

# Quantitative Geography and the SmartCensus

*Paul Longley  
Professor of Geographic Information  
Science, UCL*





Consumer  
Data  
Research  
Centre

An ESRC Data  
Investment





Consumer  
Data  
Research  
Centre

An ESRC Data  
Investment



# Free Movement of Data (1): Private/public sectors

- Big Data
  - Social survey response rates
  - Real share of *consumer* data:
    - forms
    - Time and space: 24 hour cities
  - Public versus private sector
    - The value of open platforms
  - Exhaust; metadata and provenance?
    - Linkage
    - Shared research challenges

In 1992 all the pollsters went wrong doing the same thing. [In 2015] they went wrong doing different things. Some firms contact people via telephone, others online, and they ask different questions [or use different] statistical methods.

Joe Twyman, YouGov





Consumer  
Data  
Research  
Centre

An ESRC Data  
Investment

YOUR M&S

Sainsbury's

Boots

Demographics  
USER GROUP  
Making Data Work

The co-operative

WHITBREAD

DIXONS RETAIL  
BRINGING LIFE TO TECHNOLOGY

SHOP DIRECT

goHenry

BARCLAYS

e.on

Camelot

CACI

movement  
strategies

Thomas Cook

Argos

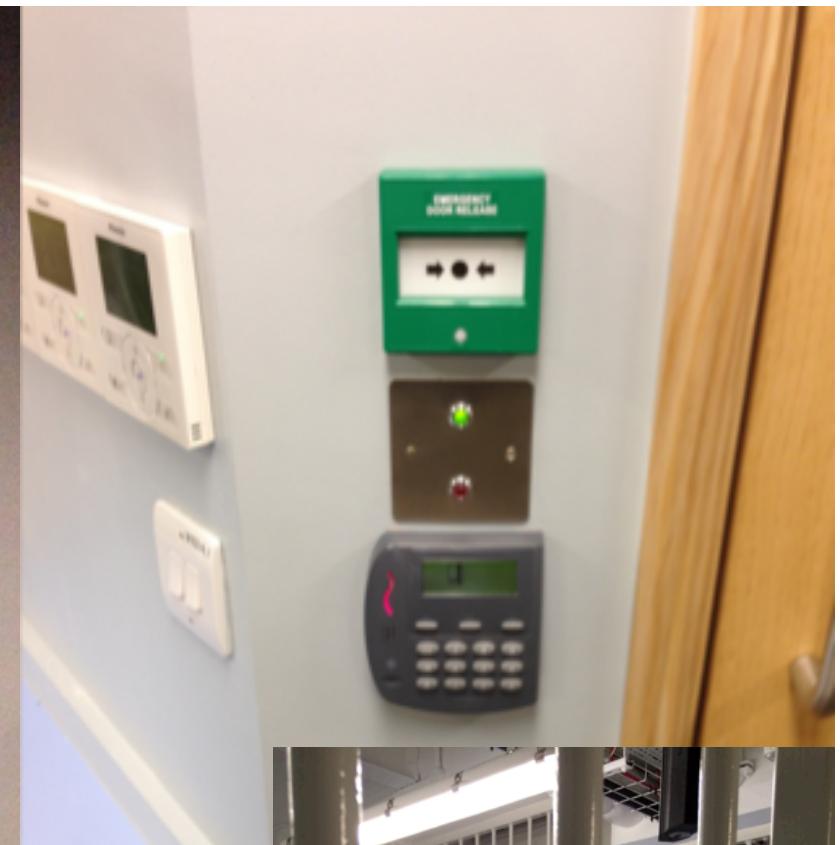
LDC

gsk  
GlaxoSmithKline

Experian™

easyJet

AXA  
PPP HEALTHCARE



Co-Production/Analytics



Data Provider

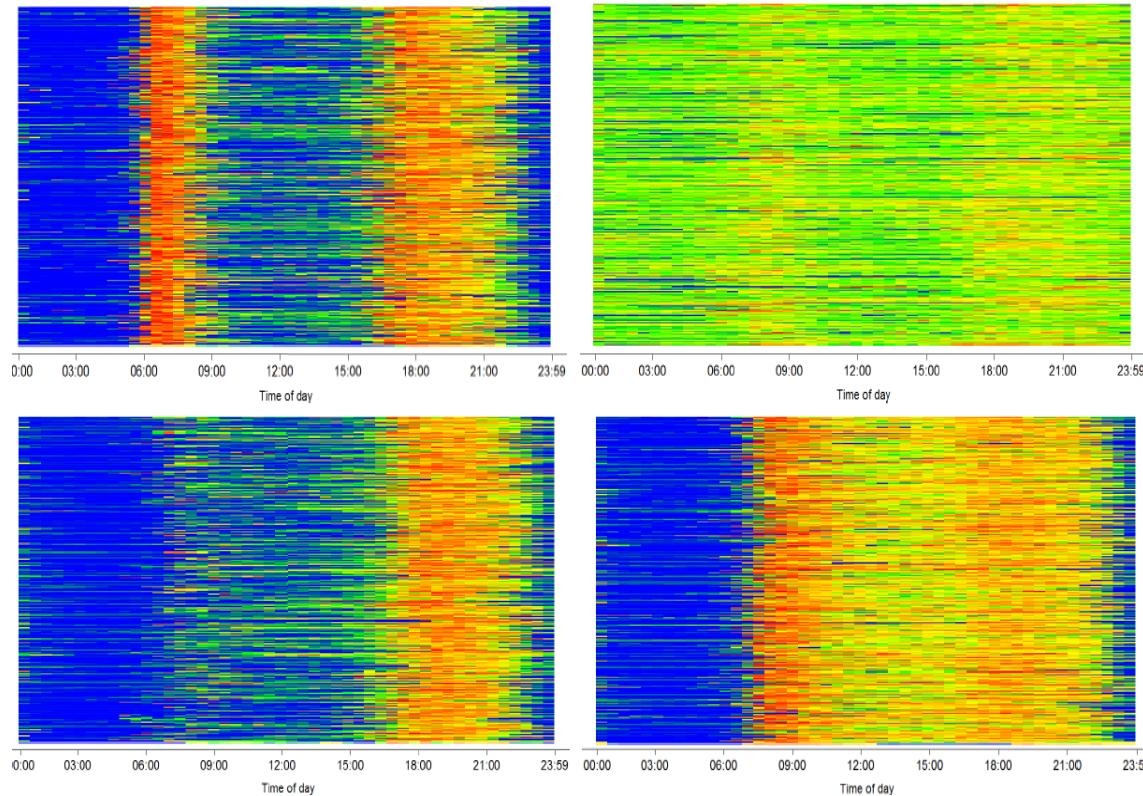
Data Science Projects

PhD Projects

MSc Projects



# Domestic Energy Providers



## Timeline

- 2014 – CDRC Dissertation student project
- 2015 – Second dissertation in the same research field
- 2015 – Co-funding of a PhD with the CDRC
- 2015 – nationwide data for the CDRC secure data service



- CDRC Dissertation project (Nicholas Samson, 2014)
  - Grouped smart meters into 4 key time profiles
  - Identified key demographic characteristics of each profile

