

Exploring vernacular perceptions of spatial entities: Using Twitter data and R for delimiting vague, informal neighbourhood units in Inner London, UK.

Luke Clasper
UNIGIS UK

The University of Salford
l.t.clasper@edu.salford.ac.uk

What is Vernacular Geography?

- Folksonomy and the informal, bottom-up, unofficial, colloquial nature of how people discuss and mentally conceive geographic place and structure.
- Individual's awareness of fuzzy, abstract geographic regions (official and unofficial) in relation to their own location.
- Important applications:
 - Emergency Services
 - Deliveries
 - In-vehicle Navigation
 - Allocations of Services and Census

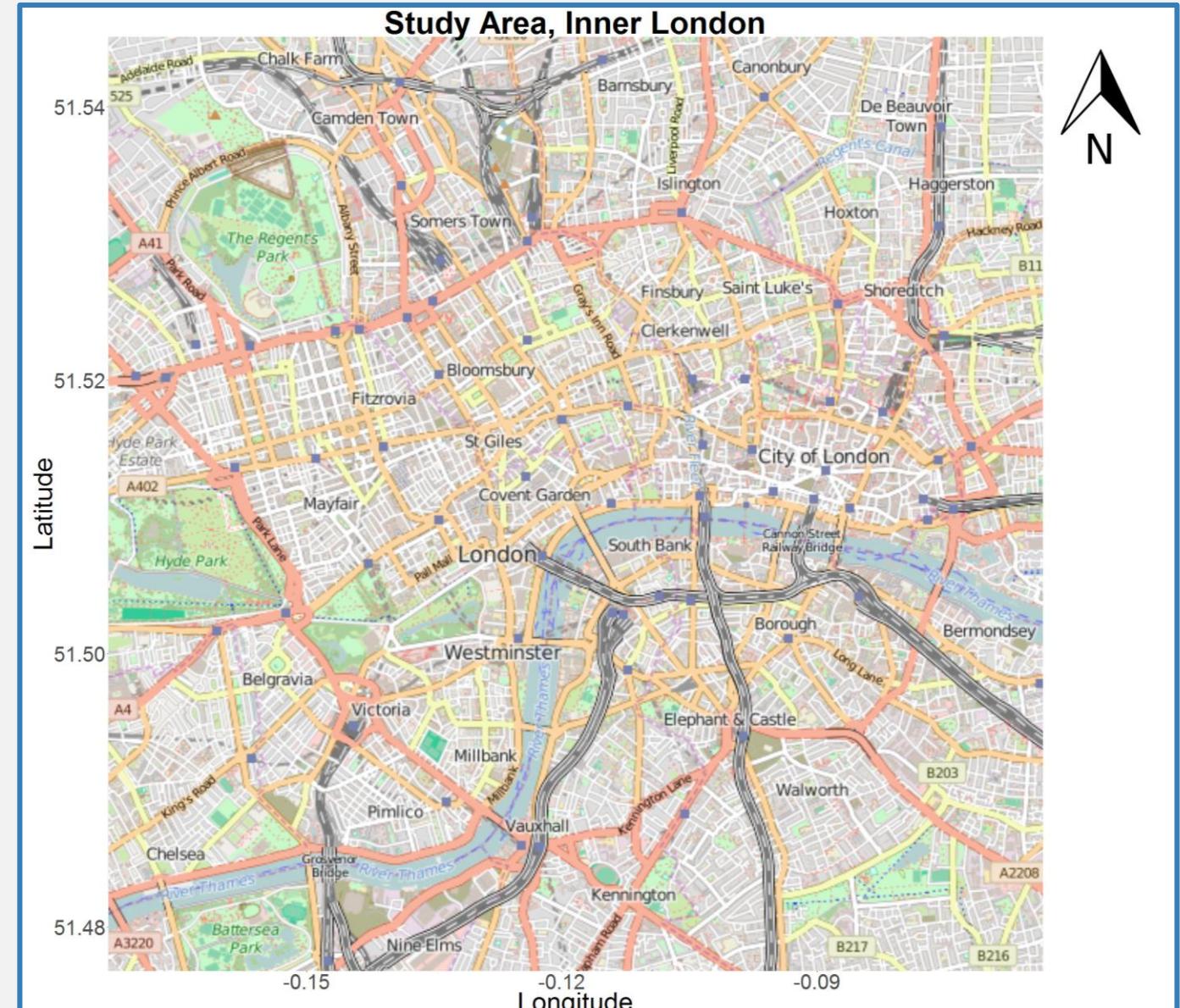
How to Capture Vernacular Geography?

- Difficult: personal, casual, vague, qualitative.
- Traditional methods like questionnaires and sketch maps.
- Web scraping.
- Flickr.
- **VGI from social media like Twitter can collect many responses from unconscious participants.**

Neighbourhoods for Study

Neighbourhood has current official boundary?
(Y/N).

- Aldgate (Y)
- Euston (N)
- Shoreditch (N)
- Barbican (N)
- Farringdon (Y)
- Soho (N)
- Bishopsgate (Y)
- Fitzrovia (N)
- Southbank (N)
- Blackfriars (N)
- Holborn (Y)
- Southwark (Y)
- Bloomsbury (Y)
- Hoxton (Y)
- Kings Cross (Y)
- Spitalfields (Y)
- Brick Lane (N)
- Lambeth (Y)
- Strand (N)
- Clerkenwell (Y)
- Leicester Square (N)
- Vauxhall (N)
- Covent Garden (Y)
- Marylebone
- Waterloo (N)
- Elephant and Castle (N)
- Paddington (N)



R Statistical Software

- Open Source.
- Language and software environment.
- Spatial capabilities.
- Research reproducibility and self documentation.
- Efficient with large datasets and repeating tasks.
- OSM (and other) base maps.
- Thousands of code libraries.
- Active community.
- Steep learning curve but rewarding.

