A Framework for Delineating the Scale, Extent and Characteristics of American Retail Centre Agglomerations: A Case Study for the Chicago Metropolitan Statistical Area during COVID-19

Patrick Ballantyne[*1], Alex Singleton[†1], Les Dolega[‡1] and Kevin Credit[^2]

[1]Geographic Data Science Lab, University of Liverpool
[2]National Centre for Geocomputation, Maynooth University

March 15, 2020

Summary

This paper presents a framework through which formal definitions and typologies of retail centres, such as those in the UK, can be extended to the US. Using the Chicago Metropolitan Statistical Area as a case study and data from SafeGraph, we present a retail centre delineation method that combines Hierarchical-DBSCAN with 'H3' and demonstrate the usefulness of a non-hierarchical approach to retail classification. In addition, we show that the dynamicity and comprehensibility of retail centres makes them an effective tool through which to better understand the impacts of COVID-19 on retail sector dynamics, through exploration of changes to retail centre visits during the pandemic. This approach demonstrates significant scope for a comprehensive classification of retail centre agglomerations for the US, providing a tool through which to monitor the evolution of American retail.

KEYWORDS: Retail, Agglomeration, COVID-19, Machine Learning

1. Introduction

Retail centres, which represent the concentration of individual retail units and their associated activities (Berman and Evans, 2013; Pavlis et al., 2018), are important tools for understanding the distribution and evolution of retail. Arguably, such tools have particular importance in the 21st century, as the retail sector faces a vast swath of challenges ranging from long-term issues like E-commerce and changing consumer behaviours, to more abrupt 'shocks' like the COVID-19 pandemic. The latter is a highly contemporary issue for retail, with numerous academic studies identifying significant alterations to consumer spending (Baker et al., 2020; Nicola et al., 2020), and retail mobility (Bonaccorsi et al., 2020).

There has in recent years been an increase in studies seeking to delineate and characterise retail centre space (Han et al., 2019; Lloyd and Cheshire, 2017; Pavlis et al., 2018), and although similar research is not absent in the US (e.g. Brown, 1992), often these studies are limited in how they delineate space or build typologies for retail centre agglomerations. In this paper, we present an empirically grounded framework through which retail centre agglomerations, characteristics and evolution can be measured, using innovative methods and new forms of data. In particular, we present a case study of the application of this framework to the Chicago Metropolitan Statistical Area (MSA), demonstrating a

[*] P.J.Ballantyne@liverpool.ac.uk
[†] Alex.Singleton@liverpool.ac.uk
[‡] L.Dolega@liverpool.ac.uk
[^] kevin.credit@mu.ie

use-case for the centres we delineate, by investigating how COVID-19 has disproportionately affected the various structures and functions of retail in Chicago.

2. Delineating Urban Retail Centres

In recent literature it has been identified as desirable to use the locational attributes of stores when defining the extent of retail centre agglomerations (Han et al., 2019; Lloyd and Cheshire, 2017; Pavlis et al., 2018). Pavlis et al. utilised a modified-DBSCAN method to delineate extents for the UK, but replicability was limited due to complex parameter specifications and computational inefficiencies. In this study, using data from SafeGraph Inc (2020), we utilise the Hierarchical-DBSCAN algorithm (Campello et al., 2013), which is arguably a more intuitive alternative. We also utilise network distance matrices and H3 geometries to generate more accurate clusters, and a summary of these steps can be seen below in Figures 1 & 2.

A major challenge with studies like this is having access to consolidated retail location data. As part of their "COVID-19 Data Consortium", SafeGraph provide researchers with open-access to their datasets. Thus, SafeGraph 'core places' were used to derive the retail locations, firstly as the only suitable open-access database available for the US, but also as it is highly accurate, being updated every month to reflect real-time openings and closures. We find that the approach used here resulted in an accurate representation of retail centre space across the MSA, identifying 1,599 retail centres, comprising a total of 54,891 individual retail locations. There were however issues with scalability in computation of the network distance matrices, and we also argue that use of building extents and land-use polygons would better capture the true footprint of these centres. However, in order to verify the centre extents delineate here, we examine them in three case study areas, such as below in Figure 3.

