Assessing the Value of Footfall Data in Retail Analytics

Terje Trasberg*, James Cheshire† and Paul Longley‡

[1]Department of Geography, University College London

January 31, 2019

Summary

Despite the benefits research has attributed to the use of data-driven analytics, retailers have been slow to integrate data into their decision making processes. Using footfall data as an example, this paper first discusses the challenges retailers are facing when attempting to apply big data for business value and then presents a practical use case for data-driven retail site selection.

KEYWORDS: Retail Geography, Big Data, Footfall, data-driven analytics

1. Introduction

The rapid growth of the e-commerce sector alongside the rising costs of operating a retail unit have motivated many retailers to evaluate the profitability of their bricks-and-mortar stores. Novel data sources such as loyalty cards and social media platforms coupled with innovative monitoring technologies provide a rich source of information for studying customer behaviour in retail areas.

However, due to the novelty of these data sources and a lack of understanding of how to utilise them, retailers have been slow to integrate big data into their decision making framework (Ridge et al., 2015).

Using footfall data as an example, this project works in collaboration with retailers to firstly identify the challenges in using data to solve real-world problems and secondly provides a practical use case for applying footfall data in retail location analytics.

2. Research Background

Recent technological developments are greatly extending our ability to monitor, and hence understand, aspects of human activity patterns. New Big Data sources are benefiting a number of applications in urban studies, town management and retail.

Our case study utilises data from the SmartStreetSensor Project, a collaboration between the Local Data Company and the ESRC Consumer Data Research Centre. This study is monitoring footfall at over 700 sites in retail areas across the UK in order to study the dynamics of pedestrian flows (Murcio et al., 2018). This dataset has already been applied to create a temporal classification of footfall patterns in retail areas (Lugomer, 2018).

Besides learning about retail activity patterns, a number of studies have tried to understand the impact of pedestrian flows on retail turnover. Earlier studies (e.g. Newby et al., 1992; Hass-Klau, 1993) focused mainly on the impact of street pedestrianisation and reported a positive correlation between general increase in footfall and turnover at retail units. A more recent data-driven study by Perdikaki et al. (2012) combined hourly footfall data with sales data and found that the number of transactions indeed increases with traffic, but that the relationship attenuated at high footfall volumes.

* terje.trasberg.15@ucl.ac.uk
† james.cheshire@ucl.ac.uk
‡ p.longley@ucl.ac.uk

3. Retailers' Challenges in using Footfall Data

In order to gain an understanding of industry views and practices in using footfall data in retail analytics, the lead author interviewed six UK-based retailers who are involved in the SmartStreetSensor project.

Despite having access to footfall data at their existing locations, none of the interviewed retailers have made use of them. Instead, the performance of retail units is essentially evaluated based exclusively upon turnover.

Drawing on the interviews, retailers struggle to understand the importance of footfall compared to the other location characteristics such as competition density and demographic profile of the population. For example, a restaurant retailer claimed that their business is more dependent on the socio-economic profile of the customers and less on the general density of the pedestrian flows.

However, despite the challenges, retailers recognised that footfall provides a unique measure of the dynamics of activity patterns in the environs of their retail sites. For instance, a Fast Food Retailer plans to implement footfall data in performance analysis, in order to calculate the ratio between opportunities (=footfall) and turnover (=sales) across its portfolio of restaurants. The Family Restaurant sees the value in footfall data in indicating irregularities such as locations which receive high footfall but are not able to convert this into sales. This could flag issues such as poor management of retail unit, inappropriate location or mismatch of restaurant offer with those passing the retail units.

4. Case Study Analysis

Another common concern of the retailers is the lack of practical use cases which implement footfall data. Therefore, working in collaboration with the Fast Food Retailer, this paper provides a practical use case for applying footfall data in retail location analysis.

The aim of the case study is to help a Fast Food Retailer analyse the expansion opportunities in the UK retail centres and determine a set of criteria for selecting potential locations. Making use of the data available for this study, we propose the workflow outlined in **Table 1**.