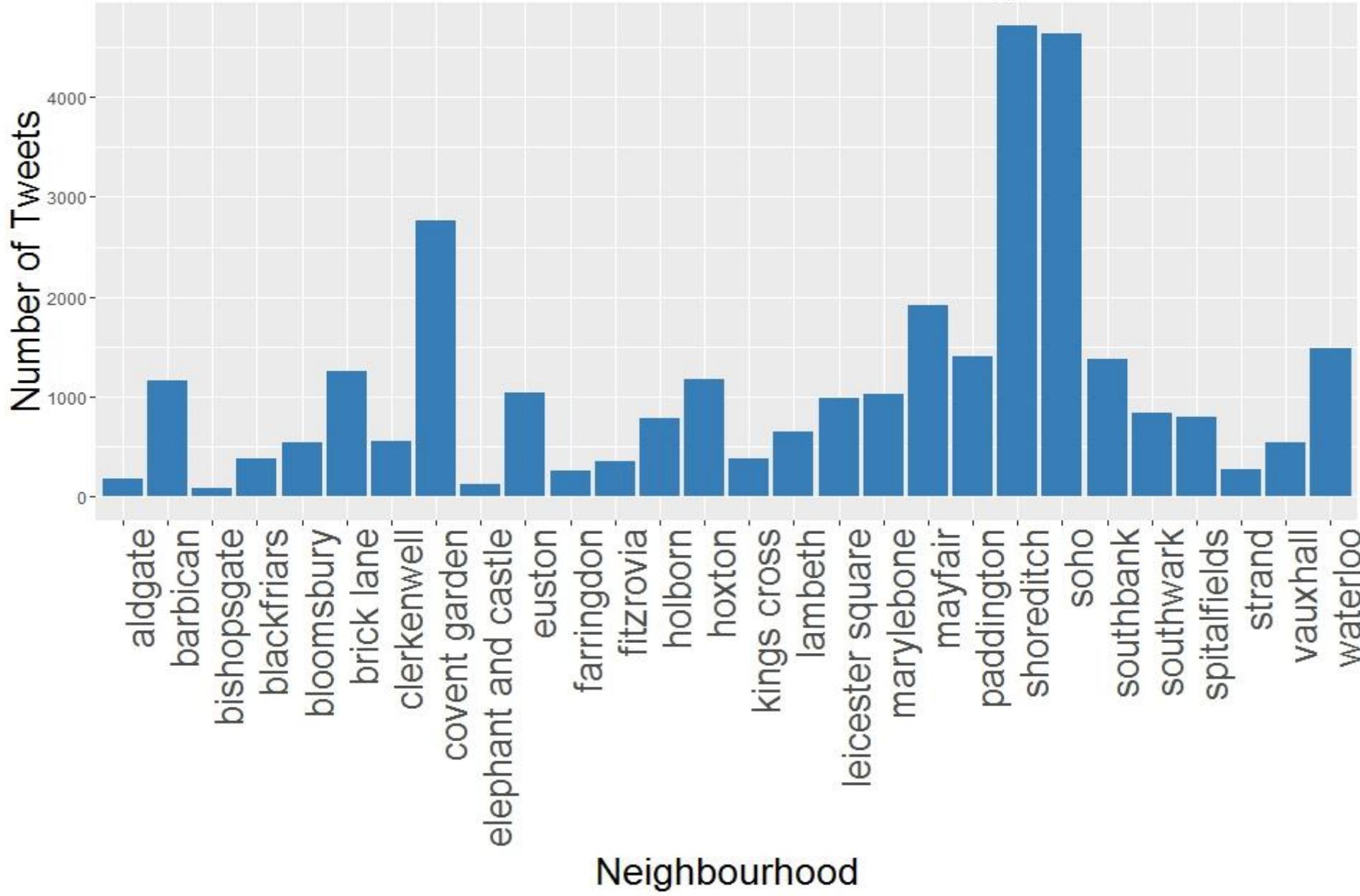
R Libraries Used

- rgdal (GDAL)
- ggmap (basemaps)
- GISTools (spatial analysis)
- sp (spatial methods)
- spatstat (point pattern analysis)
- hexbin (hexagonal binning)
- aspace (spatial point patterns)
- ggplot2 (graphics, visualisations)
- ggthemes (style themes)
- **twitteR** (connects to Twitter API)
- Httr (HTTP web data)
- RColorBrewer (colour schemes)
- geosphere (spherical trigonometry)
- reshape (data aggregation)

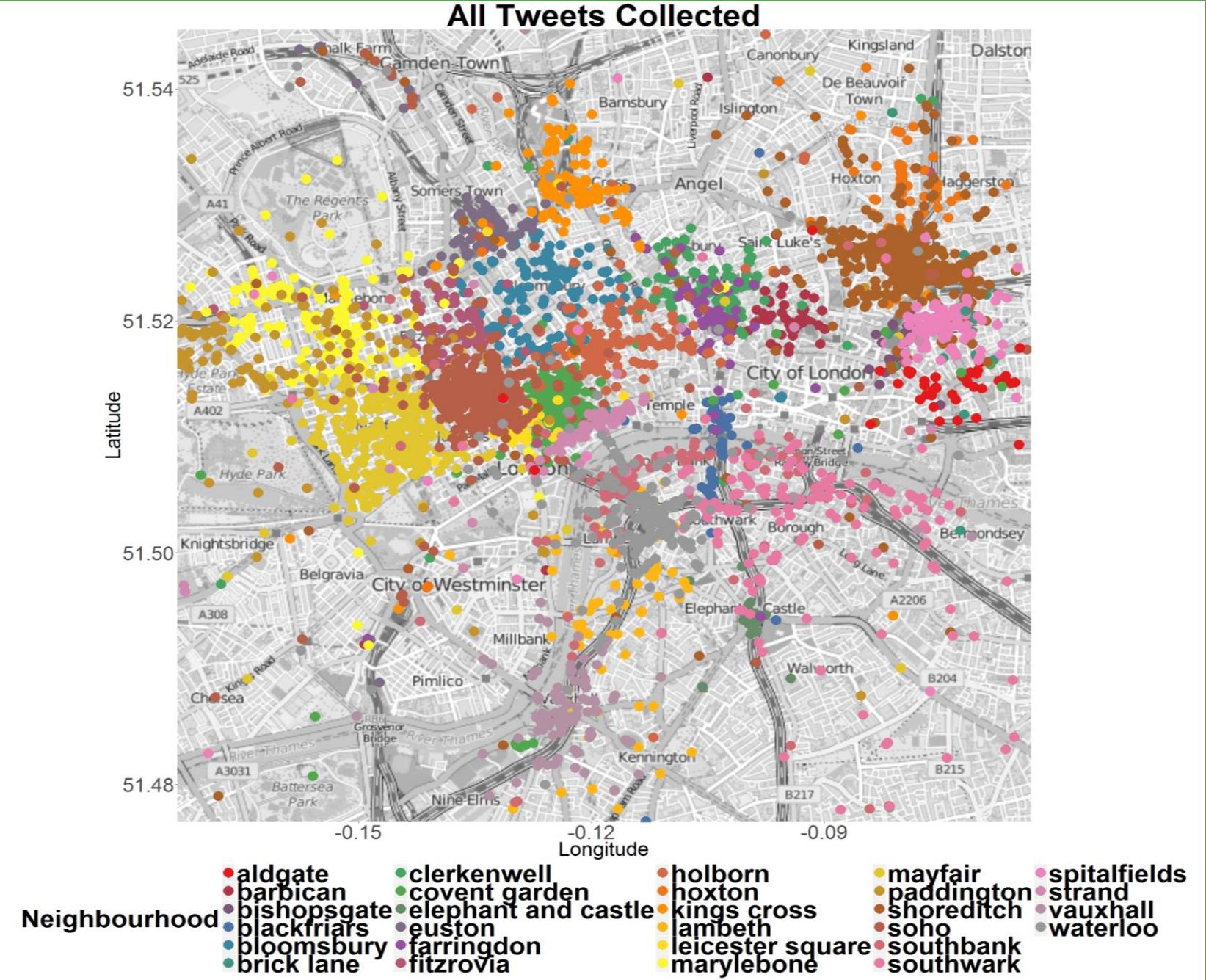
Data Collection from Twitter API

- twitteR library.
- setup_twitter_oauth() make authorised connection.
- API key, API token, API secret from Twitter.
- searchTwitter() for:
 - Looped keywords, Hastags e.g. shoreditch and #Shoreditch
 - Geocode Radius (5 miles)
 - Geocode Centroid (51.508107, -0.126449)

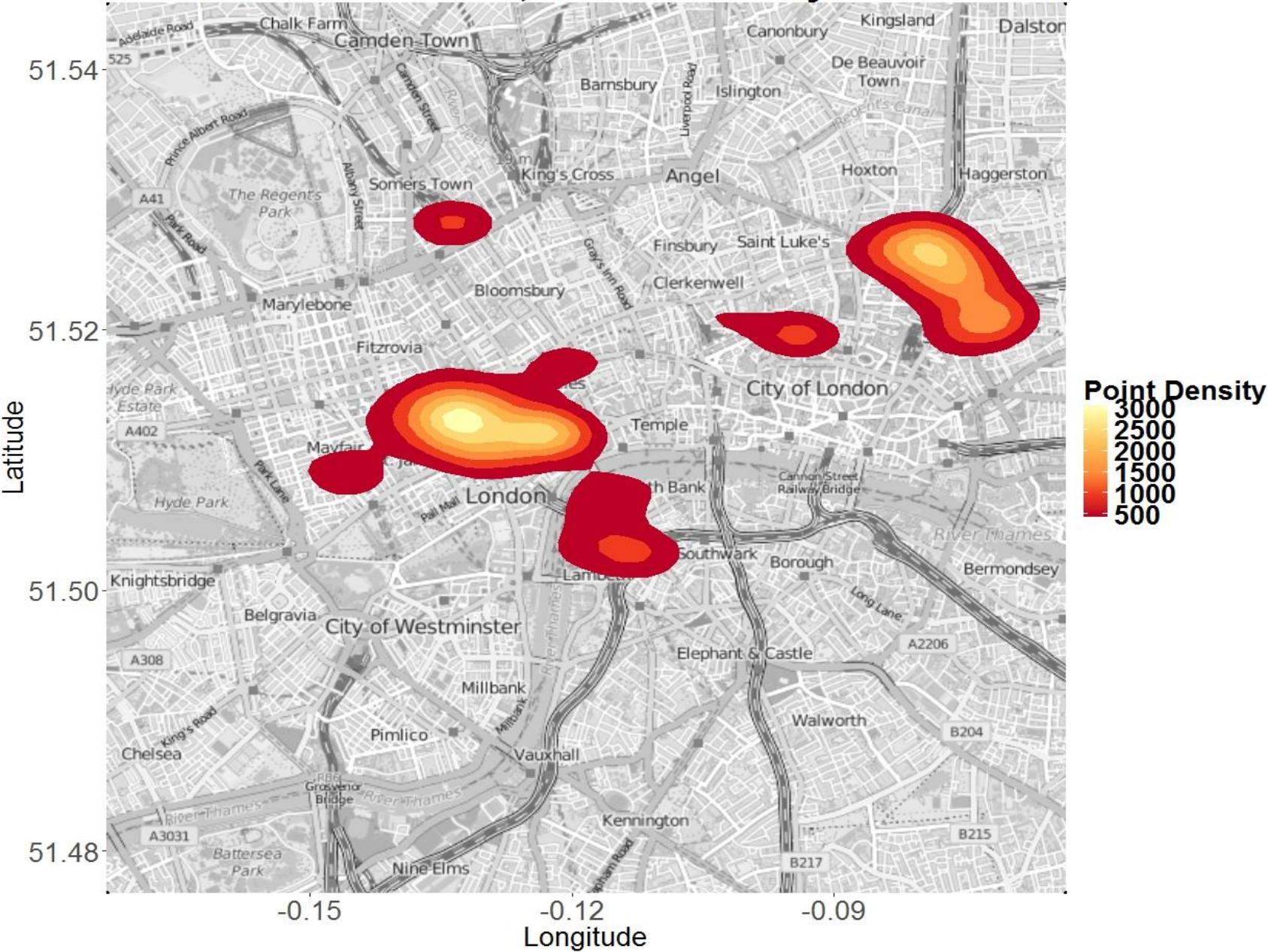
Number of Tweets Collected for each Neighbourhood