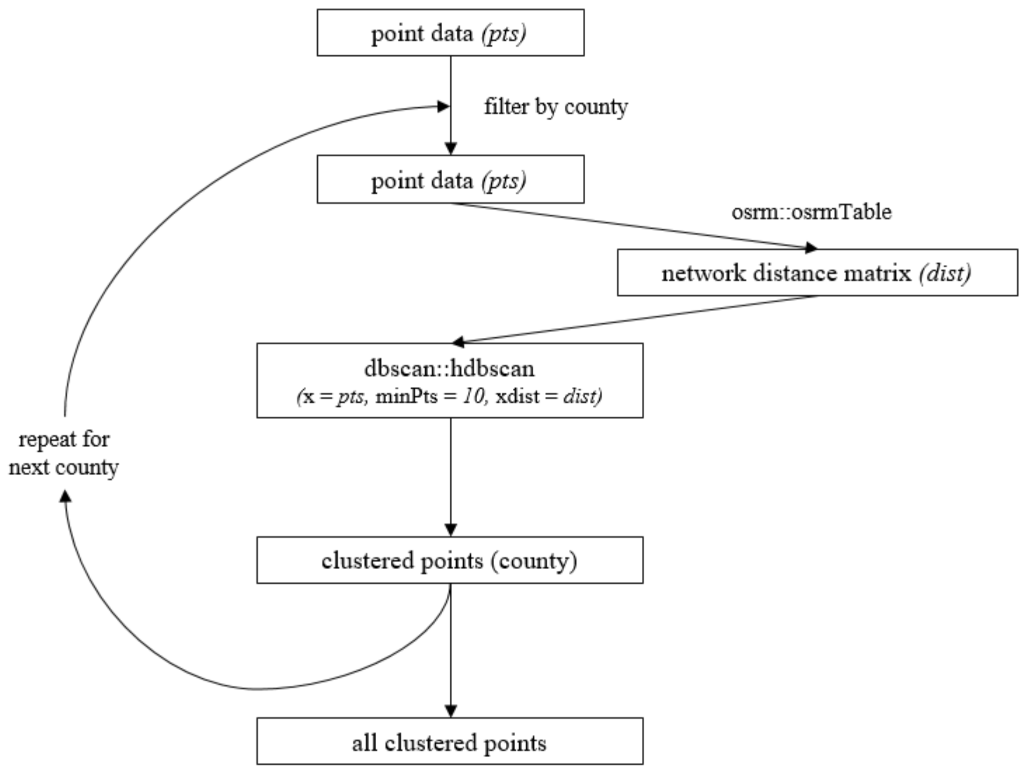

Figure 1. Iterative application of network based HDBSCAN for delineation of retail centres.