4.1. Data

Fast Food Retailer has provided data on hourly sales volumes that is, the number of transactions, for 10 of their locations, 8 of which are in London, 1 in Birmingham and 1 in Brighton. The data were made available for the year 2017.

The footfall data are made available through the Consumer Data Research Centre (CDRC). Technical details and description of the original data capture methodology have been described in earlier publications on the SmartStreetSensor Project (Murcio et al., 2018).

WORKFLOW OF THE DATA-DRIVEN SITE SELECTION
1. Cluster existing retail units based on performance (=hourly sales)
2. Determine the common (locational) characteristics of each cluster by looking at footfall patterns, workplace zone, surrounding retail mix, competition density, etc.
3. Based on common characteristics of the best performing retail units, define a set of criteria for site selection
4. Select retail centres in the SmartStreetSensor network that fulfil the criteria set in previous step
5. Exclude retail centres that have an existing Fast Food Retailer unit

Table 1 Workflow

4.2. Method

In the first instance, we focussed only on an "average weekday performance" and aggregated the data to average hourly sales and the corresponding footfall profile for 24 hours taking into account data from Monday to Friday.

To cluster similarly performing retail units, we measured the Euclidean distance between hourly sales and fed the measures into a hierarchical clustering algorithm, which generated a hierarchy of similar groups. The intergroup similarity was defined by applying Ward's method, which defines the clusters by minimizing the total within-cluster variance (Boehmke, 2018).

The clusters were then described by looking at the footfall patterns and other spatial characteristics such as the workplace zone class and competition density around the retail units in the clusters. The aim was to determine common characteristics of well-performing retail units, which were then translated into a set of site selection criteria.

The criteria are then applied to highlight the expansion opportunities for the Fast Food Retailer. The input includes all retail centres defined by Pavlis et al. (2018), which are included in the SmartStreetSensor network and have footfall data available for the year 2017. Finally, retail centres, which contain existing Fast Food Retailer unit(s) were excluded from the results.

4.3. Results

The retail units were clustered into 3 groups. Cluster 1 and 2 both include 4 retail units, which are all located in London. Cluster 3 retail units are both outside of London. Cluster 3 retail units receive in average more footfall than Cluster 2 locations; however, London locations (Cluster 2) are able to turn higher percentage of the footfall to sales.