- Tweets mirroring social activity.
- Dense Tweets in West End and East End.
- City of London, Westminster and residential areas have sparse Tweets.
- Tweets replicating underlying urban structure.
- Neighbourhood clusters forming from keyword searches.



All Tweets Collected, Kernel Density Estimation



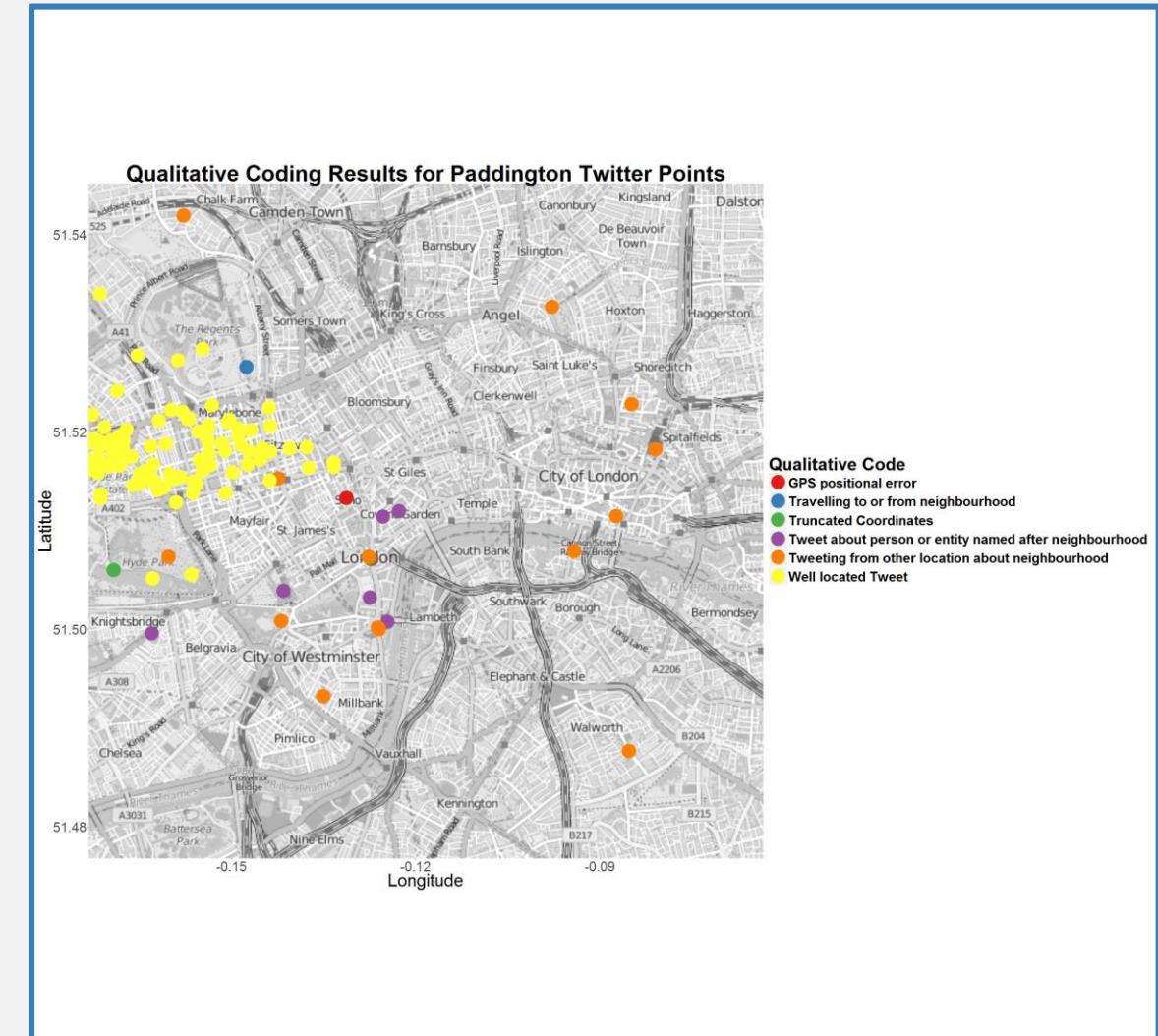
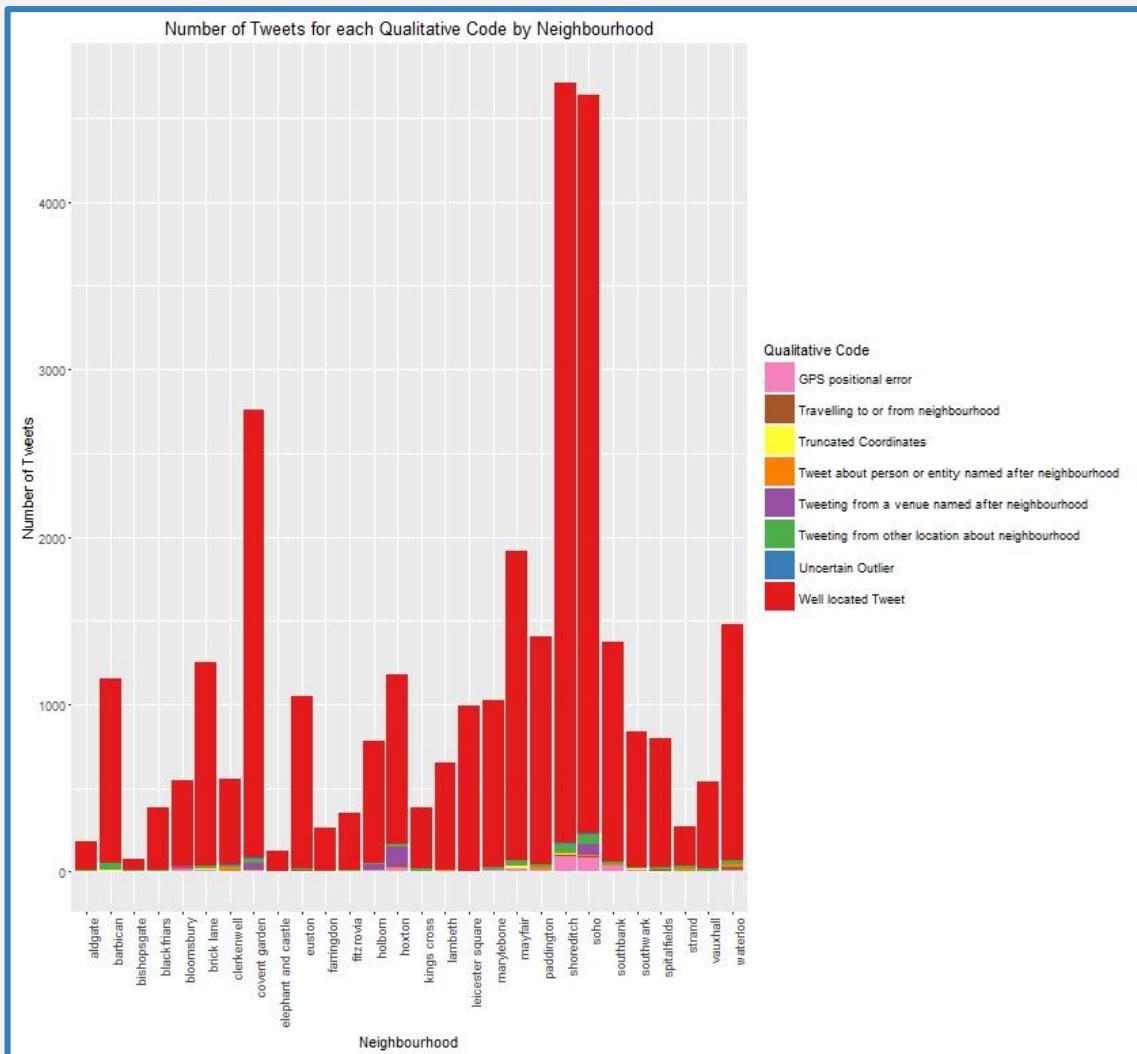
Qualitative Coding

- Social media data is unreliable (*Lovelace et al, 2016*).
- VGI from social media is qualitative.
- Qualitative coding used to improve data quality.
 - Manual examination of text topic
 - Visual analysis for precision
- A way of interpreting and filtering data in order to categorise or classify it into themes (Cope, 2003).

