

# Consumer Data New Research Opportunities

# Guy Lansley



An ESRC Data Investment



# Who we are

We are an academic led, multi-institution laboratory which discovers, mines, analyses and synthesises consumer-related datasets from around the UK. The CDRC is an ESRC Data Investment.

www.cdrc.ac.uk

maps.cdrc.ac.uk
data.cdrc.ac.uk



# CDRC-Public

Access retail, consumer and contextual information through our free public data and mapping portal.



# CDRC-Stakeholder/Archive

A download service for retail / consumer data for use in academic research.



# CDRC-Secure

Access anonymised consumer / retail data from our secure on-site facilities.













# **Ladder of Engagement'**

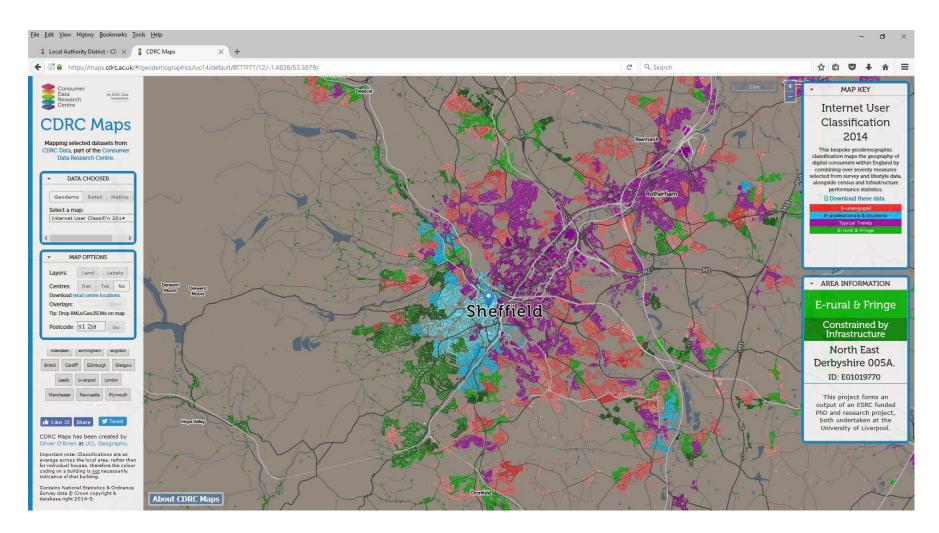
Coproduction of data Data
Sharing
with CDRC