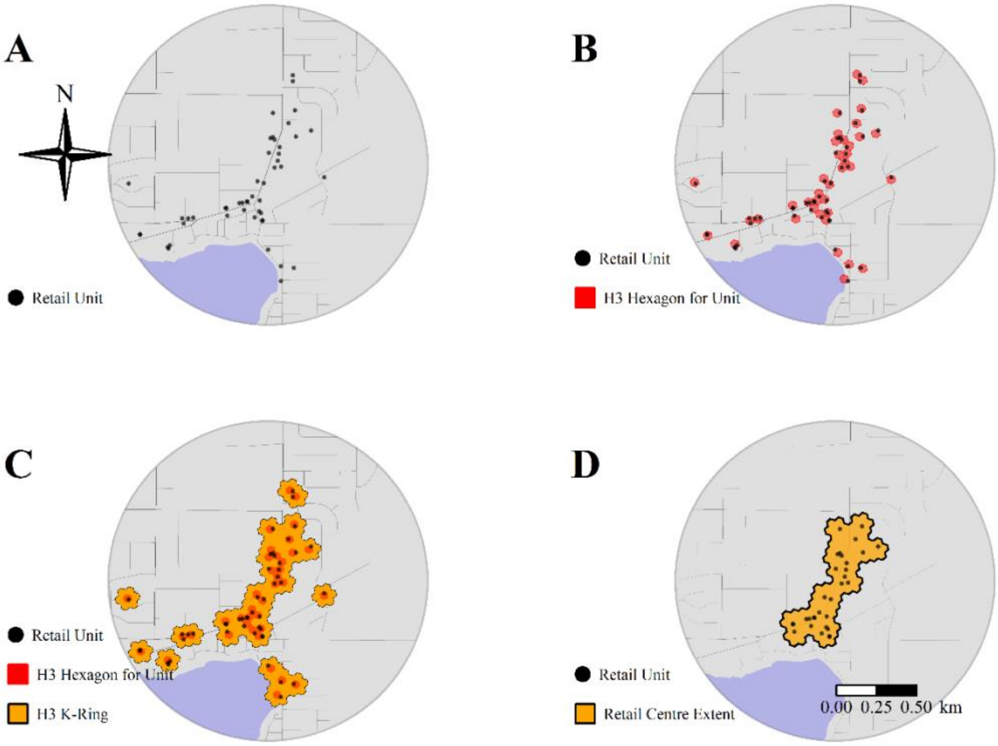

Figure 2. Cluster refinement using H3 geometries

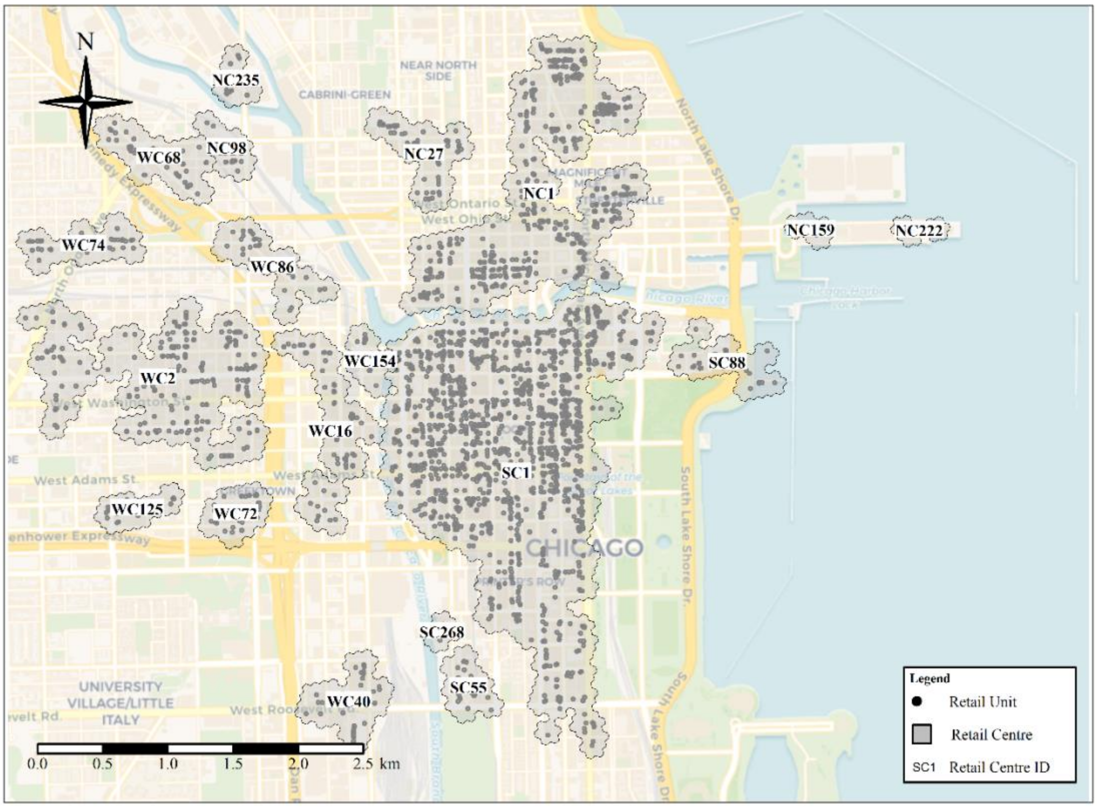

Figure 3. Delineated retail centres for Chicago CBD

3. A Typological Perspective on Retail Agglomeration within the Chicago MSA

In recent literature, there has been a call for classifications that better comprehend changes and shifts in retail provision (Grewal et al., 2017). Dolega et al. (2019) developed a ‘non-hierarchical’ typology of retail centres, using a data-driven approach and four domains (Table 1) to build a classification that arguably provides a more nuanced way of representing the relationships between centres. In this study, we adopt the approach of Dolega et al. (2019), extracting variables for each of the domains, and using unsupervised machine learning (PAM) to build the final retail classification for the 1,599 retail centres delineated in Section 2. The output of this classification was a set of five retail centre groups, containing ten nested types. The utility of these was enhanced by providing pen portraits (based on variable median values) for each type, and a summary of these can be seen below in Table 2. As in Section 2, we consider the validity of this classification in three case study areas, and an example of this can be seen below in Figure 4.