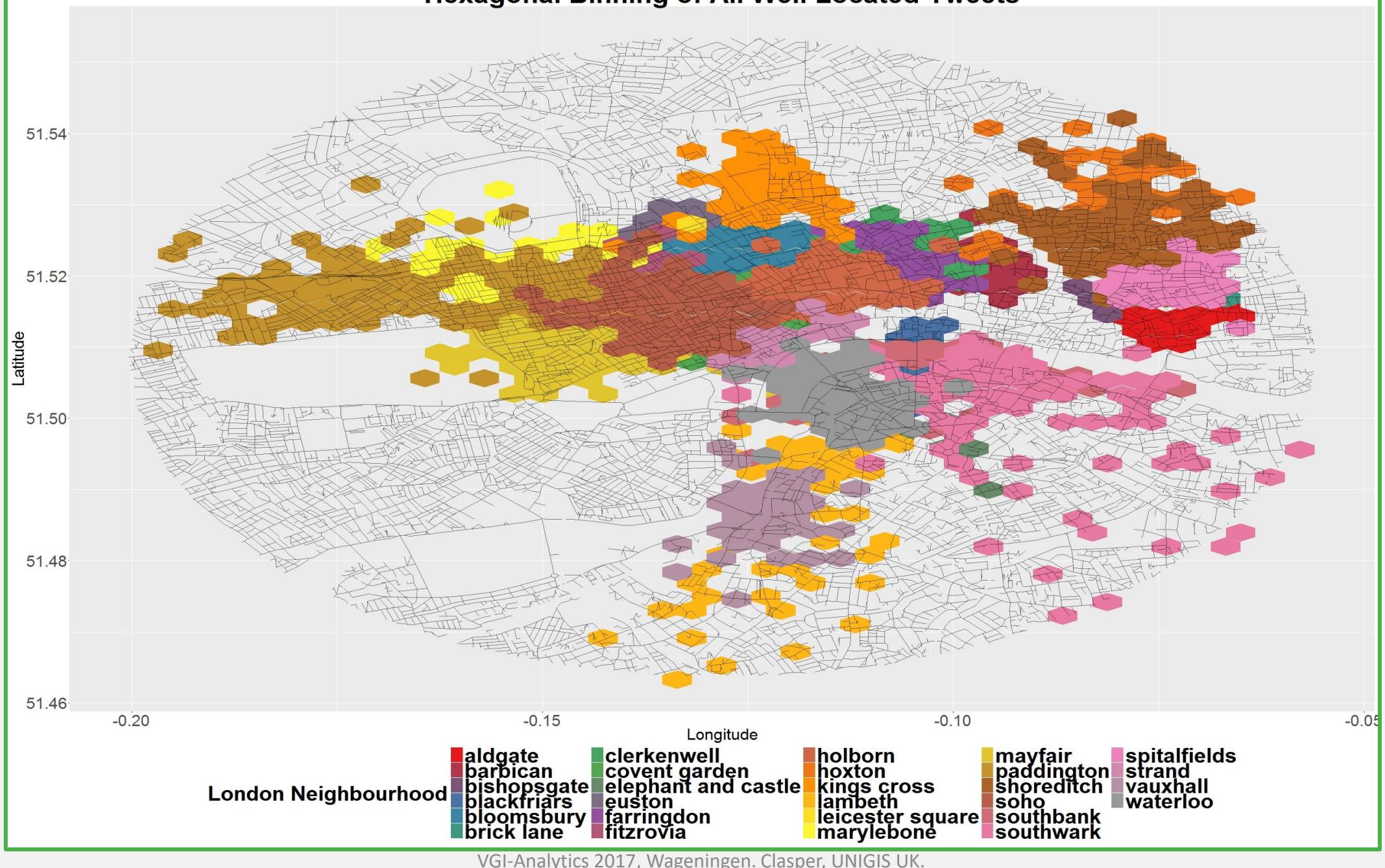
Qualitative Coding Categories

- Travelling to or from neighbourhood.
 - "Made it as far as Covent Garden en-route to Soho, gotta experience the gay night life here in London... "
- Tweeting from other locational about neighbourhood.
 - "Still feeling stuffed after yesterday's meal at @BodeansBBQ in soho. huge massive portions, can't wait to go again."
- Tweeting from a venue named after neighbourhood.
 - "Gym on a Sunday, that's how committed I am to undoing the holiday damage! (@ Soho Gyms Farringdon in London, UK)"
- Well located Tweet.
 - "I'm at Gail's Artisan Bakery in Soho"
- GPS Positional Error.
- Tweet about person or entity named after neighbourhood.
 - "Paul Strand 1890- 1976 arguably one of the greatest documentary photographers of 20th century"
- Truncated Coordinates.
- Uncertain Outlier.