Partner focused research projects

PhD Research

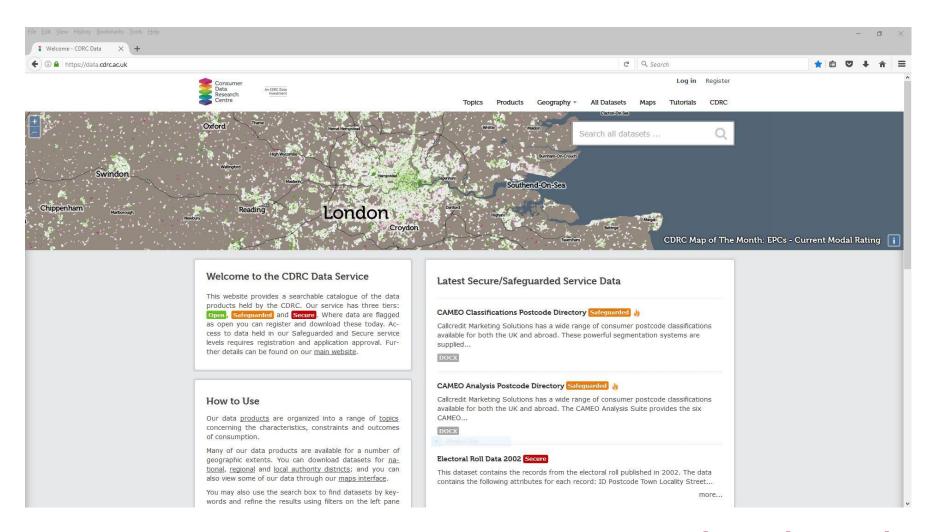
Masters Programmes

# **CDRC Maps**



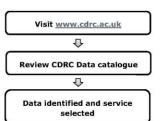


# **CDRC Data**

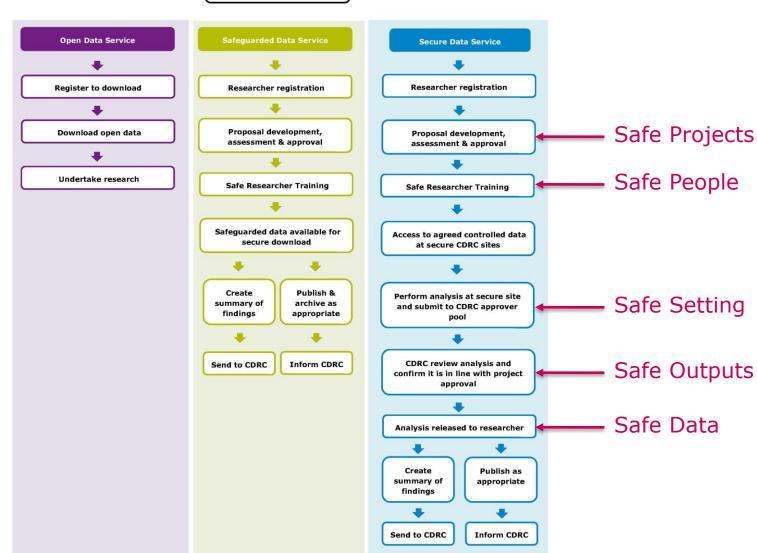


data.cdrc.ac.uk





# **User Journey**





# **Example Datasets**

Smart Meter energy consumption data from a domestic energy supplier

Electoral Roll data (1998 – 2012) and Consumer Registers (2003 – 2017)

Active Inspiration – Activity Data

Debit card transactions from a Youth Banking Card Provider

WhenFresh/Zoopla
Property Transactions
and Associated
Migration

Appliances Online Retail Data YouGov Survey Data

Synthetic Population

Online retail transactions from Shop Direct

Store transactions (linked to customer loyalty accounts) from a high street retailer

Regional Transport Provider: Ticket counts; assets, travel demand, vehicle position & transport infrastructure data

Footfall sensor data from Local Data Company

Secure Safeguarded

Acxiom – Income Data

**British Population Survey** 



# **CDRC Data Products**

CDRC 2011 Census Data Packs

Internet User Classification 2014

CDRC Wifi Hotspots

CDRC Council Tax Bands CDRC 2011 Residence-Workplace Geodata Pack

CDRC Individual Income Estimates (PAYE) CDRC 2015 OS Geodata Pack CDRC 2011 OAC Geodata Pack

Outstanding Residential Mortgage Lending by Postcode Lending by Postcode Sector

> CDRC Access to Healthy Assets & Hazards Index (AHAH)

CDRC Price Paid Data per Property Type at the LSOA scale

CDRC Fixed Broadband and Network Infrastructure

CDRC 2015 Energy Estimate Geodata Pack

Change in the Rateable Value of Retail Premises

CDRC Estimated

London Workplace Zone Classification (LWZC)

Open Data

# **CDRC Training**

- 15<sup>th</sup> Jan: Introduction to QGIS: Understanding and Presenting Spatial Data - University of Southampton (co-badged ADRC-E)
- 17<sup>th</sup> Jan: Confident Spatial Analysis University of Southampton (cobadged ADRC-E)
- 22<sup>nd</sup> Feb: Tableau Workshop: University of Leeds
- 19<sup>th</sup> Mar: Introduction to ArcGIS: University of Leeds
- 23<sup>rd</sup> Apr: Introduction to spatial data and using R as a GIS -University of Liverpool in London (co-badged ADRC-E)
- 24<sup>th</sup> Apr: Confident Spatial Analysis and Statistics in R & GeoDa University of Liverpool in London (co-badged ADRC-E)
- Date TBC: Introduction to Data Linkage Said Business School (cobadged ADRC-E)
- Date TBC: Advanced Data Linkage Said Business School (cobadged ADRC-E)
- 13-15<sup>th</sup> June: Retail Location and Customer Analysis Said Business School
- 2 days Early summer: Machine Learning and Social Science University of Liverpool, University of Liverpool in London/UCL



# Big Data



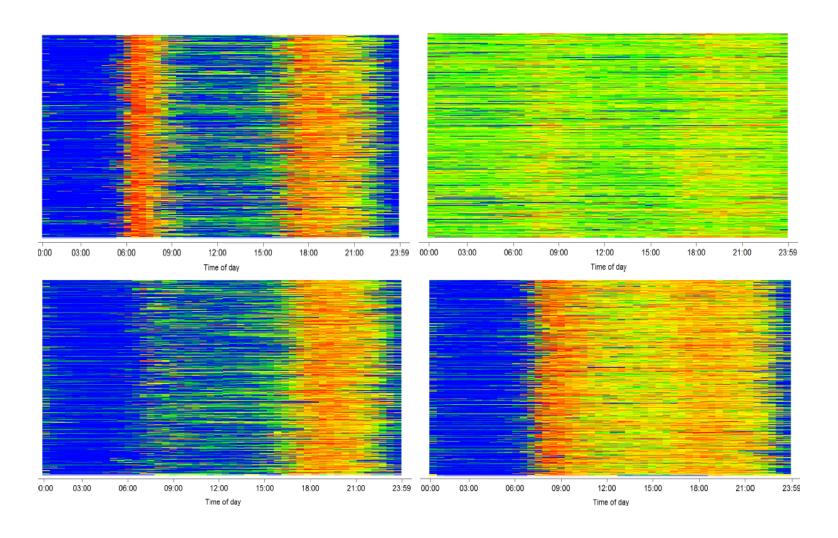
# An increasing share of data on people are being collected by commercial organisations