Table 1. The retail centre classification framework (Dolega et al., 2019), and variables used to implement the framework in this study

Domain	Domain Description	Variables
Composition	Classifying retail centres by the types of store present and purposes of shopping trip	Proportion of comparison, convenience, service and leisure units in each centre
Diversity	Focusing on the variety of goods and services offered, and the variety of ownership of stores in each centre	Proportion of independent retail units, diversity of SafeGraph 'top-categories'
Size & Function	Identifying the various roles of retail centres and the ways in which they interact with catchment geodemographics	No. of units, linearity (roeck score), median distance travelled, proportion of catchment population occupied by geodemographic groups
Economic Health	Exploring economic performance of retail centres by measuring the drivers of its vitality and viability	Median dwell time, median weekly visits

Table 2. Salient characteristics of retail centre groups and types

Supergroup	Group	Key Characteristics
1. Large Multipurpose Centres & Historic Retail Cores	1.1 Large Multipurpose Centres & Historic Retail Cores	Large no. of visits, diverse retail offering with higher proportions of service than other retail types
2. Popular Comparison Destinations	2.1 Leading Comparison Destinations	Greater visit frequency, diverse offering, serving neighbourhoods of "Wealthy Nuclear Families" and "Old Wealthy White"
	2.2 Secondary Comparison Destinations	Lower visit frequency, less diverse offering, larger number of chain retailers, serving "Middle Income Families", "Low Income Families" and Hispanic and Kids"
3. Leisure Strips	3.1 Inner City Leisure	Highest proportions of leisure and independents, concentrated in "Wealthy Urbanite" neighbourhoods
	3.2 Popular Suburban Leisure Centres	High proportions of leisure and independents, longer dwell times, "Middle Income" and "African-American Adversity" neighbourhoods
	3.3 Secondary Suburban Leisure Centres	Compact, higher proportions of chain leisure-based retail, concentrated in "Low-Income" and "Hispanics and Kids" neighbourhoods
4. Independent Service Centres	4.1 Diverse Service Centres in Affluent Neighbourhoods	Highest proportions of services, compact, large number of independents, serving neighbourhoods of "Wealthy Nuclear Families" and "Old Wealthy White"
	4.2 Service Centres in Hispanic and Low-Income Neighbourhoods	Large number of independents, longer dwell times, concentrated in "Hispanic & Kids", "Low Income" and "African-American Adversity" neighbourhoods
5. Small, Local Convenience Centres	5.1 Primary Convenience Centres	Greater visits, compact, dominant in neighbourhoods of "Wealthy Nuclear Families", but also "Old Wealthy White"
	5.2 Secondary Convenience Centres	Smaller and more linear, less visits, less chain retailers, longer dwell times, found in many different geodemographic neighbourhoods