Co-Production/Analytics

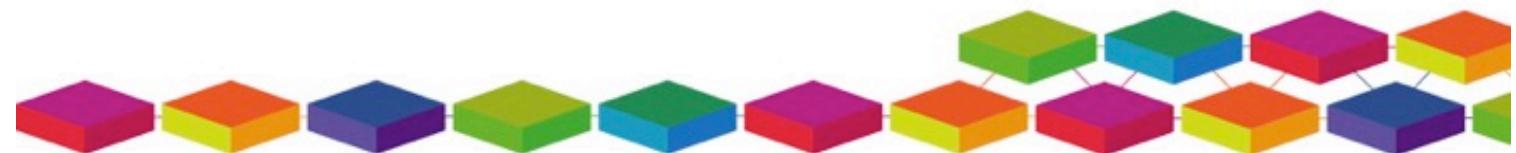


Data Provider

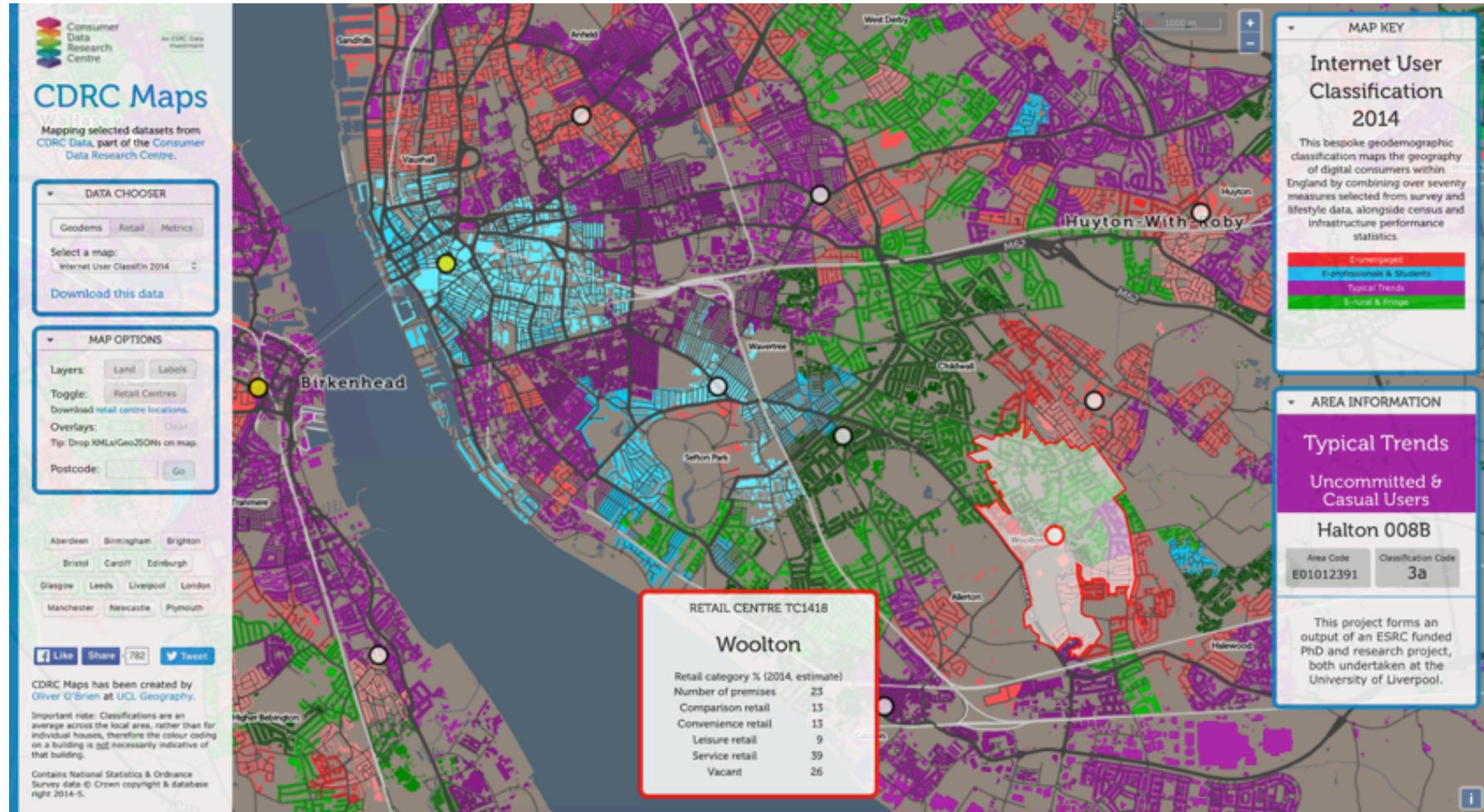
Data Science Projects

PhD Projects

MSc Projects



# Internet User Classification



Co-Production/Analytics

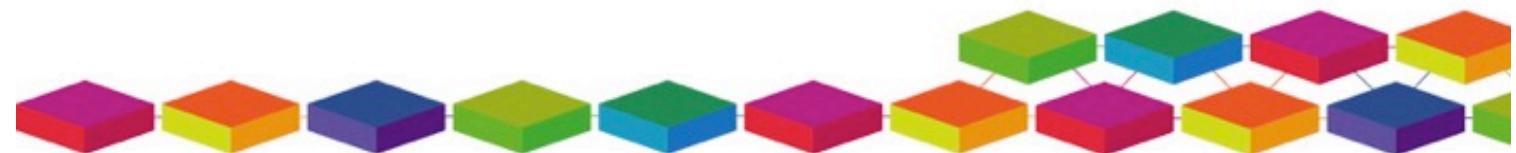


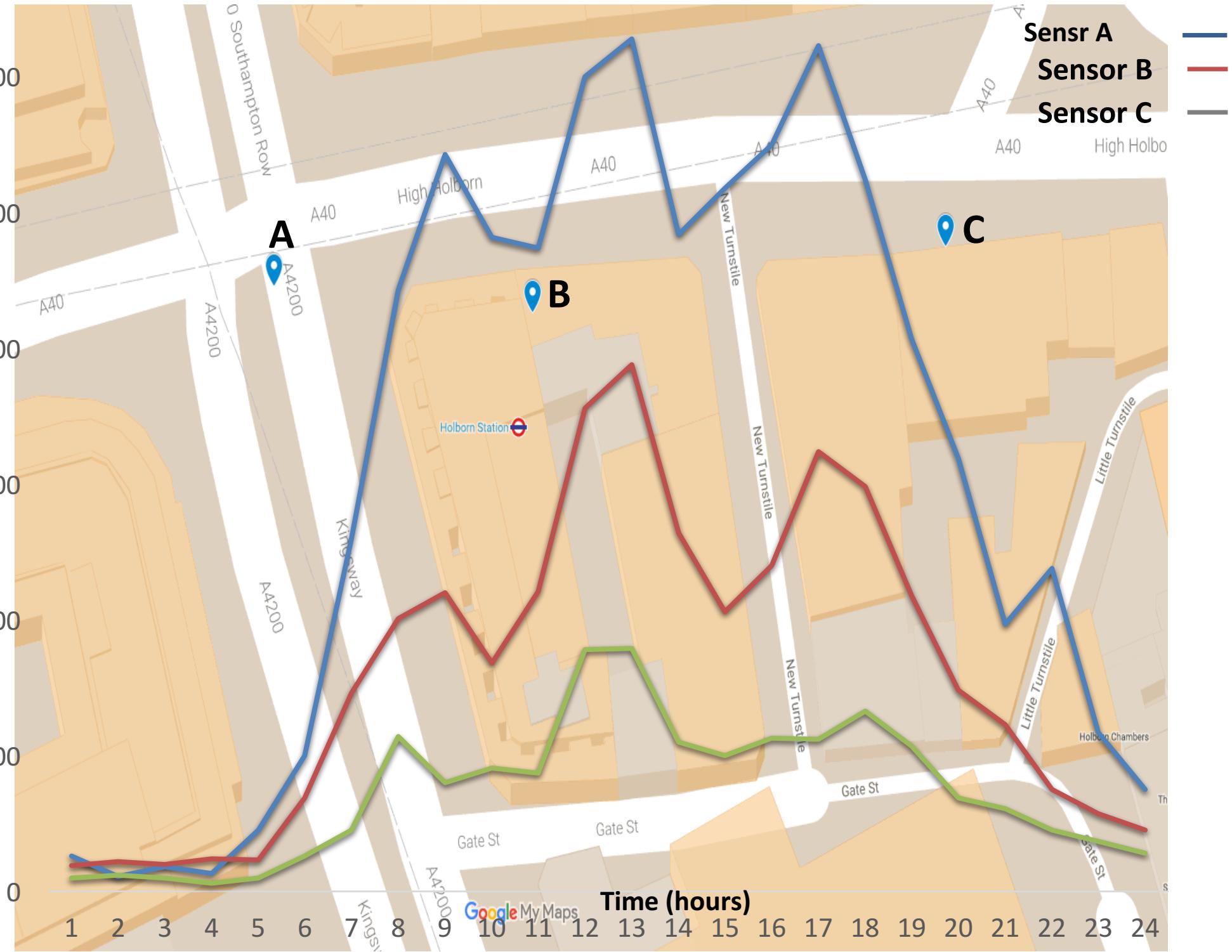
Data Provider

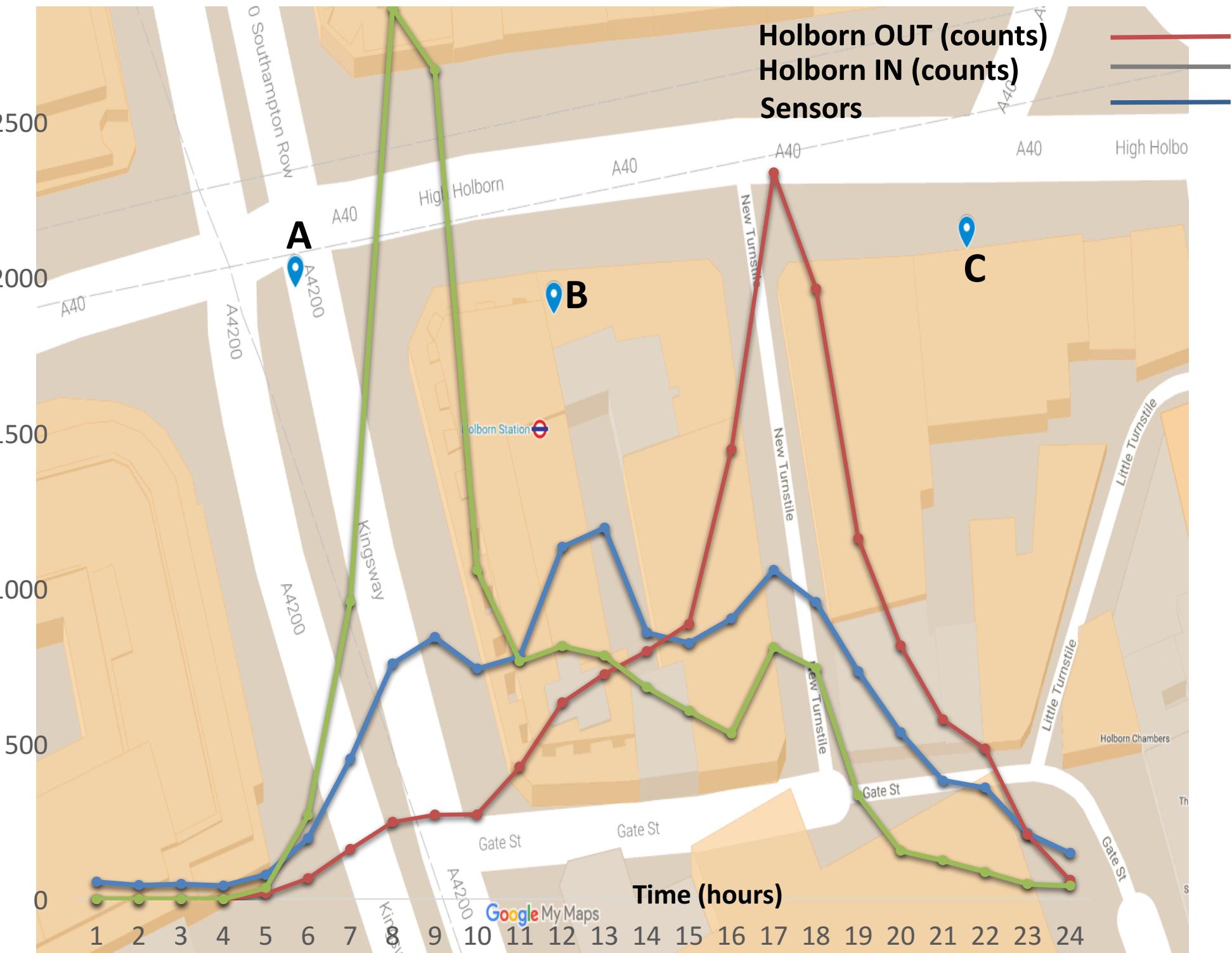
Data Science Projects

PhD Projects

MSc Projects