# **Exhaust**

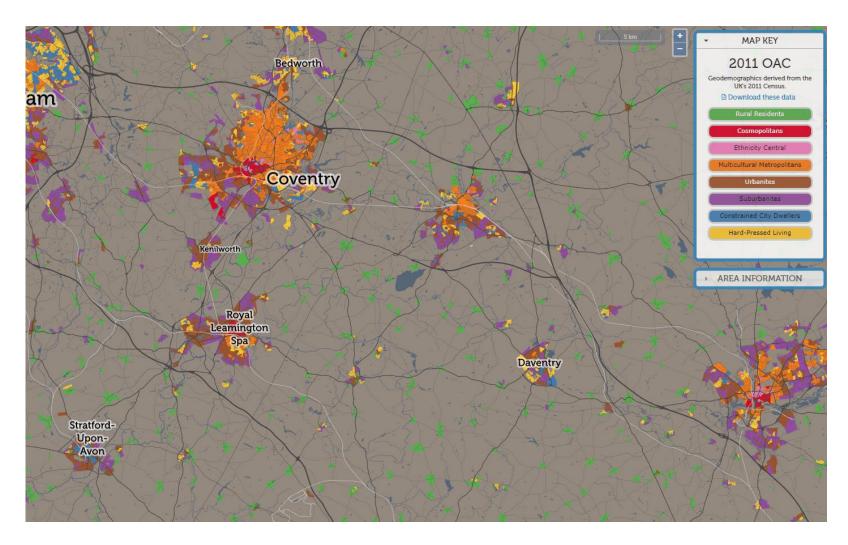


# **Smart Meter Data**

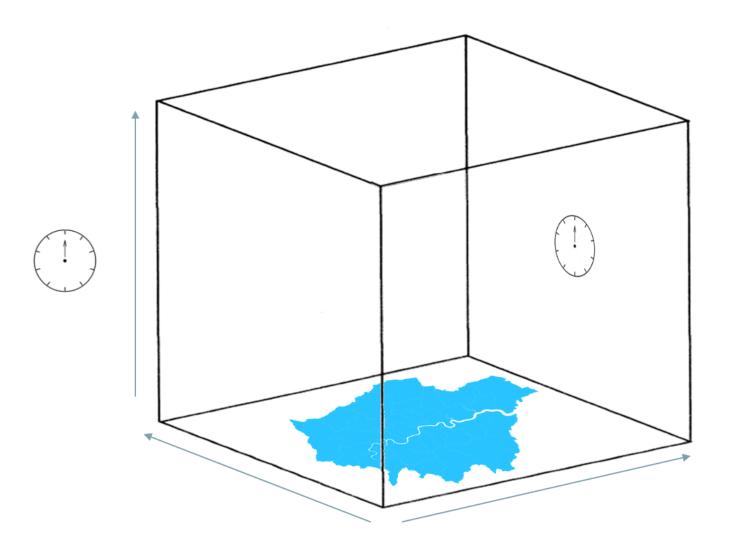


Samson, Lansley and Simpson (2014) Can smart meters save consumers and British Gas money and carbon by pinpointing which consumers are most likely and best placed to install insulation in their homes?