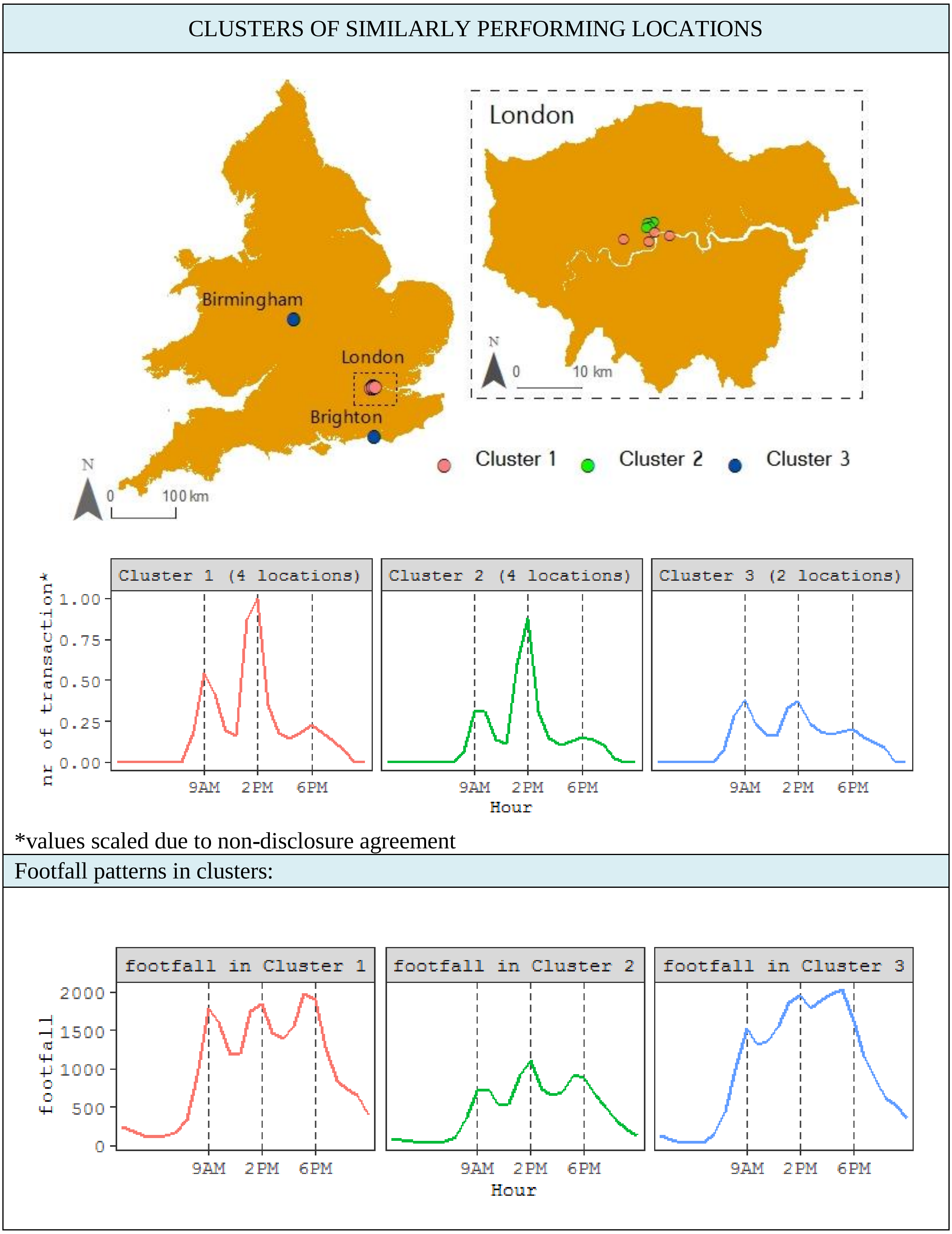

Figure 1 Similarly performing retail units

The lower conversion rate (ratio between footfall and sales) of Cluster 3 retail units can be attributed to different socio-economic profiles of the locations. For instance, according to the 2011 classification of workplace zones, all London stores fall into the category of "City and Business Parks" classification whereas Birmingham and Brighton are in "Retail" zone (ONS, 2011). Therefore, we define workplace zone "City and Business Parks" as one of the criterion for the site selection.

To find further common characteristics of well-performing retail units, we look at the surrounding retail composition using LDC’s retail locations database (LDC, 2017). Cluster 3 retail units are surrounded by more diverse retail mix. For example, the Birmingham store has 186 retail locations within a 200-meter radius (walking distance). The stores in London (Cluster 1 and 2) are surrounded by fewer retail units, but the concentration of competing locations is higher. For instance, Cluster 1 stores have in average 130 retail locations within a 200-meter radius and around 10% of those are take-away retailers. So, competition seems to boosts the sales at the Fast Food Retailer stores. Therefore, we add the average competition intensity (> 12 take-away food retailers in 200m) to the site selection criteria.

Third site selection criterion looks at the footfall at the peak hours at the well-performing retail units. Although, Cluster 1 and 2 have a similar footfall profile with three peaks, cluster 1 shows better sales numbers. Therefore, we apply a threshold footfall count of 1500 for morning and lunch peak to the site selection criteria.

The criteria are outlined in **Figure 2** together with the visualisation of high potential locations selected based on these criteria.