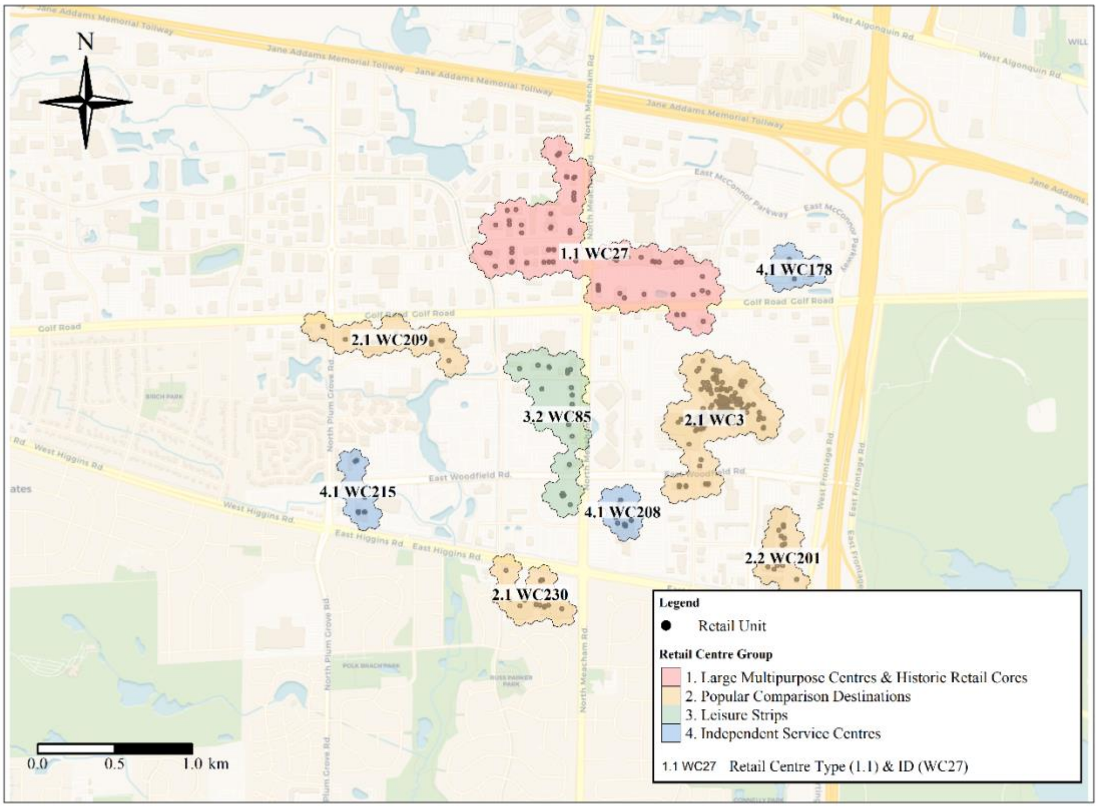

Figure 4. Delineated and classified retail centres for Schaumburg Village

4. A Case Study of the Retail Sector Response to the COVID-19 Pandemic

Owing to its nature as a highly contemporary issue, research into the impacts of COVID-19 on retail, and particularly different retail structures, is limited. Here we use the retail centres delineated for Chicago MSA and SafeGraph 'weekly patterns' data to explore the role of COVID-19 on retail sector dynamics. Patterns data was aggregated to retail centre group level, to enable exploration of how shares in mobility (%) differed between contrasting structures and functions of retail, as seen below in Figure 5.

We find evidence of a significant decrease in retail centre mobility following the issuing of the 'Stay at Home' order in Illinois (Pritzker, 2020), with notable disparities between different types of centre and different mobilities. For example, dramatic decreases in retail centre visits were identified, with centres such as 'Large Multipurpose Centres & Historic Retail Cores' losing shares in visits; whilst at the same time the 'Small, Local Convenience Centres' saw increases (Figure 5). This trend we believe suggests a general shift away from large city centre agglomerations in the early weeks of the pandemic, towards the smaller and more local ones which characteristically offered greater proportions of 'essential goods', a trend identified in other studies (Roggeveen and Sethuraman, 2020; Simbolon and Riyanto, 2020). We unpack some of these trends further, and examine the extent to which other mobilities experienced significant alterations in the early weeks of the pandemic, and how these differed between different retail centre groups.

