The Ups and Downs of London High Streets Throughout COVID-19 Pandemic: Insights from Footfall-Based Clustering Analysis

Xinglei Wang[1]
SpaceTimeLab for Big Data Analytics, University College London, UK

Xianghui Zhang
SpaceTimeLab for Big Data Analytics, University College London, UK

Tao Cheng
SpaceTimeLab for Big Data Analytics, University College London, UK

Abstract

As an important part of the economic and social fabric of urban areas, high streets were hit hard during the COVID-19 pandemic, resulting in massive closures of shops and plunge of footfall. To better understand how high streets respond to and recover from the pandemic, this paper examines the performance of London's high streets, focusing on footfall-based clustering analysis. Applying time series clustering to longitudinal footfall data derived from a mobile phone GPS dataset spanning over two years, we identify distinct groups of high streets with similar footfall change patterns. By analysing the resulting clusters' footfall dynamics, composition and geographic distribution, we uncover the diverse responses of different high streets to the pandemic disruption. Furthermore, we explore the factors driving specific footfall change patterns by examining the number of local and nonlocal visitors. This research addresses gaps in the existing literature by presenting a holistic view of high street responses throughout the pandemic and providing in-depth analysis of footfall change patterns and underlying causes. The implications and insights can inform strategies for the revitalisation and redevelopment of high streets in the post-pandemic era.

2012 ACM Subject Classification Applied computing → Sociology

Keywords and phrases High street, performance, footfall, clustering analysis, COVID-19

Digital Object Identifier 10.4230/LIPIcs.GIScience.2023.80

Category Short Paper

Funding *Xinglei Wang*: the author's PhD research is jointly funded by China Scholarship Council and the Dean's Prize from University College London.

1 Introduction

Over the past 3 years, the COVID-19 pandemic, a global public health crisis of unprecedented scale, has brought about substantial changes to urban environment [7]. Although the COVID-19 is no longer a public health emergency of international concern [9] by the time this paper was written, the long-term effects linger and continue to shape urban landscapes.

Among the most affected urban areas are spaces of consumption such as high streets which are often the heart of local communities, serving as centres for commerce, social interaction, and cultural activities. Important as they are, high streets across the UK suffered a devastating blow during the pandemic, with over 17500 chain stores and other venues closing in Great Britain [2] and footfall decreasing by over 80% [5]. Given the pivotal role of

[1] Corresponding author

12th International Conference on Geographic Information Science (GIScience 2023).
Editors: Roger Beecham, Jed A. Long, Dianna Smith, Qunshan Zhao, and Sarah Wise; Article No. 80; pp. 80:1–80:6
LIPICS Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

high streets in the economic and social fabric of urban systems, it is of great importance to explore the impacts of pandemic on high streets' performance to comprehend the broader consequences on local economies, businesses, and communities.

Existing studies have utilised various forms of data to assess the economic performance of consumption spaces, such as vacancies [3] and footfall [8]. Within the COVID-19 context, Enoch et al. used footfall data to analyse the impact of COVID-19 pandemic on six high streets in England [4]. Ballantyne et al. used a mobile phone app location dataset to examine the recent recovery of retail centres from the pandemic [1]. Although they provided empirical evidence of the impact of the pandemic, there are some limitations remained to be addressed. Firstly, the COVID-19 pandemic and its aftereffects last several years, with several rounds of national lockdowns enacted, but existing studies do not cover the whole period, thus cannot present the full picture of the responses of high streets. Secondly, existing literature only focus on the change of footfall counts, lacking in-depth analysis of the change patterns and its underlying causes. To fill the research gaps, we examine and evaluate the performance of London's high streets during the pandemic using longitudinal footfall data. Specifically, footfall data spanning over two years was calculated from a mobile phone GPS dataset and time series clustering was applied to generate multiple groups of high streets with similar patterns. By analysing the distinctive footfall change patterns of resulting clusters and their geographic distribution, we unravelled the varying responses of different high streets to the disruption and the spatial patterns of different clusters of high streets. Furthermore, we linked the clustering results to the existing typology of retail centres and gained further insight into how the different composition of clusters corresponds to their performance. Lastly, we delved into the cause of particular change patterns by looking at the number of local and nonlocal visitors.

In the following sections, we describe the dataset used in this study, followed by a brief introduction of the methods employed. We present and analyse the results, discuss their implications, and provide recommendations for policymakers and urban practitioners. We conclude the paper by summarising the main findings and pointing out directions for future work.