# Geodemographics



# **Time and Space**

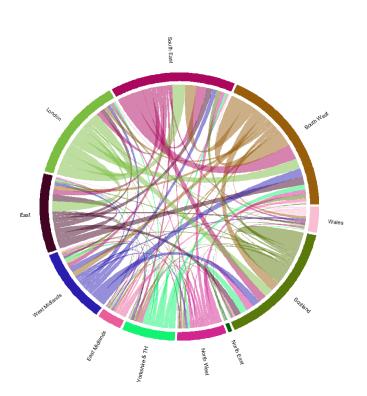




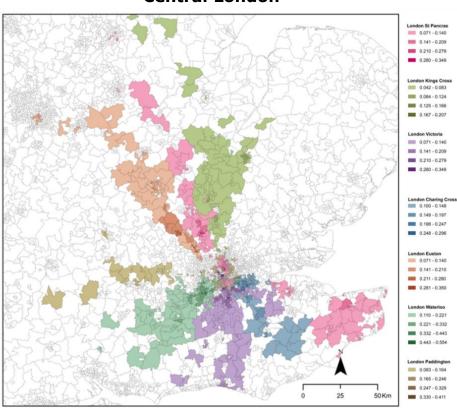
# **High Street Retailer**

# **Estimating Relocations**

Origin region > New store region



# Stores located near transport hubs, Central London



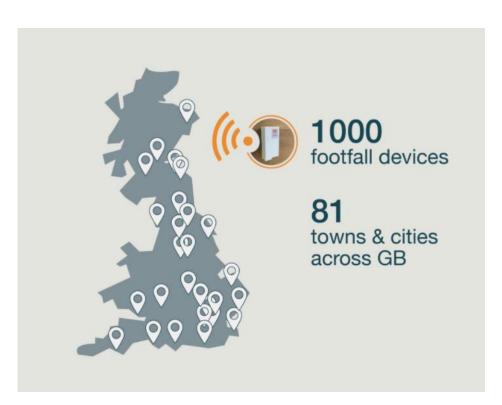
**Method Applications** 

# **High Street Retailer**

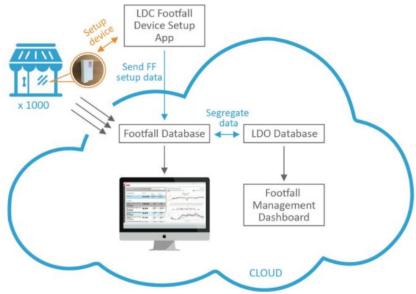




# **SmartStreetSensors**



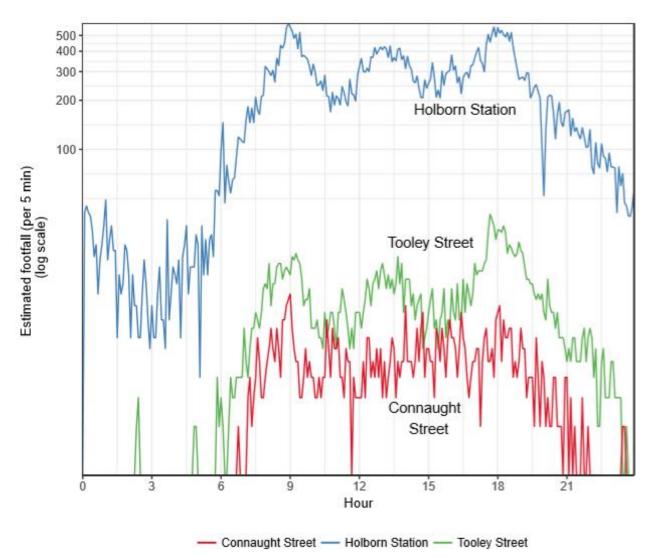
# How does it work?