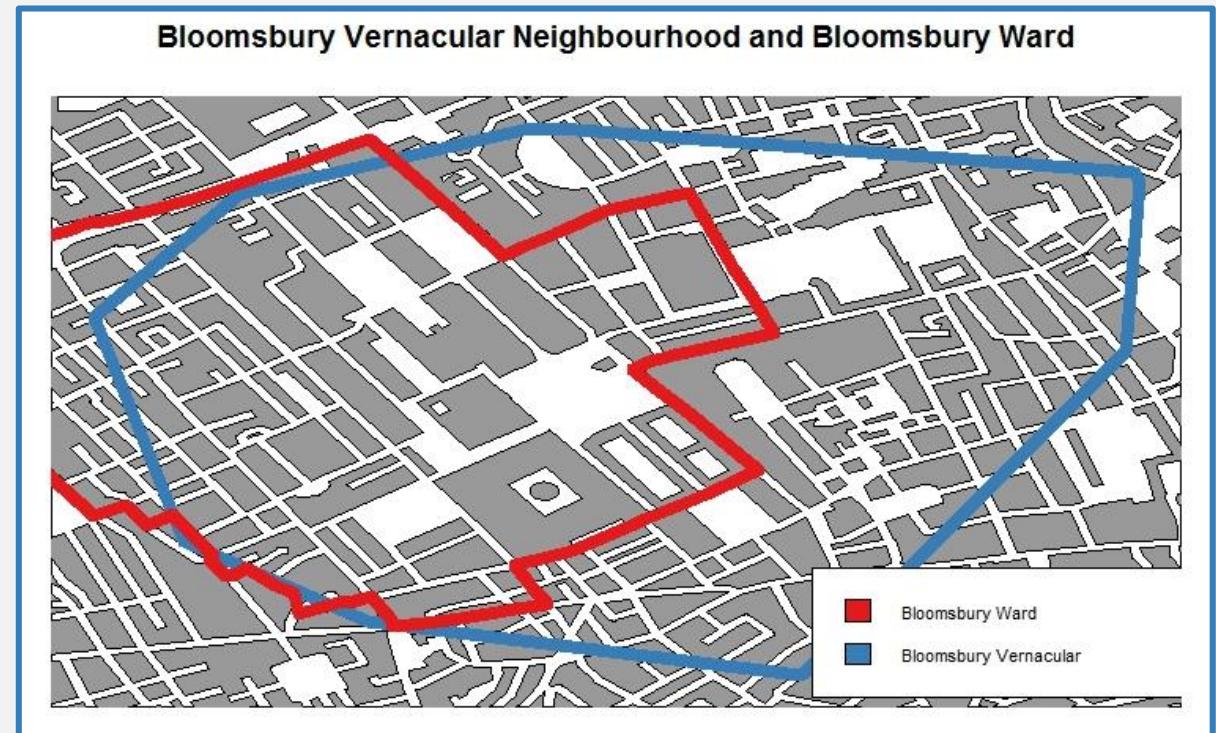
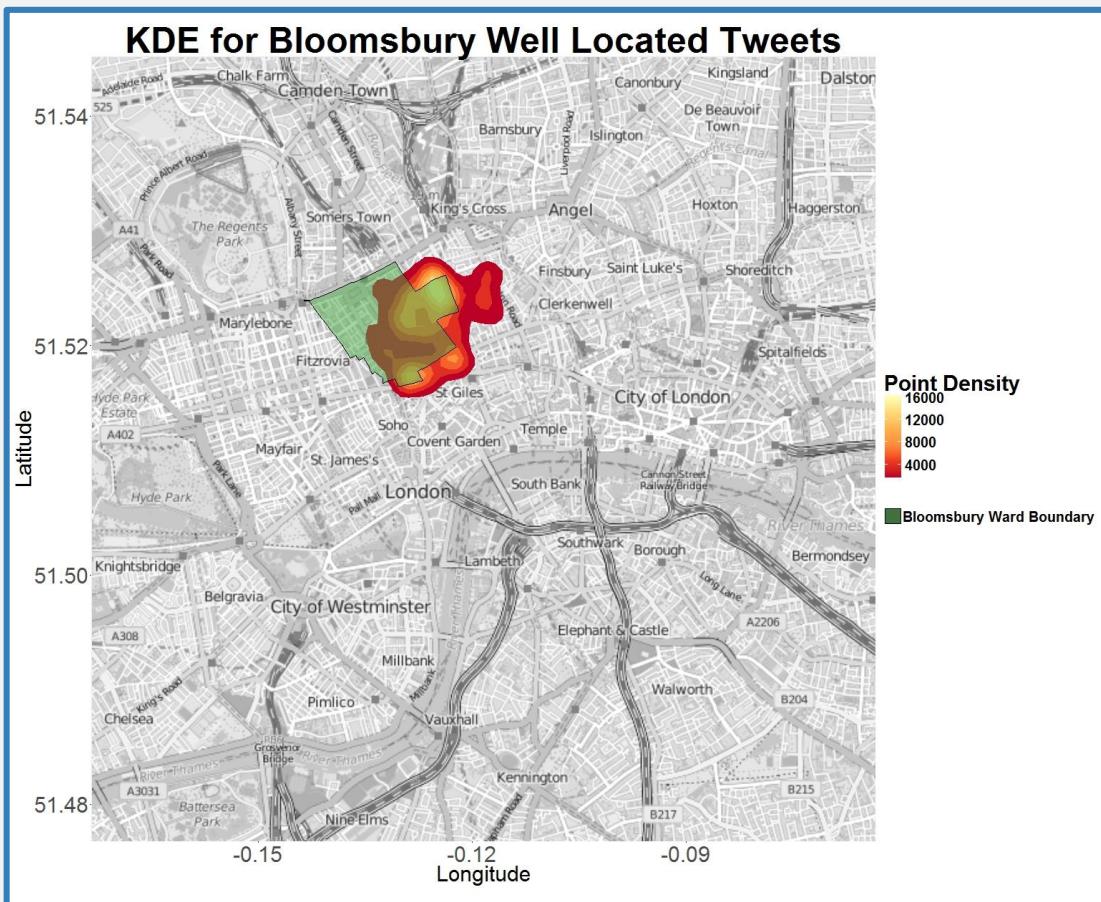
Qualitative Coding Results



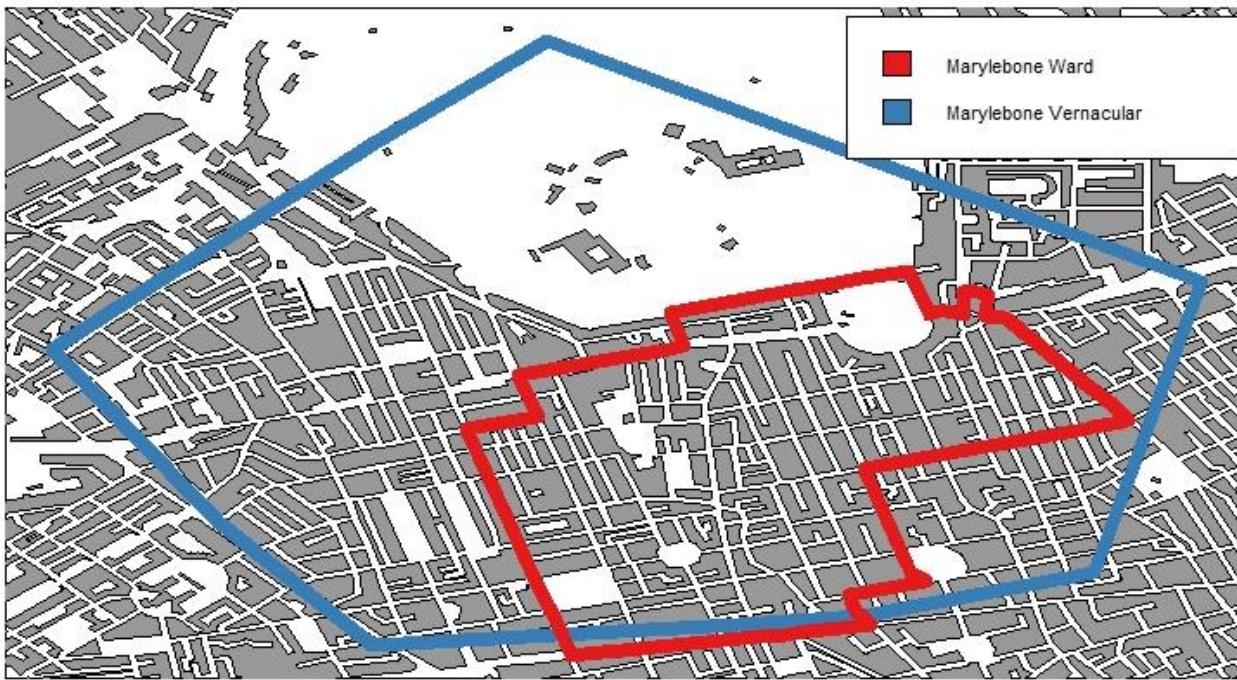
Hexagonal Binning of All Well Located Tweets



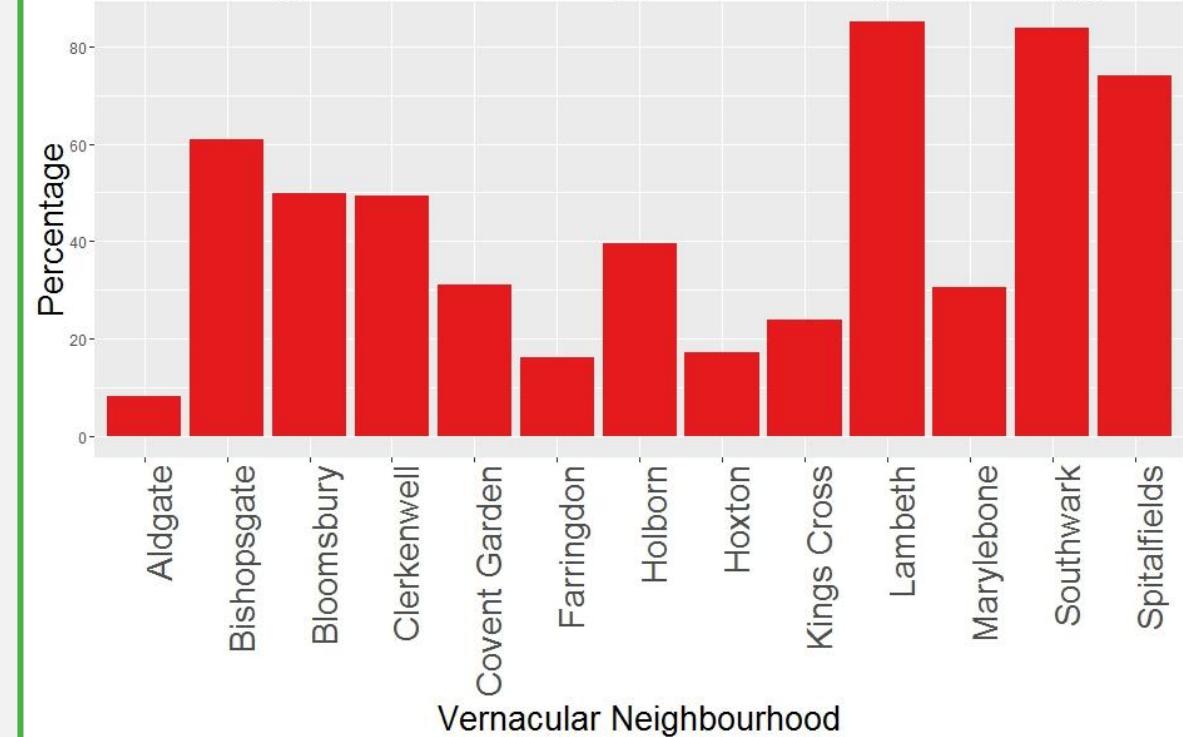
Neighbourhood Comparison to Official Administrative Boundaries