2 Data

- **Mobile phone GPS trajectory data**: it is a large-scale mobility dataset which contains millions of anonymous users' mobile phone GPS trajectory data (collected from tens of location-based service apps) provided by Location Sciences under GDPR compliance. The dataset spans 3 years, and we define our study period from the first Monday of February 2020 (03/02/2020) to the last Sunday of April 2022 (24/04/2022), spanning 812 days (116 weeks). The number of unique devices in London in February 2020 exceeds 610,000. The data collection method and sampling rate over the whole country remains consistent throughout the study period.
- **High street boundary dataset**: provided by the Greater London Authority[2], this is a shapefile containing the boundaries of 616 London high streets located outside the Central Activity Zone.
- **Lower Layer Super Output Areas (LSOAs)**: It is a geographic hierarchy designed to improve the reporting of small area statistics in England and Wales. This study utilised the LOSAs dataset created in the most recent 2021 census and only those within Greater London area were included.

[2] `https://data.london.gov.uk/dataset/gla-high-street-boundaries`

3 Methods

In this section, we present the workflow of footfall calculation and give a brief introduction to the K-means clustering method.

Footfall calculation

The workflow consists of the following steps:

1. Home detection: obtain the LSOA-level home location of each individual, which is denoted as *home_lsoa*. Here, the home of a person is defined as the LSOA where they generate the greatest number of GPS points during night time (e.g., 22:00-07:00).
2. Stop detection: to get the stop which is where people remain stationary for more than a specific amount of time (we set 5 minutes as the threshold in this study).
3. Identity inference: infer the identities (being one of resident and non-resident) of the people visiting a certain high street. If the *home_lsoa* of a person is one of the LSOAs that intersects with the high street, then this person is considered as a local resident, otherwise, a non-resident.
4. Footfall calculation: join people's stops with high street boundaries and calculate the footfall w.r.t resident and non-resident by day and sum the 2 types to get the overall footfall.

The output is daily footfall counts on each of the high street in London. We further aggregate the daily footfall into weekly ones by summing over 7 days of each week. This step can not only smooth the time series but also reduce the length of it, which can improve the results of clustering. We also normalise the footfall counts to convert them into relative values between 0 and 1, which is an essential pre-processing step for clustering.

K-means time series clustering

We employ a K-means time-series clustering algorithm to cluster the time-series of weekly overall footfall on the high streets. We use the Euclidean distance as the metric for clustering. One advantage of this method is that the distance between any two objects is not affected by the addition of new objects to the analysis, which may be outliers [6]. Elbow method and the Silhouette score are used to identify the optimal cluster number K.

4 Results

High street clusters and their spatial distribution and composition

Three high street clusters are identified based on the time-series pattern in footfall. The share of high streets in each cluster are 39 %, 33 % and 28 %, respectively. Figure 1 shows the footfall time series for the entire study period across the three clusters. Cluster 1 had the smallest drop when the first national lockdown came into effect. But after the third national lockdown, it remained the lowest while the other two (especially Cluster 3) recovered better. Cluster 2 and 3 exhibited similar trend, but Cluster 3 surpassed Cluster 2 (and Cluster 1) significantly during the "stepping out of lockdown" period, reaching its pre-pandemic level. The results demonstrate the varied ability of high streets to weather crises.

Figure 2 shows the spatial distribution of the clusters where some degree of spatial clustering is notable. Most of the high streets in Cluster 1 are located in inner London area, while Cluster 3 finds more high streets in outer London. Combining the spatial distribution

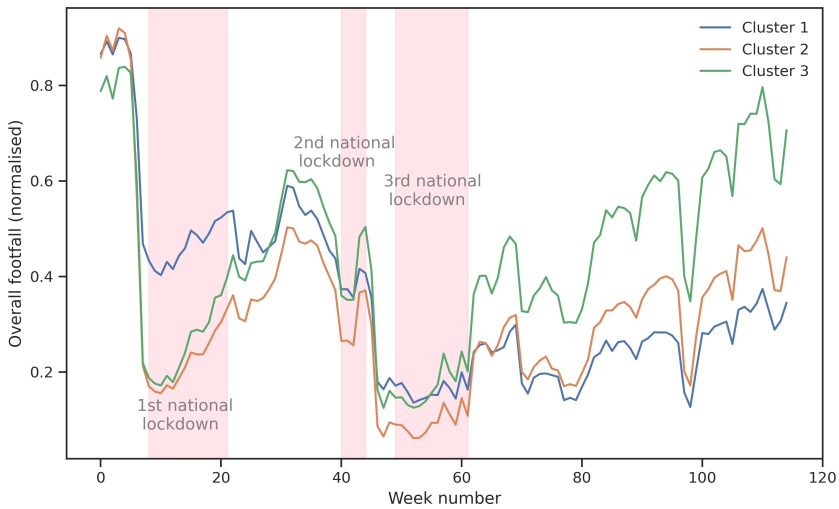

Figure 1 Time-series pattern of the three identified high street clusters (the pink shades indicate the period of three national lockdowns).

and performance, We can draw the crude conclusion that high streets located at the periphery of the city tend to recover better than those closer to the city centre. High streets in Cluster 2 are more evenly distributed, but are relatively bigger in sizes.