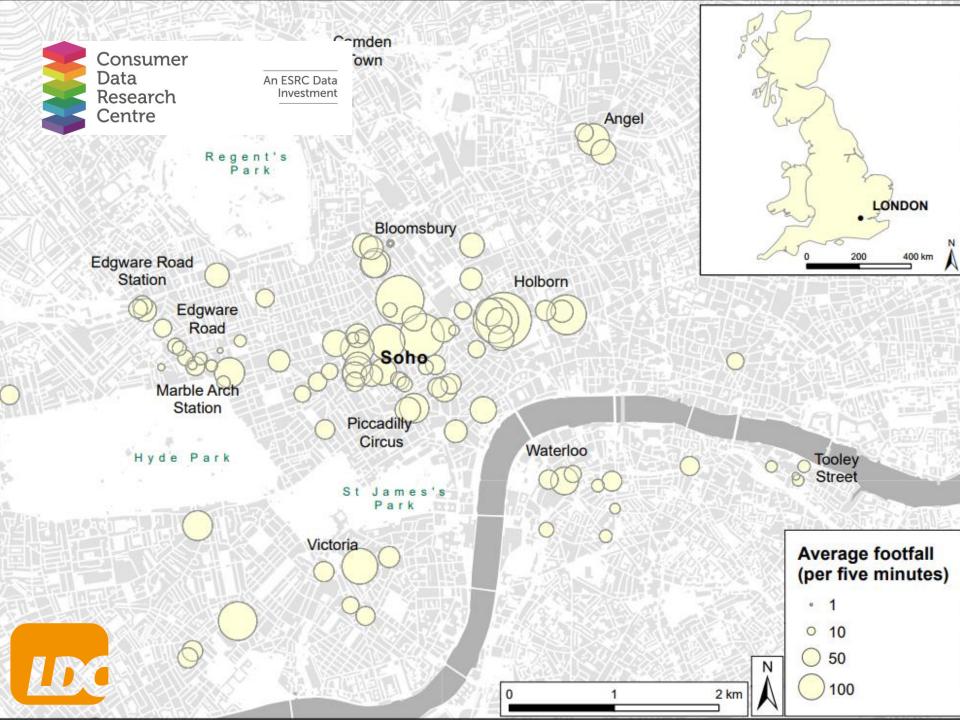




# **SmartStreetSensors**

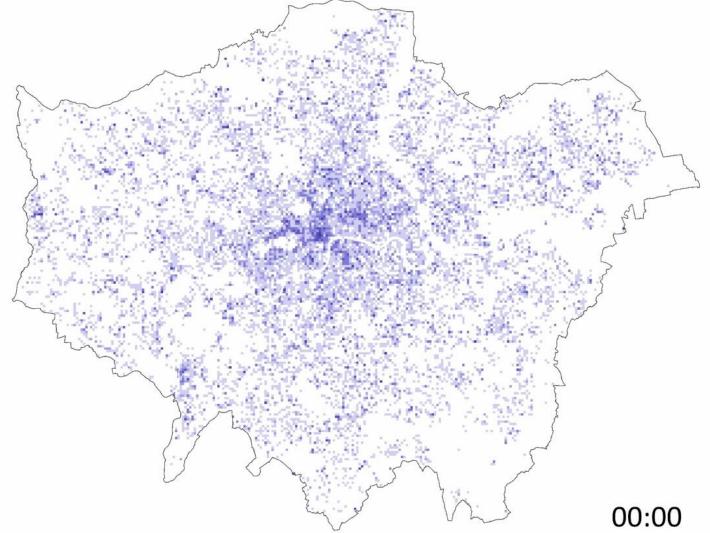






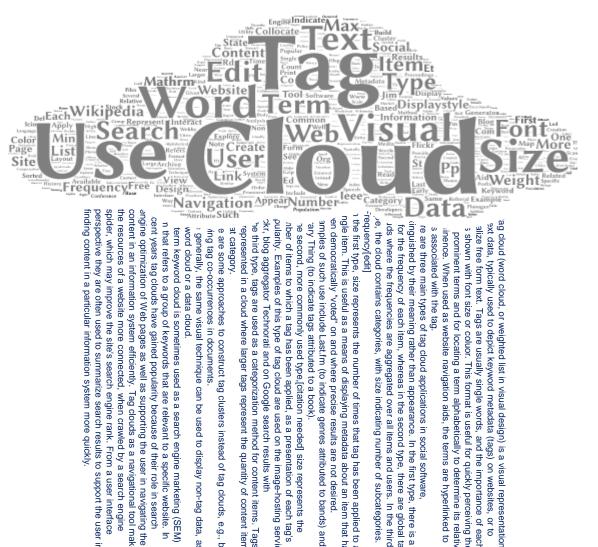


# **Geotemporal Demographics**



Longley, Lansley and Adnan (2015) The geotemporal demographics of Twitter usage. *Environment and Planning A*, 47 (2), 465-484

# **Text Mining**



ngle item. This is useful as a means of displaying metadata about an item that has the first type, size represents the number of times that tag has been applied to a Je, the cloud contains categories, with size indicating number of subcategories. en democratically "voted" on and where precise results are not desired. ckr, blog aggregator Technorati and on Google search results with pularity. Examples of this type of tag cloud are amples of such use include Last.fm (to indicate genres attributed to inguished by their meaning rather than appearance. In the first type, there is a for the frequency of each item, whereas in the second type, there are global tag uds where the frequencies are aggregated over all items and users. In the third nber of items to which a tag has been applied, as a presentation of each tag's ary Thing (to indicate tags attributed to a book) he third type, tags are used as a categorization method for content items. Tags re are three main types of tag cloud applications in social software represented in a cloud where larger tags represent the quantity of content items ne second, more commonly used type,[citation needed] size represents the used on the image-hosting service

s shown with font size or coluor. This format is useful for quickly perceiving the

Tags are usually single words, and the importance of each

inence. When used as website navigation aids, the terms

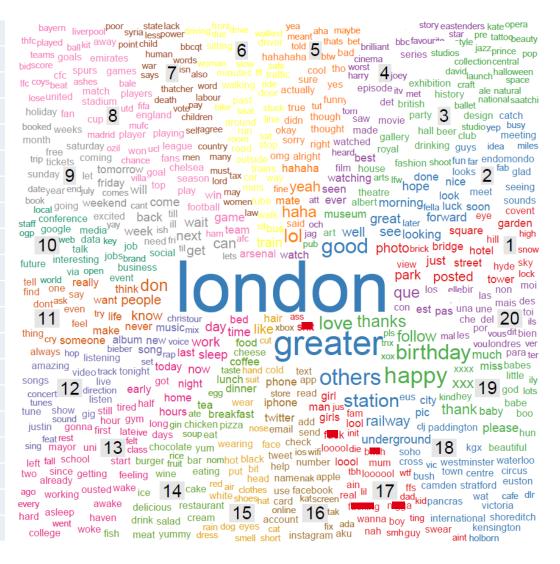
term keyword cloud is sometimes used as a search engine marketing (SEM) word cloud or a data cloud. generally, the same visual technique can be used to display non-tag data, as ing tag co-occurrences in documents e are some approaches to construct tag clusters instead of tag clouds, e.g., by

spider, which may improve the site's search engine rank. From a user interface the resources of a website more connected, when crawled by a search engine engine optimization of Web pages as well as supporting the user in navigating the n that refers to a group of keywords that are relevant to a specific website. In cent years tag clouds have gained popularity because of their role in search