# Free Movement of Data (2): Open Data and Access

- Economic case for Open platforms, e.g. TfL
- Apportioning the cost of infrastructure creation and maintenance, e.g. NMOs
- Disclosure control and new modes of access, e.g. EthnicityEstimator
- Implications for the practice of science



# Free Movement of Data (3): Policy Issues

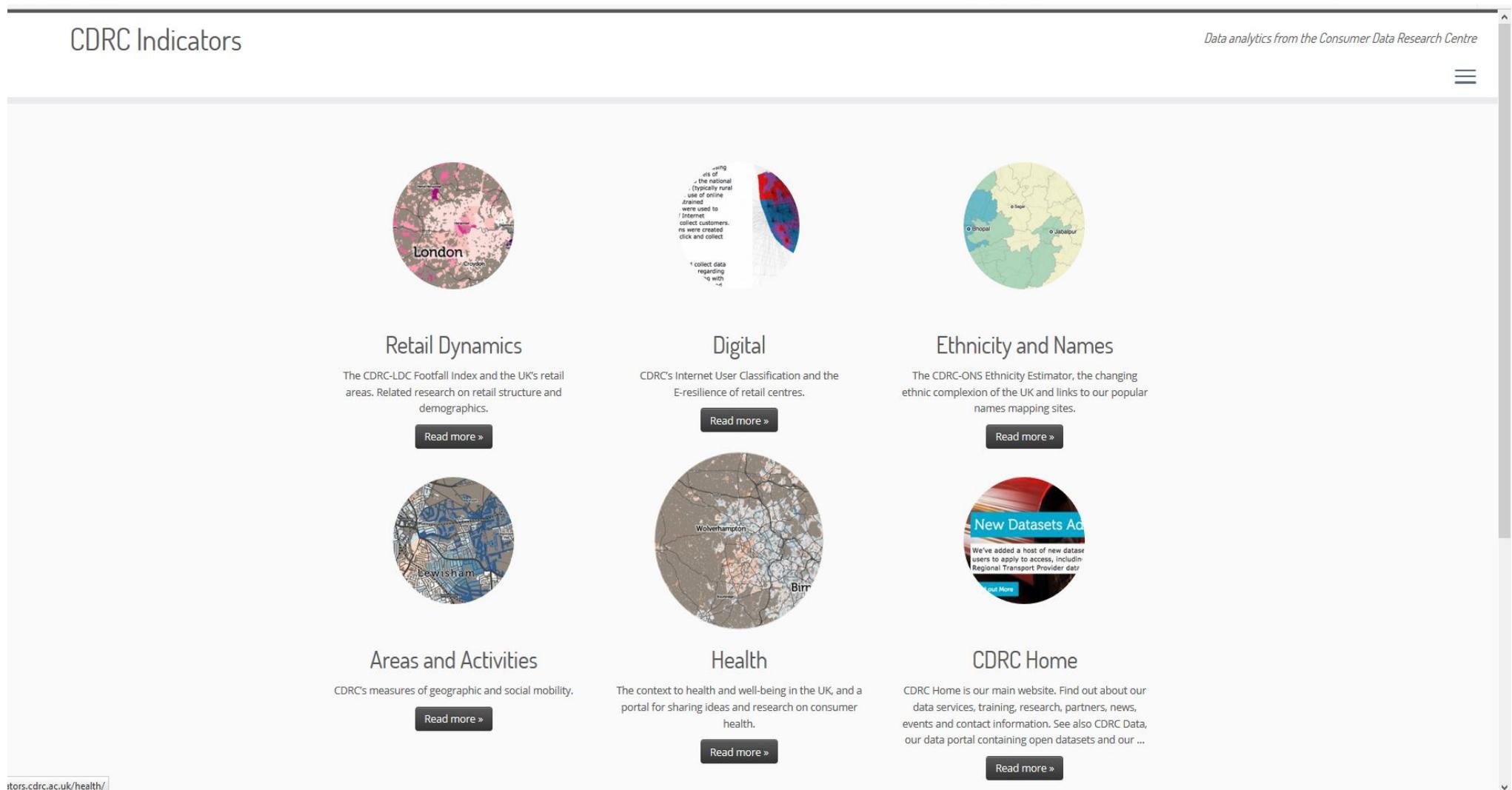
- Public and private goods: infrastructure and industrial strategy