To gain more information about what types of high streets each cluster contains, we further look into the composition of each cluster by linking the high streets with Retail Centre Typology provided by CDRC [3]. Figure 3 shows the composition of each cluster, where we can see that Cluster 1 has the highest proportion of small local centres compared to Cluster 2 and 3. In Cluster 2, only 38.7% of the high streets are small local centres, the lowest among the three clusters, while higher percentages of district centre and town centre are found in this cluster. As for Cluster 3, the composition is very similar to Cluster 1, but the proportion of high streets located in the outer London area is much higher than that of Cluster 1 (referring to Figure 2).

Local and nonlocal visitors

It is of great interest to us to uncover the underlying cause that made the three clusters affected so disproportionately by the pandemic. In particular, the compositions of Cluster 1 and 3 are very similar, yet they have such distinctive responses during multiple rounds of lockdowns and after-lockdown recovery period. With the question in mind, we calculated the number of local visitors (residents) and nonlocal visitors (non-residents) in each cluster and present the result in Figure 4.

Clearly, the stronger resilience Cluster 1 showed during the first national lockdown is owing to the preservation of local residents, while its downfall in the recovery phase is largely due to the continued loss of local residents (possibly because of people moving out of city [10]). The rise of Cluster 3 after the third national lockdown is much explained by the rapid increase in the number of both local residents and nonlocal visitors.

[3] https://data.cdrc.ac.uk/dataset/retail-centre-boundaries-and-open-indicators

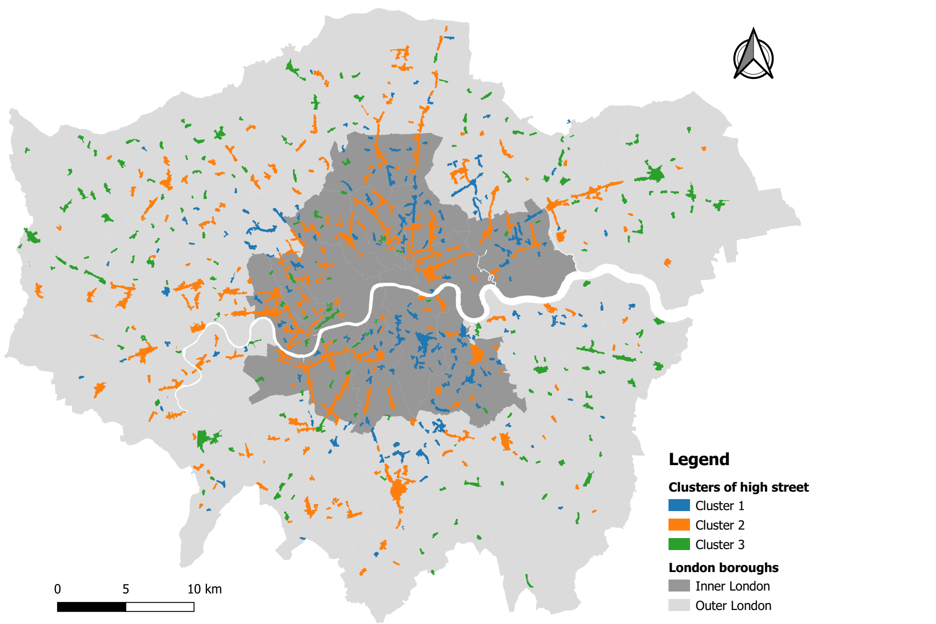

Figure 2 Spatial distribution of high streets in three clusters.

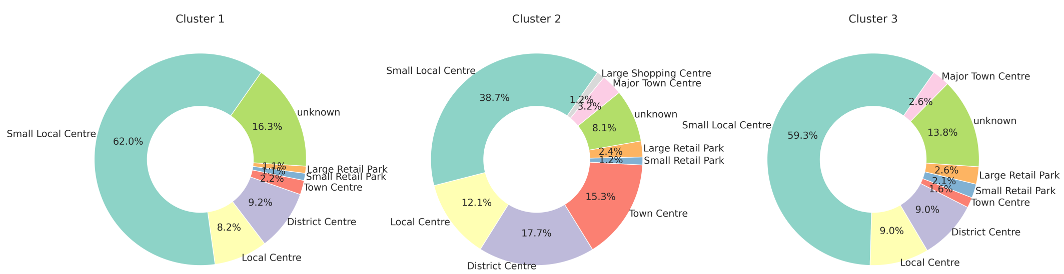

Figure 3 The proportion of different high streets in three clusters.