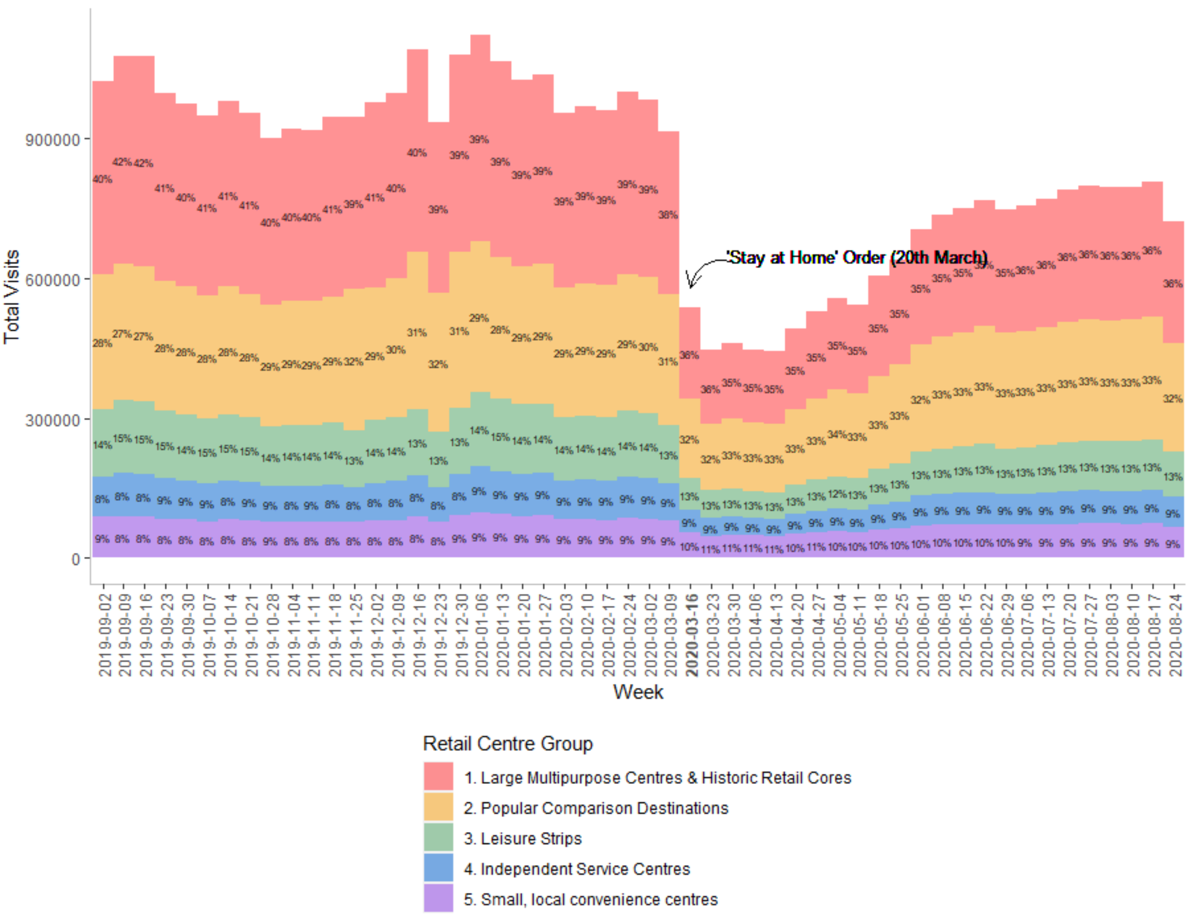

Figure 5. Change in total retail centre visits over the 12-month period, disaggregated by retail centre group to illustrate change in share (%).

5. Conclusions

In this study we have, for the first time, extended a data driven framework for the derivation of retail centre agglomerations (as in Pavlis et al., 2018 and Dolega et al., 2019) to the US, specifically considering the Chicago MSA. We have also demonstrated a significant use-case for the centres, using them to exhibit changes in retail mobility during the COVID-19 pandemic, highlighting in particular some interesting disparities between different retail structures and functions. We find SafeGraph 'core places' useful in generating a comprehensive insight into the spatiality of retailing in Chicago, offering much better coverage than equivalent open-source datasets (e.g. OpenStreetMap).

We propose that there is significant scope for a similar study of the scale extent and characteristics of retail centre agglomerations, but for the entirety of the US. Future research is however needed to ensure a computationally more scalable approach to retail centre delineation. The benefit of such an extension from a substantive perspective is that this would generate more comprehensive insights into the spatiality of local, regional and national retail provision, and provide a set of tools through which to understand how American retail as a whole, continues to transform in the wake of the 'American Retail Apocalypse' and the COVID-19 pandemic.

Acknowledgements

We would like to thank SafeGraph Inc. for permitting access to various SafeGraph datasets (core places, weekly place patterns) as part of their COVID-19 Data Consortium.