# Next Steps

## • CDRC research indicators

CDRC Indicators

Data analytics from the Consumer Data Research Centre



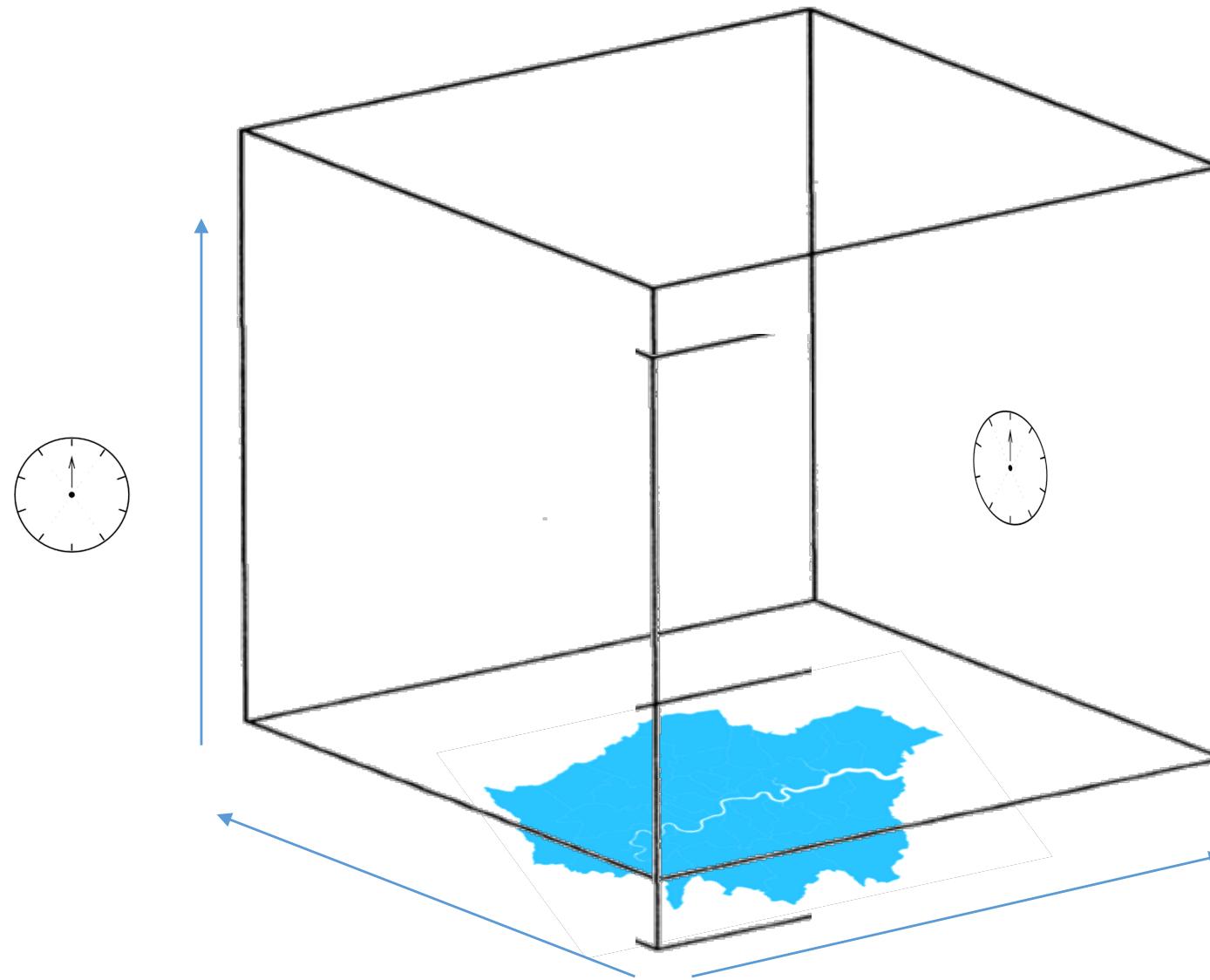
The screenshot shows a grid of nine research indicator cards:

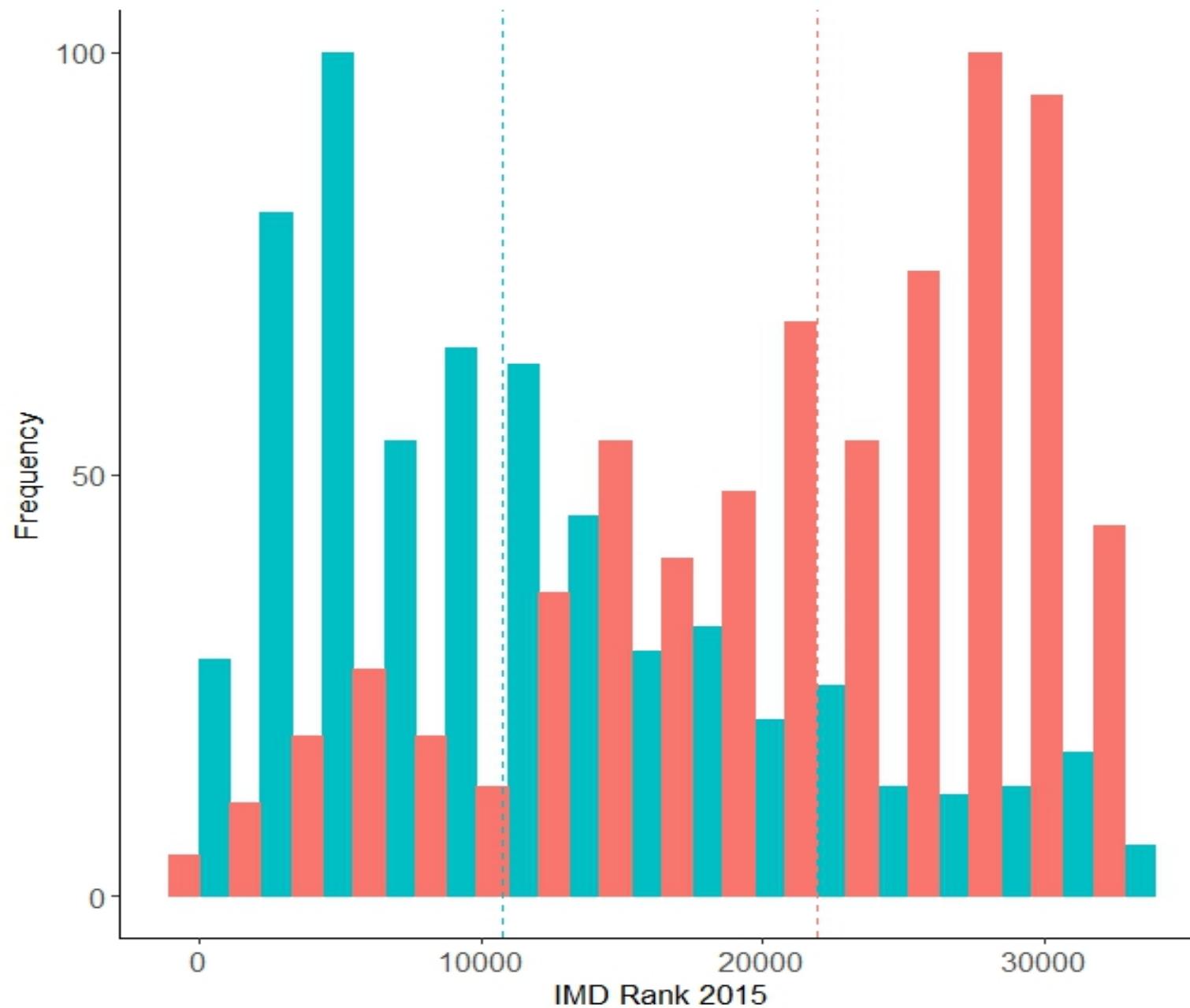
- Retail Dynamics**: The CDRC-LDC Footfall Index and the UK's retail areas. Related research on retail structure and demographics. [Read more »](#)
- Digital**: CDRC's Internet User Classification and the E-resilience of retail centres. [Read more »](#)
- Ethnicity and Names**: The CDRC-ONS Ethnicity Estimator, the changing ethnic complexion of the UK and links to our popular names mapping sites. [Read more »](#)
- Areas and Activities**: CDRC's measures of geographic and social mobility. [Read more »](#)
- Health**: The context to health and well-being in the UK, and a portal for sharing ideas and research on consumer health. [Read more »](#)
- CDRC Home**: CDRC Home is our main website. Find out about our data services, training, research, partners, news, events and contact information. See also CDRC Data, our data portal containing open datasets and our ... [Read more »](#)