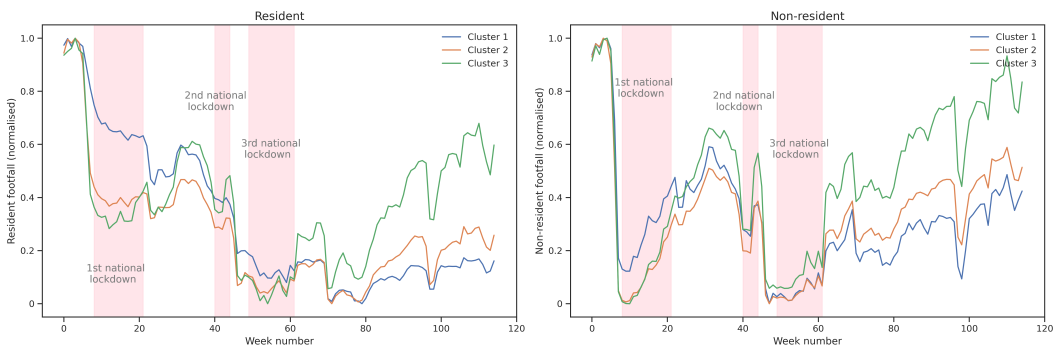

Figure 4 The local (resident) and nonlocal (non-resident) visitors in three clusters.

5 Discussion and conclusions

By analysing the full trajectory of high street footfall, we made a significant discovery that the resilience of high streets towards pandemic is very complicated both in space and time: in general, high streets located at the periphery of the city have recovered better and those which endured better through lockdowns may not recover well in post-pandemic period. We also made the first attempt to uncover the underlying cause of such varied responses of high streets. The interesting finding that local residents is the "key to success" highlights the importance of community engagement for London high streets. Policymakers and local authorities might consider organising more local events and activities, as well as launching initiatives such as community-led regeneration projects, to help strengthen the high streets' attractiveness, bring back residents and create a sense of ownership and pride among them.

In conclusion, this paper is a first step towards the quantification and clustering analysis of high streets performance throughout the COVID-19 pandemic. By identifying the variations in footfall among different high streets, it provides evidence-based insights for decision-making processes related to urban regeneration, infrastructure development, and the formulation of policies that support local businesses. Policymakers can tailor interventions and allocate resources more effectively, ensuring a targeted approach that addresses the unique characteristics and needs of each high street. For future work, we will incorporate more contextual features into our clustering analysis, such as catchment demographics and built environment information to investigate the mechanisms in which high street response to disruptions.

References

1 Patrick Ballantyne, Alex Singleton, and Les Dolega. Using unstable data from mobile phone applications to examine recent trajectories of retail centre recovery. *Urban informatics*, 1(1):21, 2022.

2 BBC. Pandemic impact 'yet to be felt' on high streets. `https://www.bbc.co.uk/news/business-56378667`, 2021. Accessed: 2023-05-22.

3 Les Dolega and Alex Lord. Exploring the geography of retail success and decline: A case study of the liverpool city region. *Cities*, 96:102456, 2020.

4 Marcus Enoch, Fredrik Monsuur, Garyfalia Palaiologou, Mohammed A Quddus, Fiona Ellis-Chadwick, Craig Morton, and Rod Rayner. When covid-19 came to town: Measuring the impact of the coronavirus pandemic on footfall on six high streets in england. *Environment and Planning B: Urban Analytics and City Science*, 49(3):1091–1111, 2022.

5 Retail Gazette. Uk footfall drops to lowest level on record. `https://www.retailgazette.co.uk/blog/2020/04/uk-footfall-drops-to-lowest-level-on-record/`, 2020. Accessed: 2023-05-22.

6 Ali Javed, Byung Suk Lee, and Donna M Rizzo. A benchmark study on time series clustering. *Machine Learning with Applications*, 1:100001, 2020.

7 Ka Yan Lai, Chris Webster, Sarika Kumari, and Chinmoy Sarkar. The nature of cities and the covid-19 pandemic. *Current Opinion in Environmental Sustainability*, 46:27–31, 2020.

8 Christine Mumford, Cathy Parker, Nikolaos Ntounis, and Ed Dargan. Footfall signatures and volumes: Towards a classification of uk centres. *Environment and Planning B: Urban Analytics and City Science*, 48(6):1495–1510, 2021.

9 WHO. Who chief declares end to covid-19 as a global health emergency. `https://news.un.org/en/story/2023/05/1136367`, 2023. Accessed: 2023-05-22.

10 Elias Willberg, Olle Järv, Tuomas Väisänen, and Tuuli Toivonen. Escaping from cities during the covid-19 crisis: Using mobile phone data to trace mobility in finland. *ISPRS international journal of geo-information*, 10(2):103, 2021.