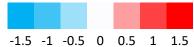
# **Twitter Topics**

1	Photography and Sights
2	Optimism, Kindness and Positivity
3	Leisure and Attractions
4	TV and Film
5	Humour and Informal Conversations
6	Transport and Travel
7	Politics, Beliefs and Current Affairs
8	Sport and Games
9	Anticipation and Socialising
10	Business, Information and Networking
11	Pessimism and Negativity
12	Music and Musicians
13	Routine Activities
14	Food and Drink
15	Body, Appearances and Clothes
16	Social Media and Apps
17	Slang and Profanities
18	Place and Check-Ins
19	Wishes and Gratitude
20	Foreign and Other





# **Twitter Topics**



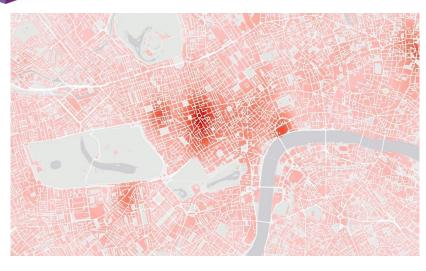
Photography and Sights Optimism, Kindness and Positivity Leisure and Attractions TV and Film **Humour and Informal Conversations Transport and Travel** Politics, Beliefs and Current Affairs **Sport and Games Anticipation and Socialising** Business, Information and Networking **Pessimism and Negativity** Music and Musicians **Routine Activities** Food and Drink Body, Appearances and Clothes Social Media and Apps Slang and Profanities Place and Check-Ins Wishes and Gratitude Foreign and Other

Hour 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

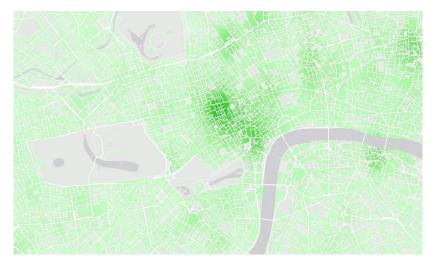
**All Tweets** 

Lansley and Longley (2016) The geography of Twitter topics in London. *Computers, Environment and Urban Systems*, 58, 85-96