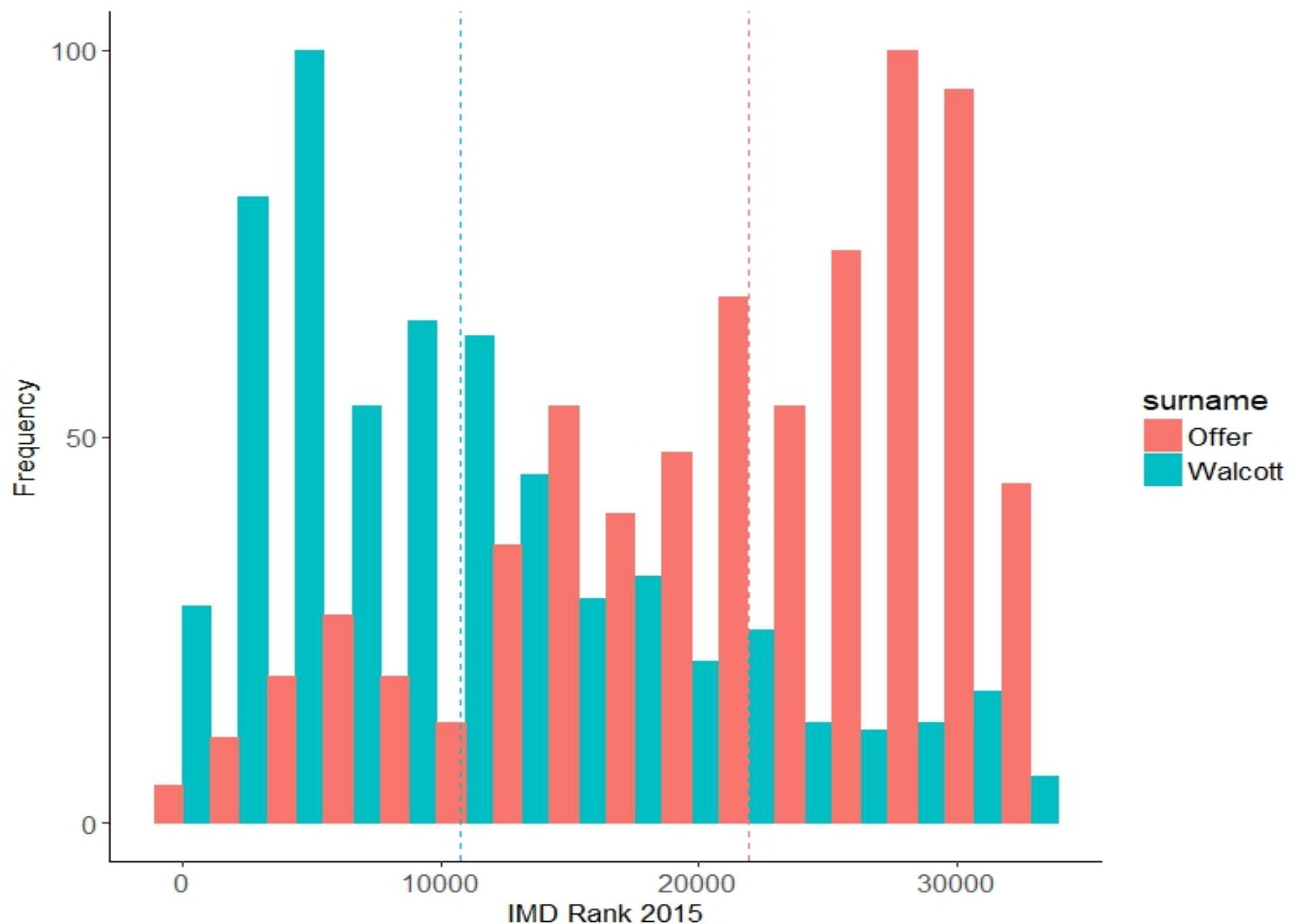
[View the full website](#)

# Free Movement of Data (3): Policy Issues

- Public and private goods: infrastructure and industrial strategy
  - Competitive economies / competition
  - No open market for data
  - New silos: A, A, A, F, M
- UK Digital Economies Act
  - Passporting for research?
- Linkage:
  - Indicators, e.g. health, social mobility, carbon footprints
  - Preservation (of what?); curation
  - Research exemptions







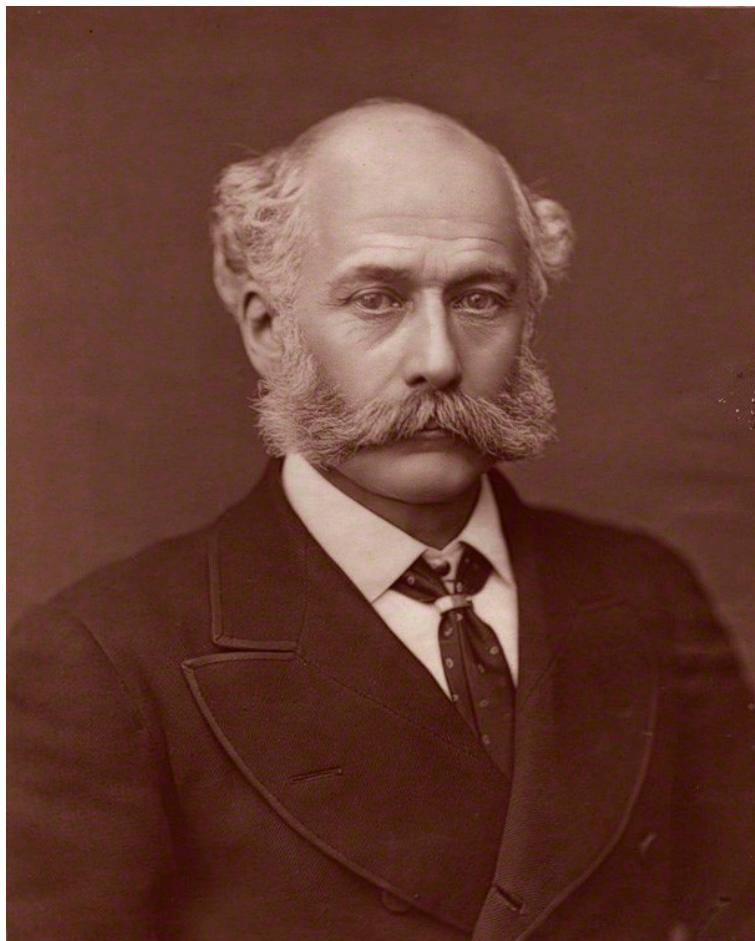
: ‘Surnames and Social Mobility’, Gregory Clark and Neil Cummins

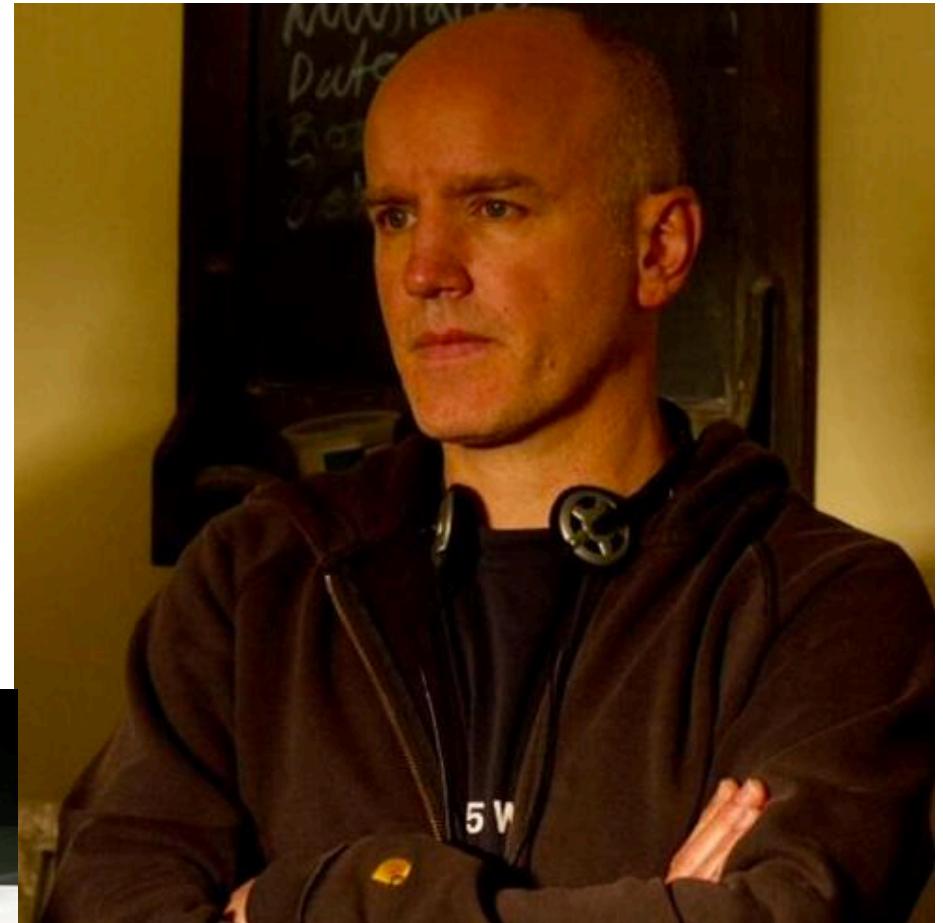
<http://www.econ.ucdavis.edu/faculty/gclark/ecn110a/readings/Surname%20Mobility%202013.pdf>

‘using educational status in England 1170-2012 as an example, ... the true underlying [intergenerational] correlation of social status is in the range 0.75-0.85. Social status is more strongly inherited even than height.’

This ‘stems from the nature of inheritance of characteristics within families. Strong forces of familial culture, social connections, and genetics must connect the generations. There really are quasi-physical “Laws of Inheritance.”’