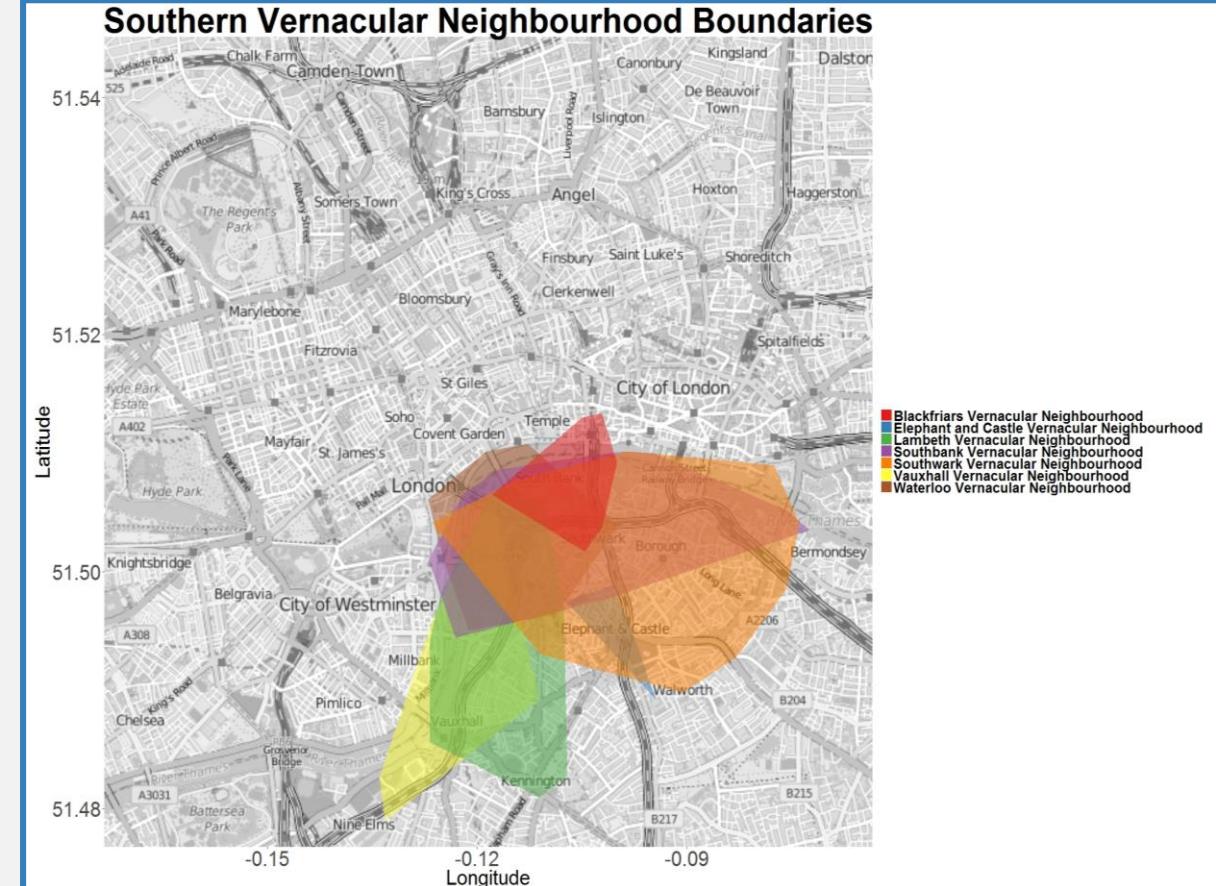
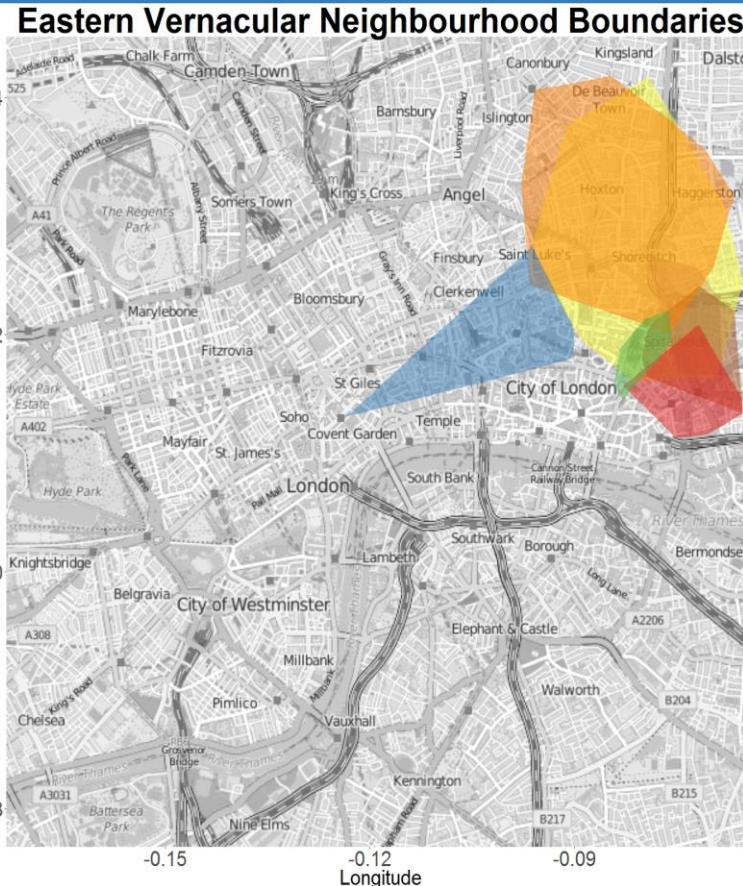
Marylebone Vernacular Neighbourhood and Marylebone Ward