# **Twitter and Space**



Fashion and Shopping



Nightlife

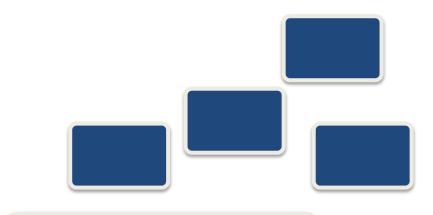


Museums and Galleries



**Shows and Entertainments** 

# **Consumer Registers**



Public Version of the Electoral Register

+

Consumer data from several sources

2003 - 2017

1998 - 2017

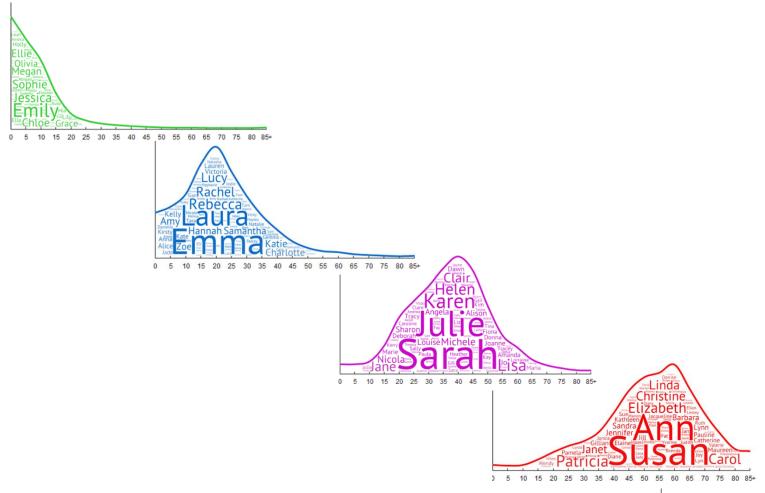


# **Identity**



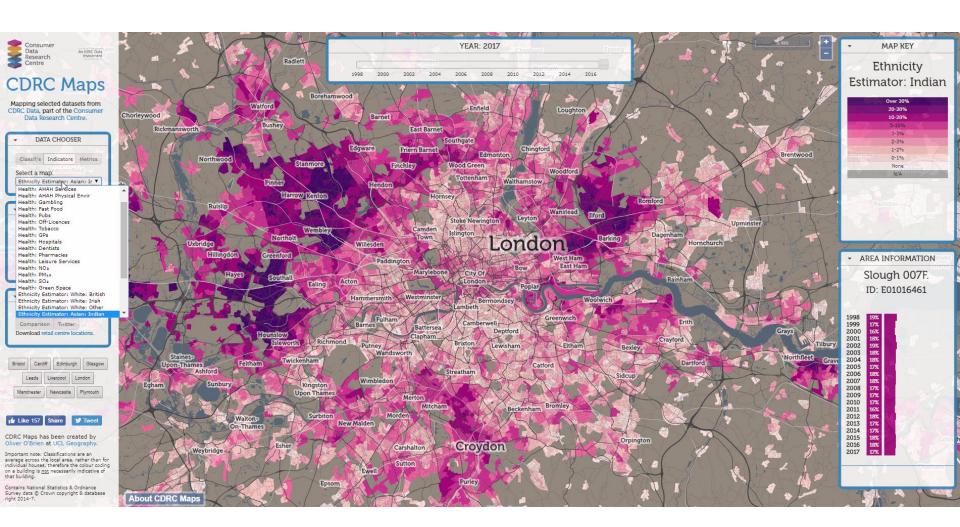


# Age



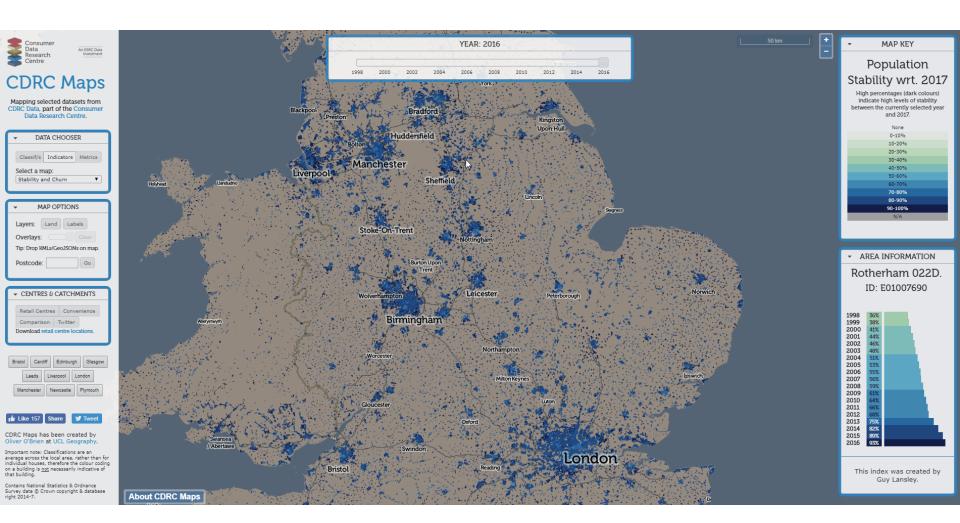


# **Ethnicity**



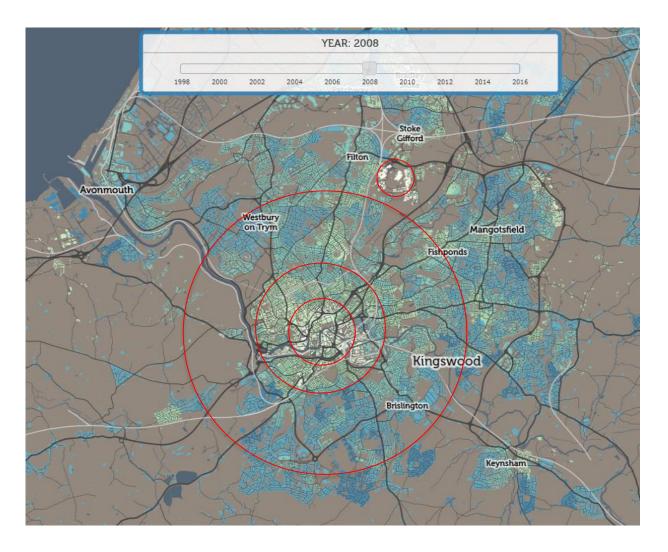


# **Population Churn**



CDRC Maps Beta https://maps.cdrc.ac.uk/index\_beta.php#/indicators/

# **Population Churn**

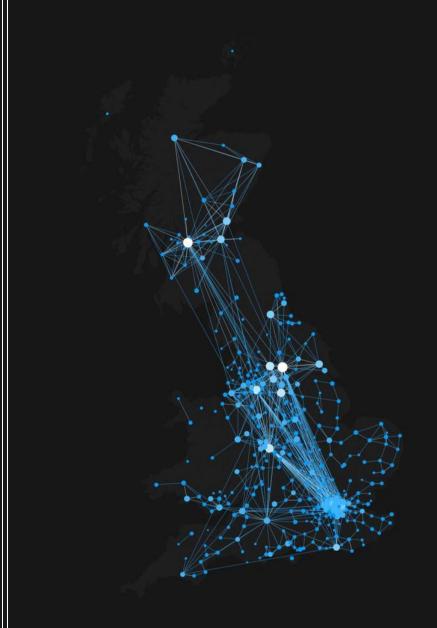




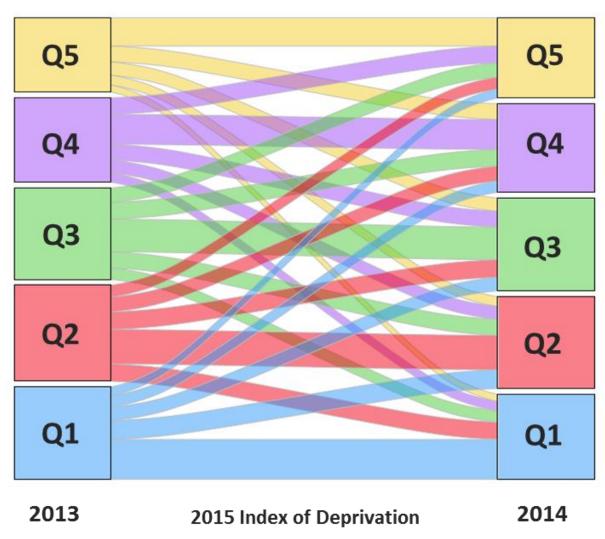


- Flows between large urban districts
- Flows between neighbouring districts
- Median move of 20.97 miles

# **Estimating Migration**

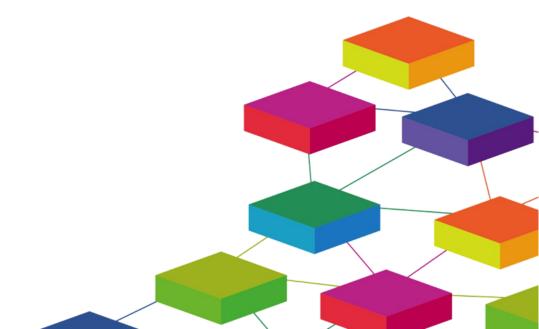


# **Social Mobility**



Lansley and Li (2018) Consumer registers as spatial data infrastructure and their use in migration and residential mobility research. *Forthcoming...* 

# The CDRC Masters Research Dissertation Programme



# **Timeline**

# **January**

Projects are advertised on the CDRC website