# Bazalgette

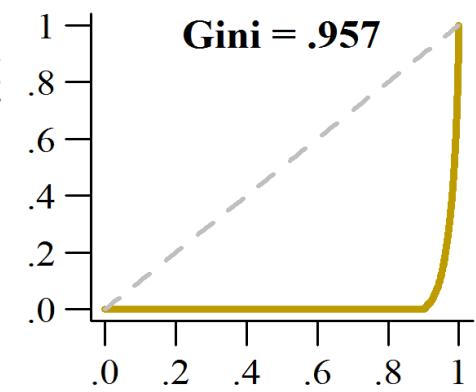
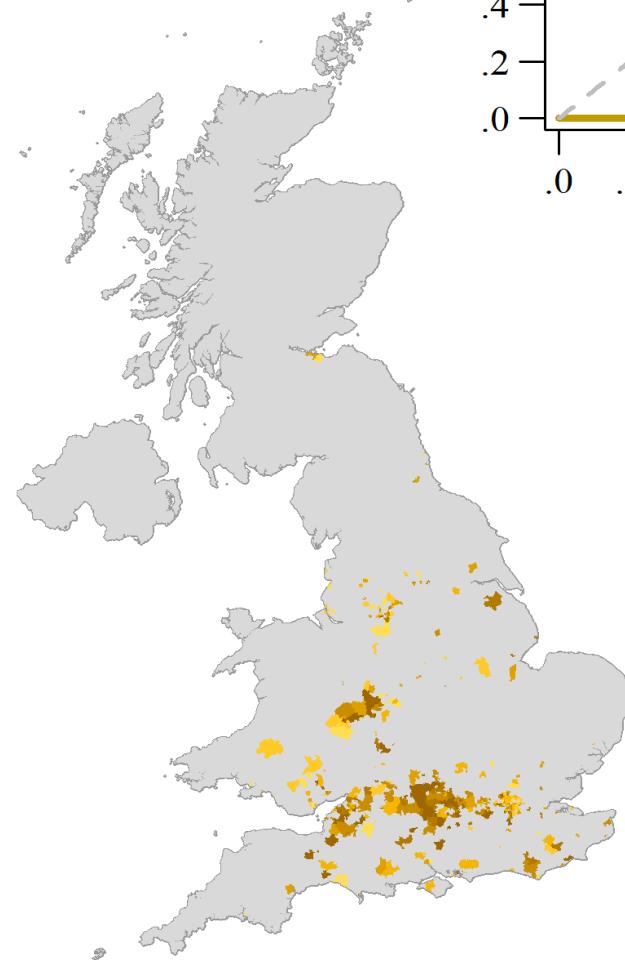
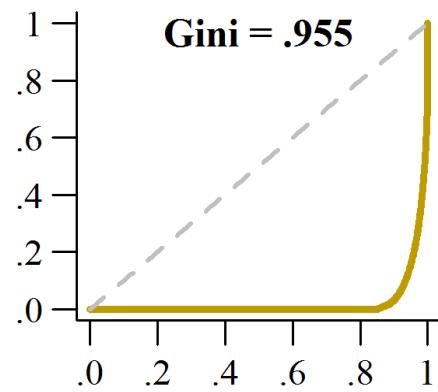
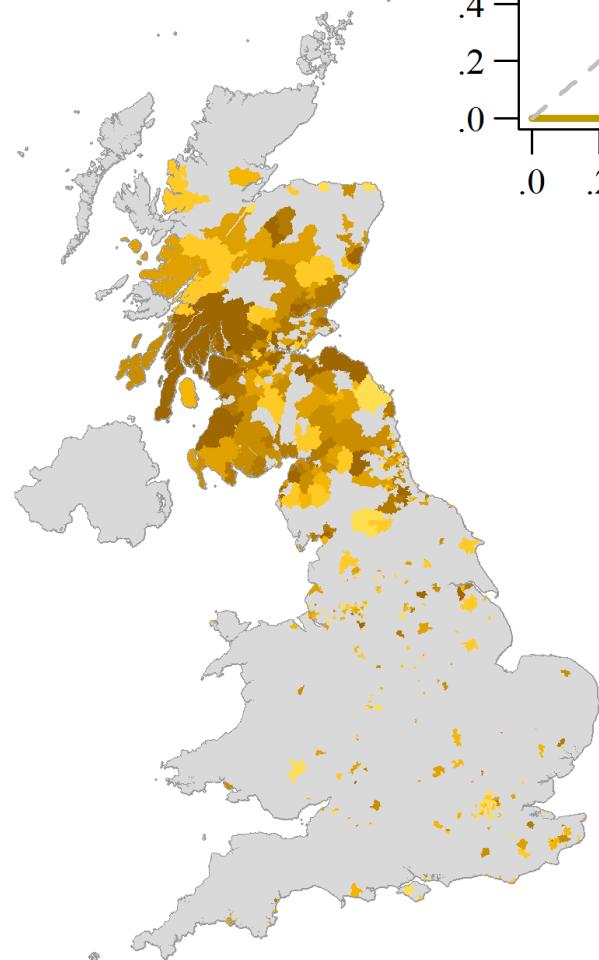




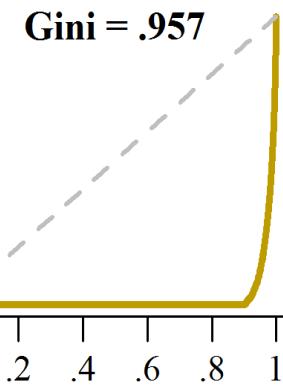
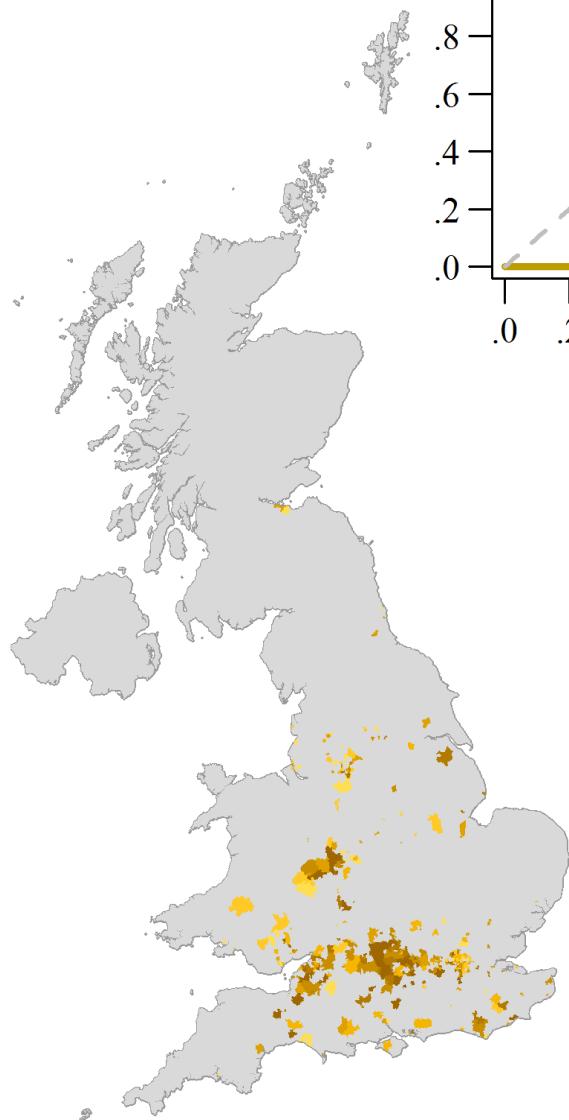
You can't plan a career in too much detail": Simon Bazalgette (Mary Turner)