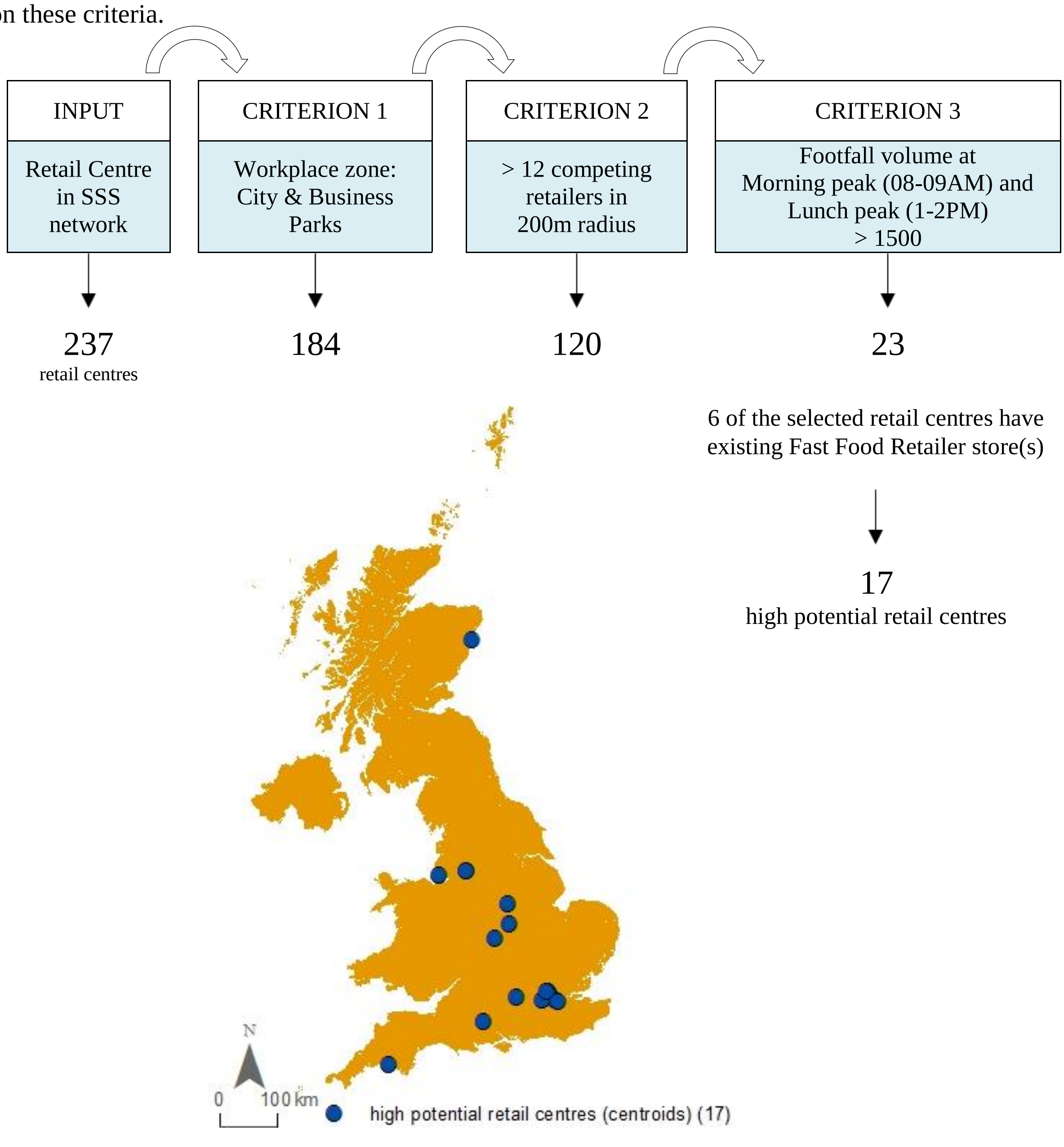

Figure 2 Results

Of the 237 retail centres only 23 fulfill the site selection criteria. There is at least one existing Fast Food Retailer store in 6 of the selected retail centres. Those 6 retail centres include in total 27 Fast Food Retailer stores (for instance, 11 in central London), which makes up over half of the existing Fast Food Retailer units. Therefore, the selected criteria seem to correspond well to the criteria used to select the existing locations.