Percentage of Vernacular Polygon Intersecting Ward Polygon

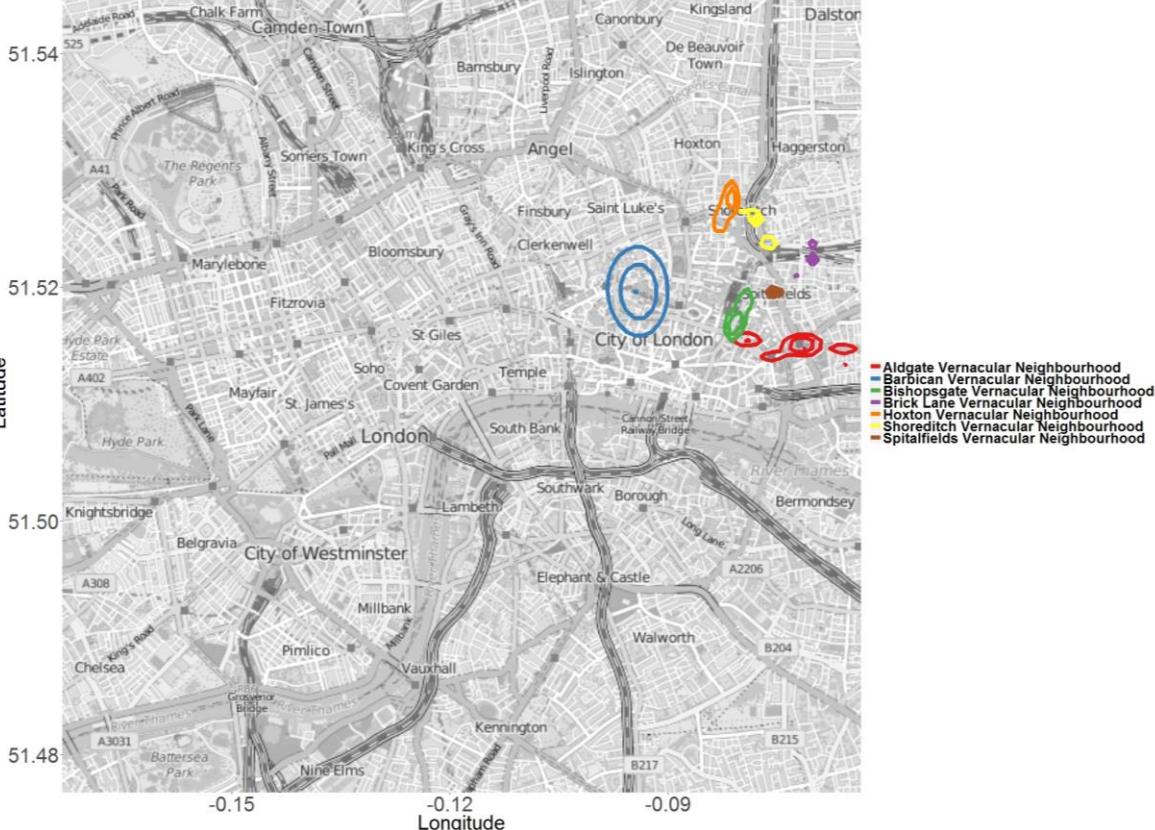


Vernacular Neighbourhood Delimitation by Convex Hull



Vernacular Neighbourhood Origins

Eastern Vernacular Neighbourhood Centres

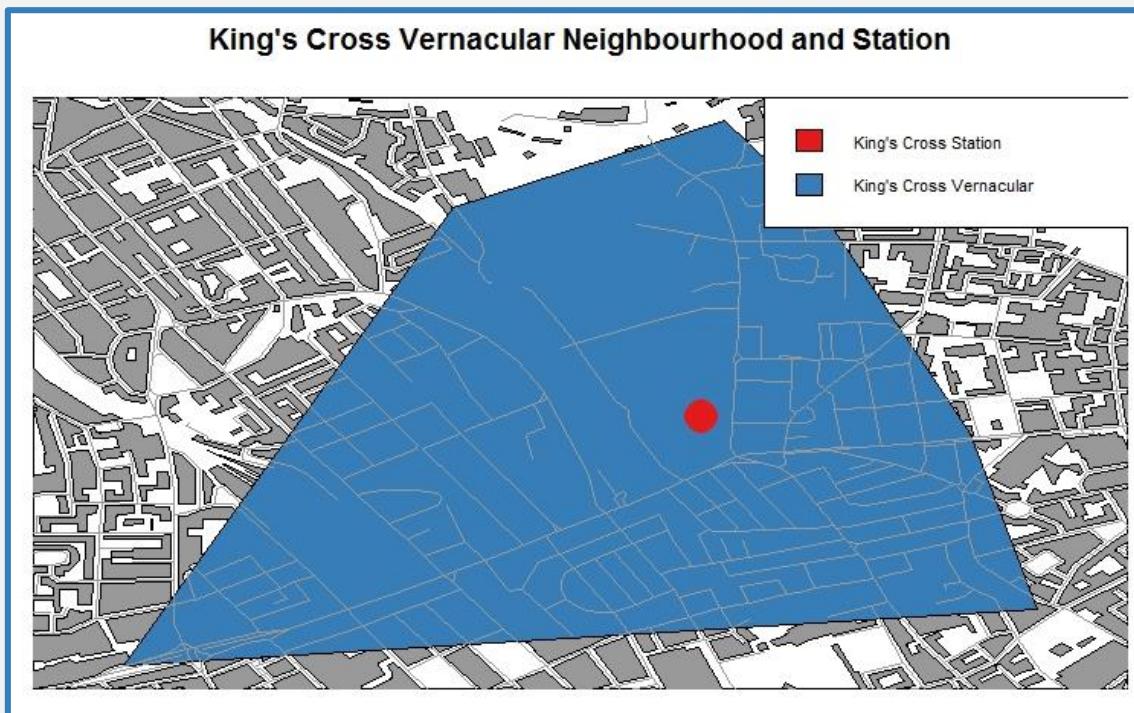


Southern Vernacular Neighbourhood Centres



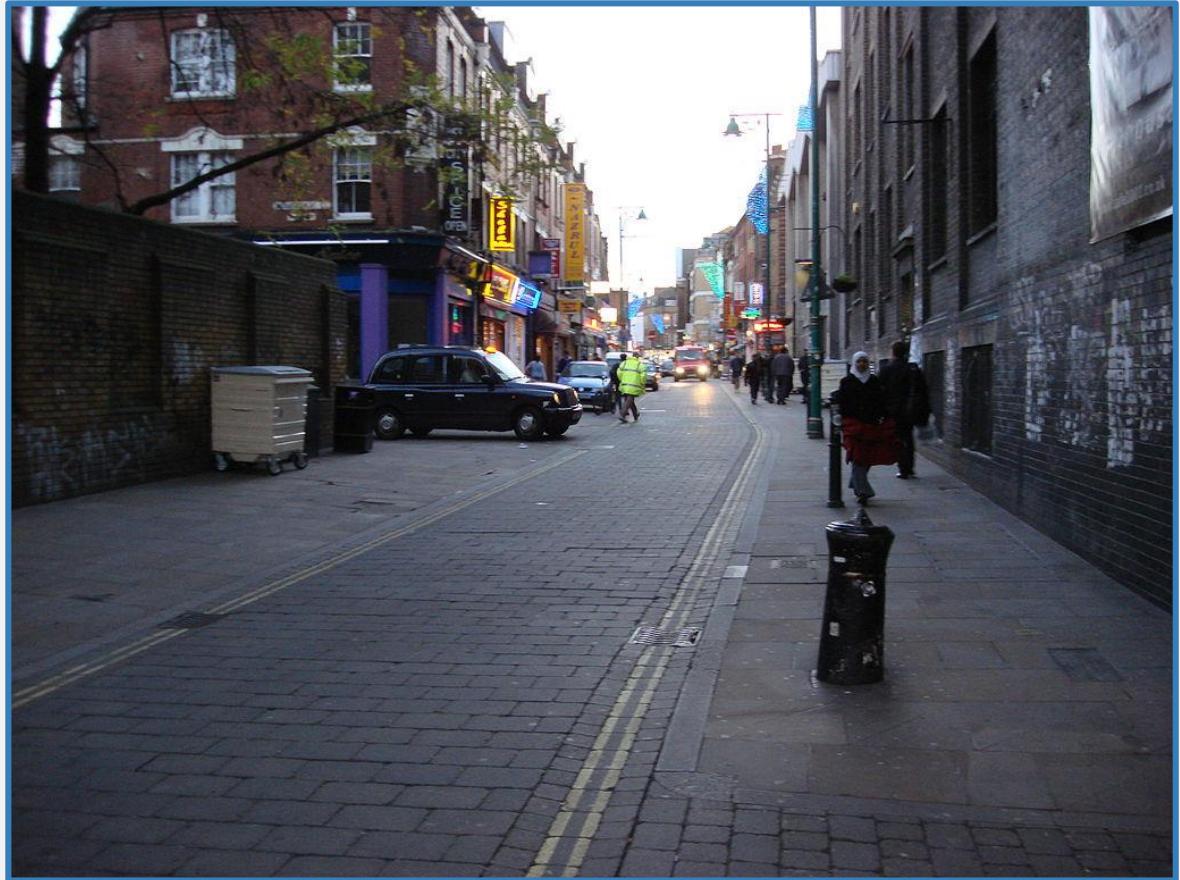
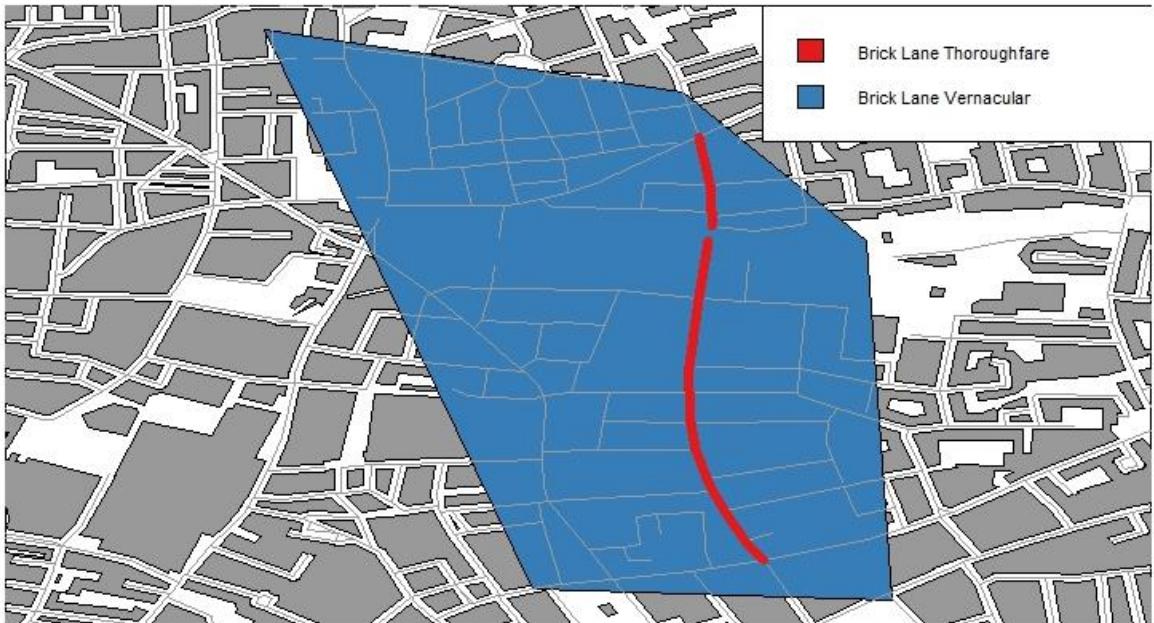
Lynch's (1960) Elements Form Basis of Vernacular Neighbourhoods