## Surname: 'BAZALGETTE'

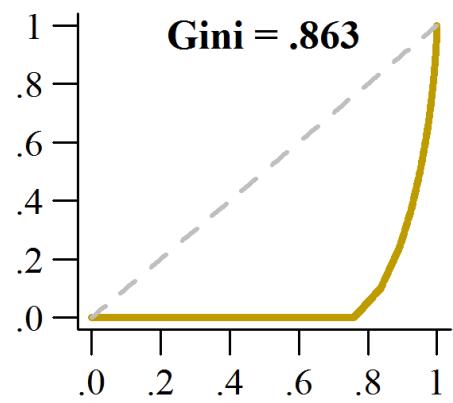
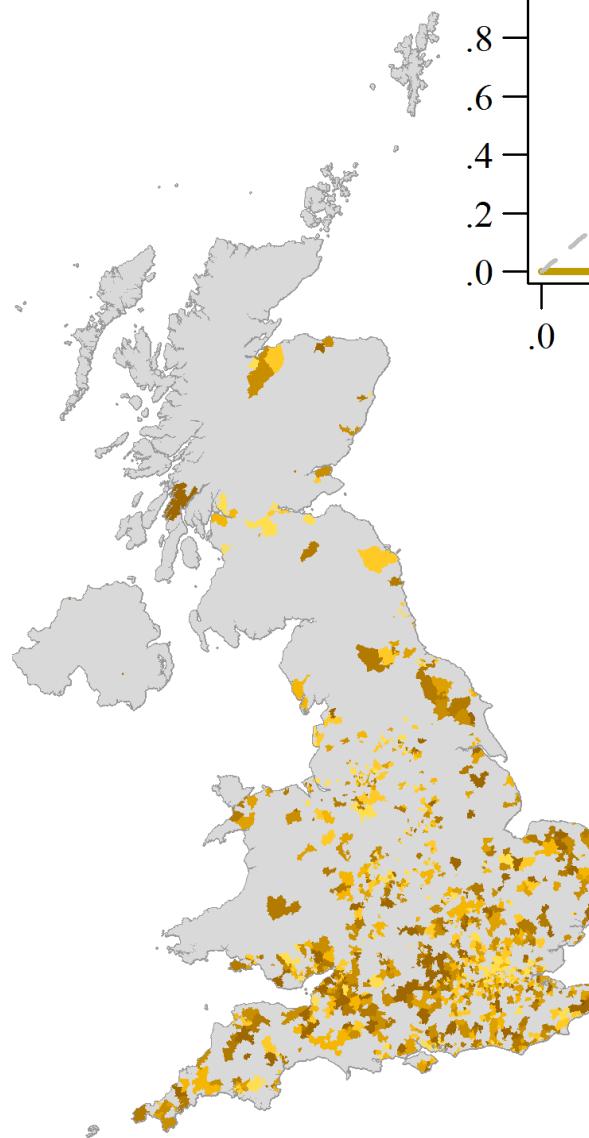
Forename	Surname	Address	Postcode
JANE	BAZALGETTE		BH1***
GUY	BAZALGETTE		BH1***
EMMA	BAZALGETTE		BH1***
ELIZABETH	BAZALGETTE		TW9***
VICTORIA	BAZALGETTE		LS1***
MARK	BAZALGETTE		DT1***
LUKE	BAZALGETTE		DT1***
MARIE	BAZALGETTE		DT1***
ELEANOR	BAZALGETTE		SW8***
LEE	BAZALGETTE		SA3***
ROBERT	BAZALGETTE		SA3***
ROBIN	BAZALGETTE		SA3***
RUTH	BAZALGETTE		PL1***
EMILY	BAZALGETTE		N15***
MARY	BAZALGETTE		BH1***
RICHARD	BAZALGETTE		BH1***