In total the analysis identified 17 new high potential retail centres, which have similar characteristics to the existing well performing retail units.

The analysis could be further developed by adding additional criteria such as requirements for available retail space and rental costs.

5. Conclusion

The study aimed to define the challenges and potentialities of implementing footfall data in retail location analysis. The results of the case study showed that footfall data can be successfully use to understand the performance of existing retail units and to narrow down the number of high potential sites. However, the analysis also showed that footfall data should be analysed in the context of their spatial environment (e.g. workplace zone, socio demographic profile). Making decisions based on only footfall data can be misleading as density of pedestrian flows reveal movement patterns, but do not help to understand why some retail areas are inherently more popular and profitable than the others.

6. Acknowledgements

This work is funded by the UK ESRC Consumer Data Research Centre (CDRC). We would like to thank the Local Data Company for providing footfall data for this project and the CDRC's industry partners who have contributed to this study.

References

Boehmke B (2018) 'Hierarchical Cluster Analysis, University of Cincinnati Business Analytics R Programming Guide'. Available at: https://uc-r.github.io/hc_clustering (Accessed: 12 November 2018).

Hass-Klau C (1993) 'Impact of pedestrianization and traffic calming on retailing', *Transport Policy (1), pp. 21-31*

Lugomer K (2018) 'Towards a Comprehensive Temporal Classification of Footfall Patterns in the Cities of Great Britain'. *GIScience 2018 (43). pp. 1-6*

Murcio, R., Soundararaj, B. and Lugomer, K. (2018) 'Movements in Cities: Footfall and its Spatio-Temporal Distrib', in Longley, P., Cheshire, J., and Singleton, A. (eds) Consumer Data Research. UCL Press, pp. 85–95

Newby L, Spencer-Wort S and Wiggins P (1992) '*Paved with Gold?: A Study of the Economic Impact of Pedestrianisation and Its Relevance to Leicester'.*

Office for National Statistics (2011) 'A Classification of Workplace Zones from the 2011 Censuses for the United Kingdom'. Available at: http://cowz.geodata.soton.ac.uk/ (Accessed: 12 November 2018)

Pavlis M, Dolega L and Singleton A (2018) 'A Modified DBSCAN Clustering Method to Estimate Retail Center Extent'. Geographical Analysis (50), pp. 141-161. doi:10.1111/gean.12138

Perdikaki O, Kesavan S and Swaminathan J M (2012) 'Effect of Traffic on Sales and Conversion Rates of Retail Stores', *Manufacturing & Service Operations Management*, 14(1), pp. 145–162. doi: 10.1287/msom.1110.0356.

Ridge, M., Johnston, K. A. and O'donovan, B. (2015) 'The use of big data analytics in the retail industries in South Africa', *African Journal of Business Management Full*, 9(19), pp. 688–703. doi: 10.5897/AJBM2015.7827.

The Local Data Company (2017) *LDC's Field Researched Points of Interest.*

Biography

Terje Trasberg is a Geography PhD student at University College London (UCL). Her research is primarily focused on retail geography and investigating the potential variables driving retail locations profitability and effectiveness. She previously studied Geography BSc at University of Salzburg, Austria and Geospatial Analysis MSc at UCL.

James Cheshire is Senior Lecturer in Quantitative Geography at University College London. He is interested in geovisualisation of mobile populations, and has written best-selling books on this theme.

Paul Longley is Professor of Geographic Information Science at University College London, where he also directs the Consumer Data Research Centre. His research interest are grouped around socioeconomic applications of GISc, with particular emphasis upon the geo-temporal demographics of city populations.