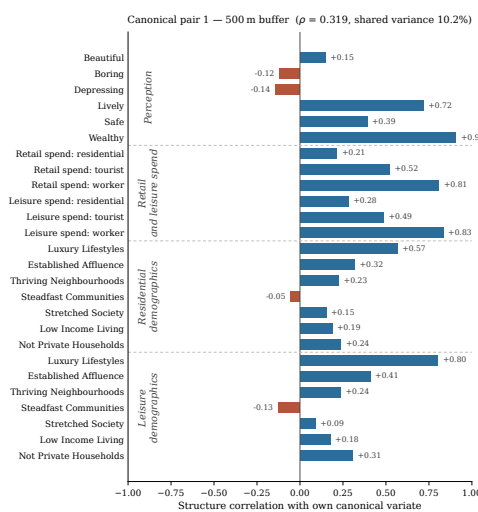


Fig. 1: Structure correlations, first canonical pair (500 m, 78 RFs). Wealthy and Lively link most strongly to worker and tourist spending.

Spatial Coverage: Example Cities

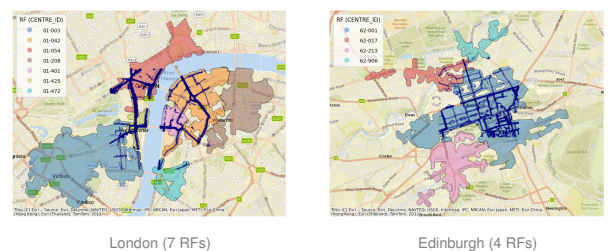


Fig. 2: Coloured zones = Retail Footprint catchments; dark lines = image locations within each 500 m buffer.

Example Street-Level Images (Global Streetscapes)



Selected References

- Hou Y, et al. (2024). Global Streetscapes. *ISPRS J. Photogramm.* 215: 216–238.
- Bach F, Jordan M (2005). A probabilistic interpretation of CCA. *Tech. Rep. 688*, UC Berkeley.
- Dolega L, Pavlis M, Singleton A (2016). Estimating retail centre attractiveness. *Appl. Geogr.* 74: 82–97.
- Winkler S (2024). Stepwise CCA tests. *Statist. Comput.*

Acknowledgements

With thanks to Dr Seppo Virtanen (University of Leeds) for academic supervision, and to CACI Ltd for the Retail Footprint data and industry guidance. Global Streetscapes dataset: Hou et al. (2024), CC BY-SA 4.0.