Nodes and Landmarks



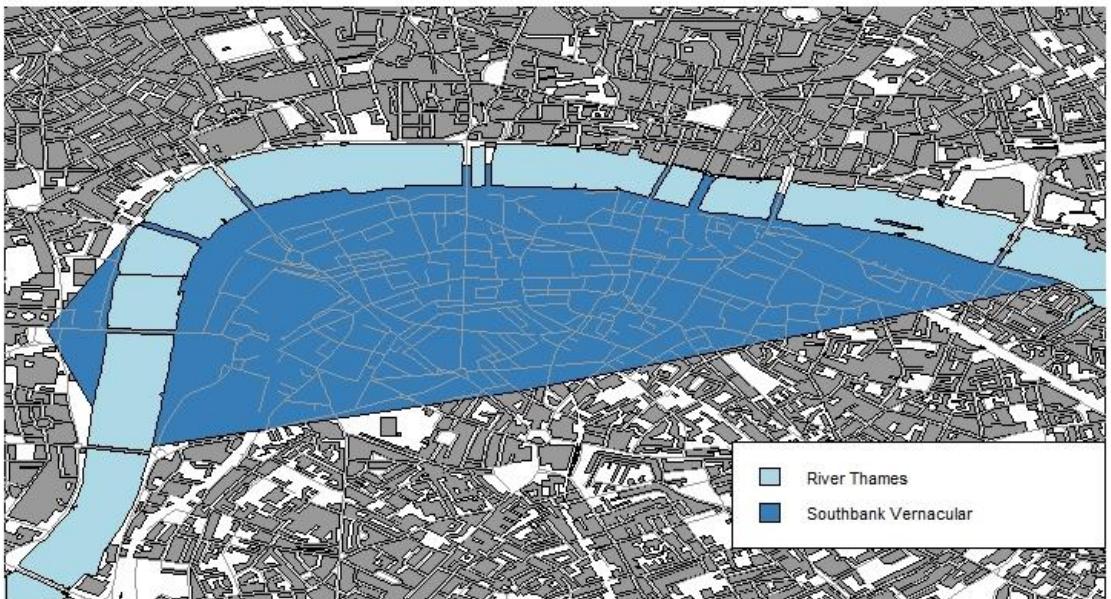
Path Elements, Lynch (1960)

Brick Lane Vernacular Neighbourhood and Thoroughfare

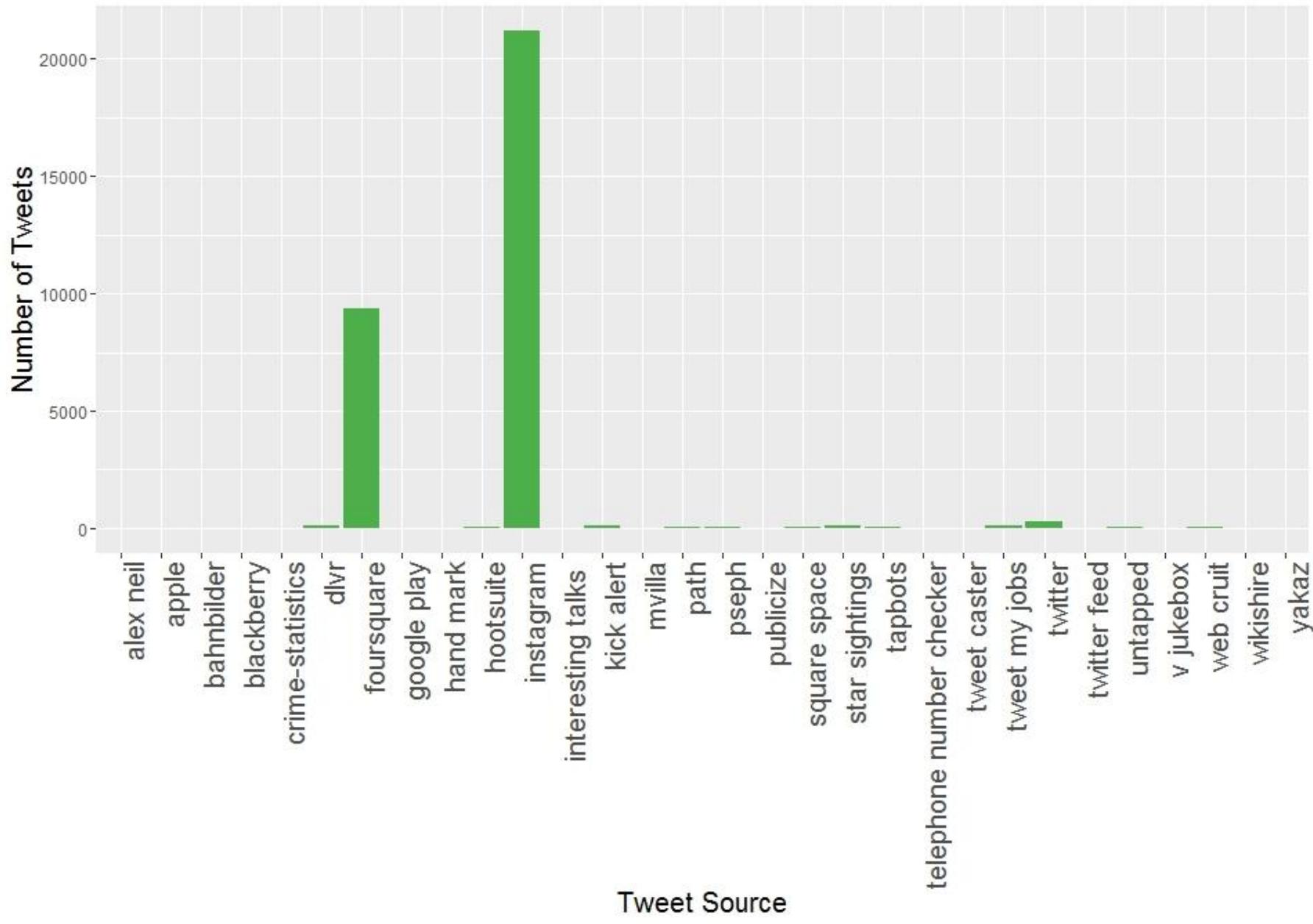


Edge Elements, Lynch (1960)

Southbank Vernacular Neighbourhood and River



Source of Tweets



Limitations of Study and Further Work.

- Qualitative Coding process subjective and time consuming.
- Twitter API Limits number of Tweets harvested.
- 31,692 Tweets collected but only 14,832 users.
- Socio/economic/cultural demographics of Twitter users.
- Machine learning algorithm to sort Tweets (in R).
- Precision and Recall (in R) to test data quality.
- Test data quality by comparing to results from Flickr.
- Apply to other cities, e.g. Paris.
- Link to OSM.

References and Acknowledgements

- Cope, M. (2003) Coding Transcripts and Diaries. In Clifford, N., French, S. & Valentine, G. (Editors) Key Methods in Geography. Chapter 27: 440-452. London: Sage.
- Lovelace, R., Birkin, M., Cross, P. & Clarke, M. (2016) From Big Noise to Big Data: Towards the verification of Large Data sets for Understanding Region Retails Flows. Geographical Analysis 48: 59-81.
- Lynch, K. (1960) The Image of the City. Cambridge (Massachusetts): The MIT Press.

Contains:

- © OpenStreetMap contributors (<http://www.openstreetmap.org/copyright>).
- Ordnance Survey OpenData © Crown Copyright/database right 2017.
- Thames and Lambeth Bridge. © Justin Norris (2007) (<https://creativecommons.org/licenses/by/2.0>).
- Brick Lane, London, E1. Looking south. © Oxyman (2007) (<https://creativecommons.org/licenses/by/2.0>).
- King's Cross Western Concourse. © Colin/London (2012) (<https://creativecommons.org/licenses/by/2.0>).