References

Baker, S. R., Farrokhnia, R. A., Meyer, S., Pagel, M., & Yannelis, C. (2020). How Does Household Spending Respond to an Epidemic? Consumption During the 2020 COVID-19 Pandemic. *SSRN Electronic Journal*, (March). https://doi.org/10.2139/ssrn.3565521

Berman, B., & Evans, J. R. (2013). *Retail Management: A Strategic Approach* (12th Editi). Cambridge: Harvard University Press.

Bonaccorsi, G., Pierri, F., Cinelli, M., Flori, A., Galeazzi, A., Porcelli, F., … Pammolli, F. (2020). Economic and social consequences of human mobility restrictions under COVID-19. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(27), 15530–15535. https://doi.org/10.1073/pnas.2007658117

Brown, S. (1992). *Retail Location: A Micro-Scale Perspective*. Ashgate Publishing Limited.

Campello, R. G. ., Moulavi, D., & Sander, J. (2013). Density-Based Clustering on Hierarchical Density Estimates. In Pei, Tseng, Cao, Motoda, & Xu (Eds.), *Advances in Knowledge Discovery in Databases* (pp. 160–173). Springer. https://doi.org/10.1007/978-3-319-13212-9

Dolega, L., Reynolds, J., Singleton, A., & Pavlis, M. (2019). Beyond retail: New ways of classifying UK shopping and consumption spaces. *Environment and Planning B: Urban Analytics and City Science*, 1–30. https://doi.org/10.1177/2399808319840666

Grewal, D., Roggeveen, A. L., & Nordfält, J. (2017). The Future of Retailing. *Journal of Retailing*, *93*(1), 1–6. https://doi.org/10.1016/j.jretai.2016.12.008

Han, Z., Cui, C., Miao, C., Wang, H., & Chen, X. (2019). Identifying spatial patterns of retail stores in road network structure. *Sustainability (Switzerland)*, *11*(17). https://doi.org/10.3390/su11174539

Lloyd, A., & Cheshire, J. (2017). Deriving retail centre locations and catchments from geo-tagged Twitter data. *Computers, Environment and Urban Systems*, *61*, 108–118. https://doi.org/10.1016/j.compenvurbsys.2016.09.006

Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., & Al-jabir, A. (2020). The socio-economic implications of the coronavirus pandemic (COVID-19): A review. *International Journal of Surgery*, *78*, 185–193.

Pavlis, M., Dolega, L., & Singleton, A. (2018). A Modified DBSCAN Clustering Method to Estimate Retail Center Extent. *Geographical Analysis*, *50*(2), 141–161. https://doi.org/10.1111/gean.12138

Pritzker, J. B. Executive Order 2020-32 (COVID-19 Executive Order No. 32) (2020). Retrieved from https://www2.illinois.gov/Pages/Executive-Orders/ExecutiveOrder2020-32.aspx

Roggeveen, A. L., & Sethuraman, R. (2020). How the COVID-19 Pandemic May Change the World of Retailing. *Journal of Retailing*, (January).

SafeGraph Inc. (2020). SafeGraph COVID-19 Data Consortium. Retrieved December 7, 2020, from https://www.safegraph.com/covid-19-data-consortium

Simbolon, R., & Riyanto, S. (2020). How Retail Survive Against Pandemic of COVID-19 : An Insight from Optical Retailer. *International Journal of Current Science and Multidisciplinary Research*, *3*(05), 124–132.

Biographies

Patrick Ballantyne is a PhD student at the Geographic Data Science Lab, University of Liverpool. His research focuses on quantitative and spatial approaches to understanding retail environments, with a particular interest in those outside the UK.

Prof. Alex Singleton is a Professor of Geographic Information Science at the University of Liverpool. His research focuses on applications of Urban Analytics and Geographic Data Science, with a particular focus on Geodemographics.

Dr. Les Dolega is a Lecturer in the Department of Geography and Planning at the University of Liverpool. His research focuses on GIS, retail and consumption spaces and economic geography.

Dr. Kevin Credit is a Lecturer at the National Centre for Geocomputation at Maynooth University. His research focuses on better understanding how urban spatial structure and transportation systems influence economic, environmental, and social outcomes using quantitative approaches and large open-source datasets.