# Jan - Apr

The student application process is open. The CDRC will forward selected CVs and cover letters.

# May - Aug

Research is undertaken either at the students own institution or on-site with the industry sponsor

# **September**

Dissertations are submitted to the CDRC and industrial sponsors

# October

Participants are invited to an Academic conference where the prizes are awarded



# **Sponsors (2012-17)**









The **co-operative** 

John Lewis





WHITBREAD





**British Gas** 





SHOP

DIRECT





































# **Students (2012-17)**

























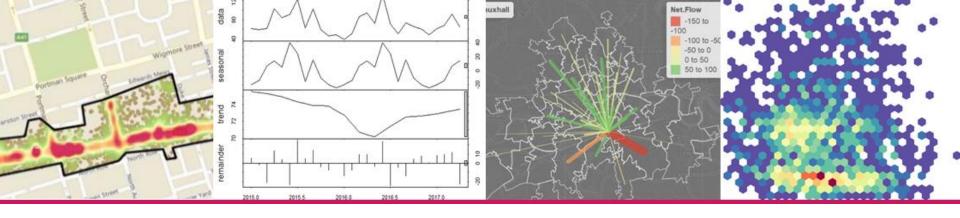


UNIVERSITY OF LONDON











The programme offers students a unique opportunity to apply their skills to real-world problems and to gain valuable experience working with commercial datasets



# **Outputs**













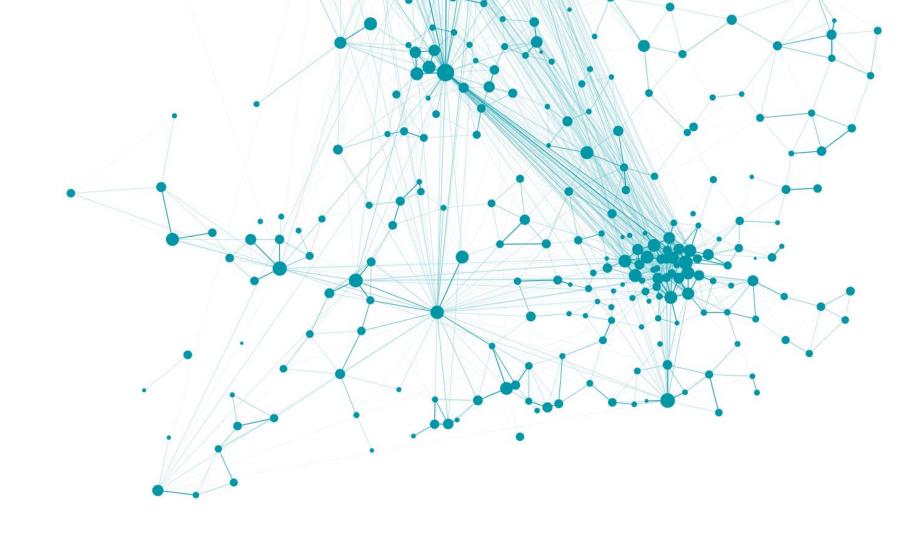




CACI



Short examples of previous projects are available online www.cdrc.ac.uk/retail-masters/archive



Guy Lansley g.lansley@ucl.ac.uk @GuyLansley

www.cdrc.ac.uk/retail-masters