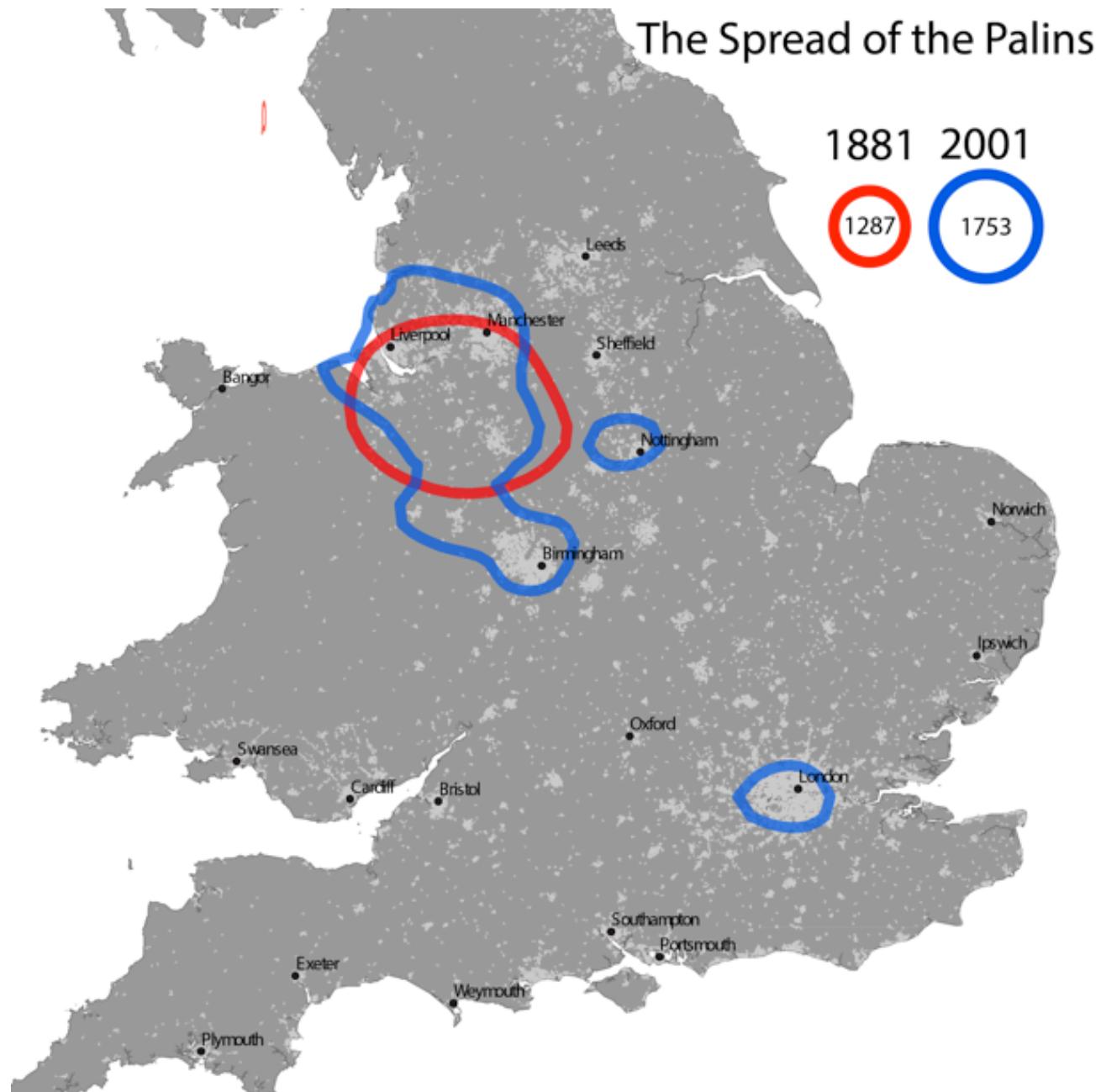


**Thatcher (1881)**

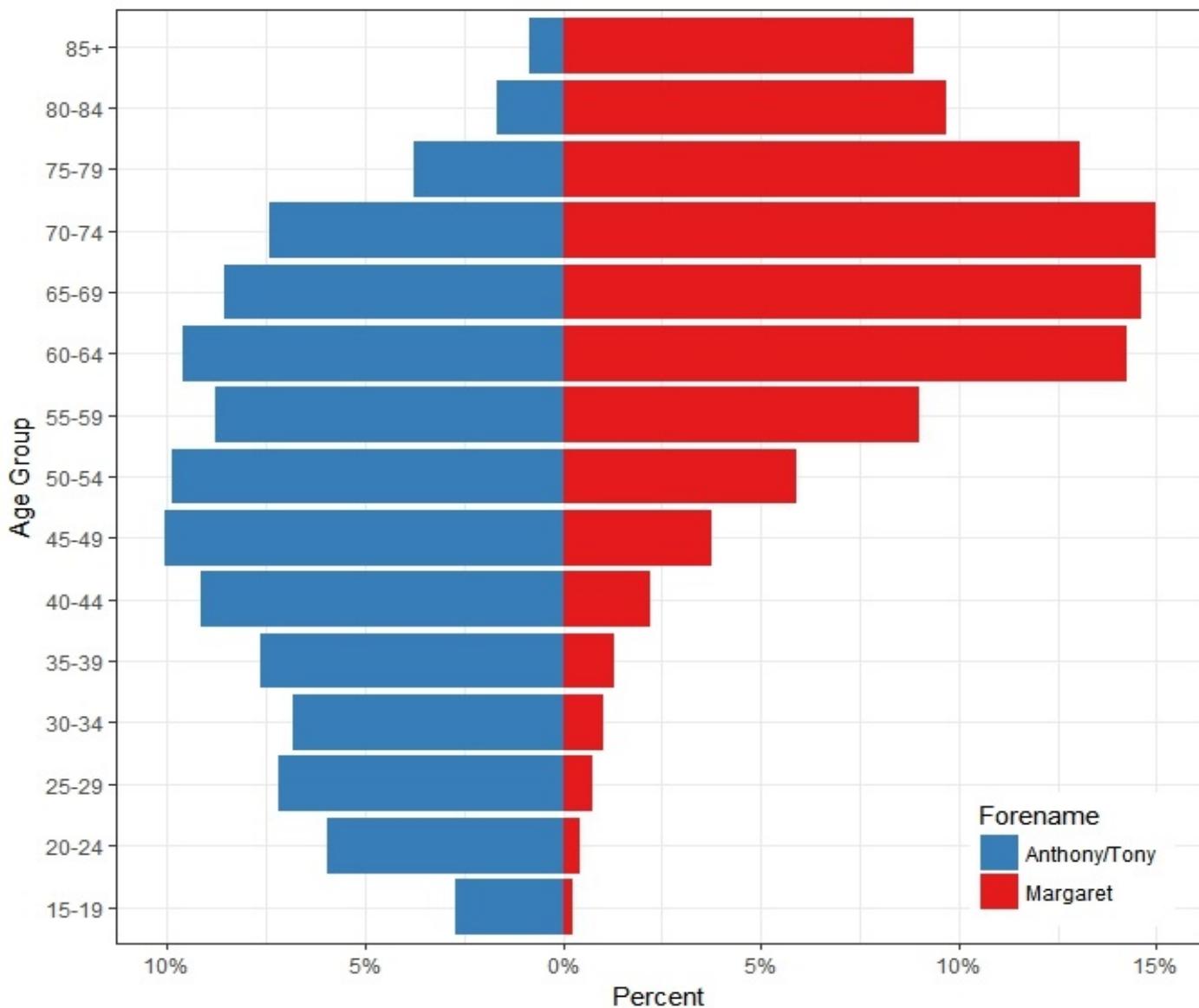


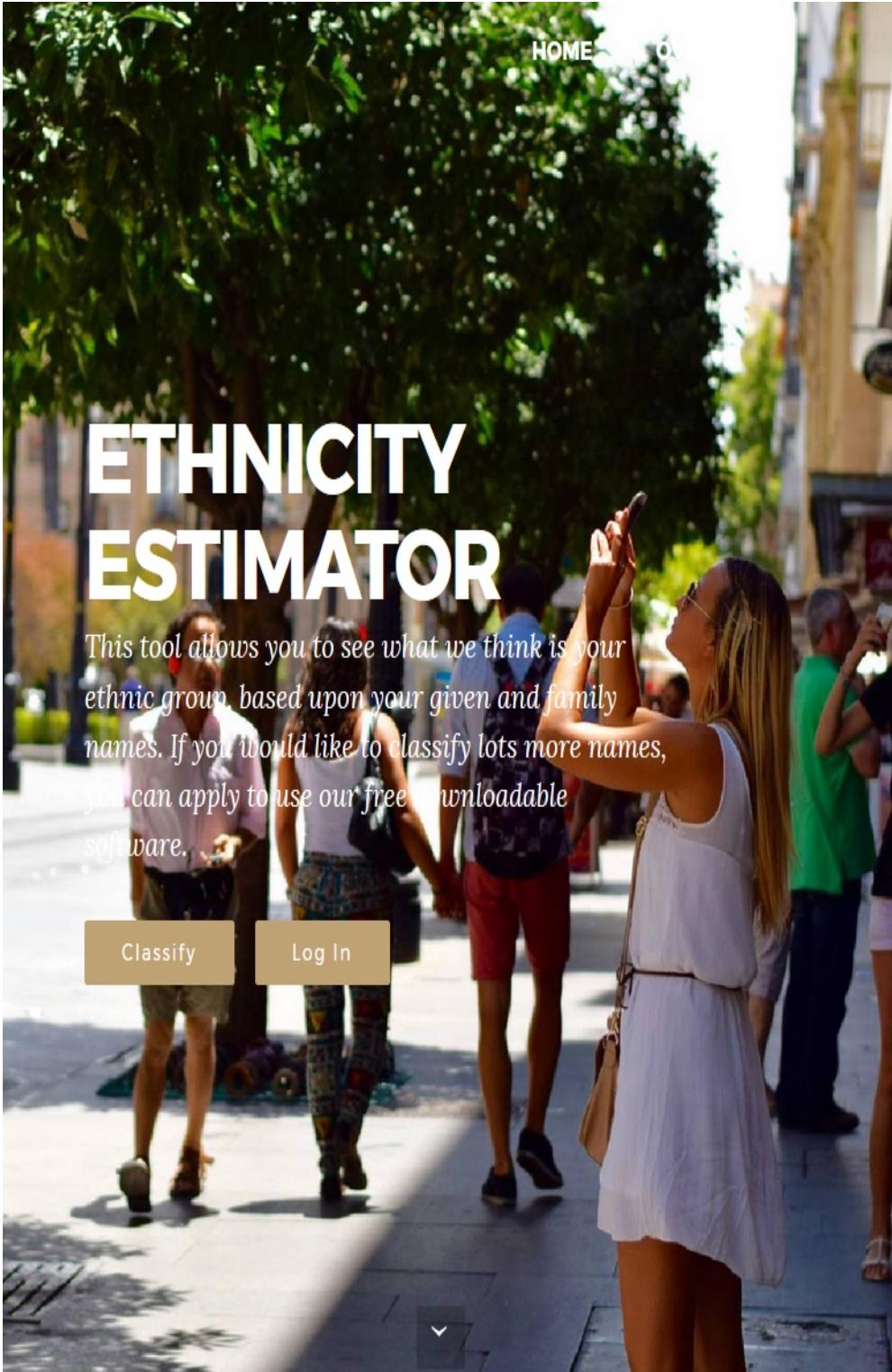
**Thatcher (2016)**





Courtesy:  
James  
Cheshire





**ETHNICITY ESTIMATOR**

This tool allows you to see what we think is your ethnic group, based upon your given and family names. If you would like to classify lots more names, you can apply to use our free downloadable software.

[Classify](#) [Log In](#)

HOME

HOME

OVERVIEW

## Classify a name

This Website is part of a project devised by the [Consumer Research Data Centre \(CDRC\)](#) in association with the [UK Office for National Statistics \(ONS\)](#). The CDRC is funded by the [Economic and Social Research Council \(ESRC\)](#).

Results from searches and user feedback are used to improve our ongoing research.

Type a Forename and a Surname in any style you want: upper, lower case; any special character. You need to fill at least one of the two fields.

Forename

Surname

Muhammad

Gill

Age (optional)

0

Gender (optional)

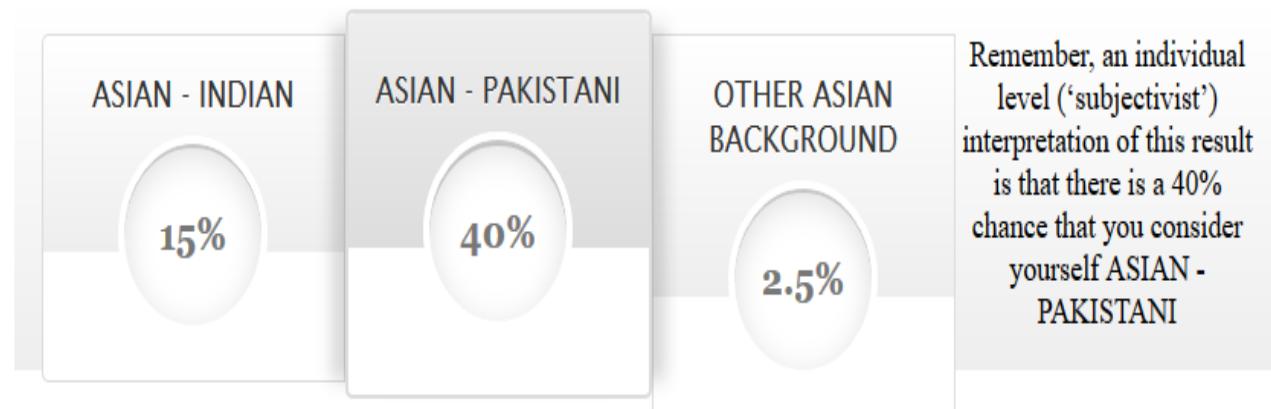
---

**CLASSIFY**

### About Ethnicity Estimator

Our classifier is based on analysis of combined names data from consumer sources assembled by CDRC to survey and survey data for England and Wales held securely by ONS. **The classifier does not use or contain any personal data** Predicting ethnicity cannot be an exact science – not least because some people share identical names despite having quite different ethnicities. We acknowledge this by reporting the three most likely estimated ethnicities for a name based on a likelihood score ranging from 0 to 1. For example, the result **"White British (35%); Asian – Pakistani (25%); Asian – Indian (12.5)"** can be understood as 'We think that there is a 25% chance that you describe yourself as White British, a 35% chance that you describe yourself as Pakistani and an 12.5% chance that you describe yourself as Indian (An alternative, frequentist, interpretation of the same numbers is that 35% of people who have your precise name describe themselves as White British, and so on. The numbers do not sum to 100% if there is a residual possibility of membership of one or more other ethnic groups and we do not report very small percentages). We can get things wrong and people themselves may see things differently. So we would really appreciate your feedback using the drop down menus provided in the next page. We have used a particular classification of ethnicity that is widely used in the UK by statistical agencies – but if you think this is unhelpful in any way please let us know this, too.'

## The ethnicity for Muhammad Gill has been estimated as:



Do you agree? Feedback will help us in our future research, so please let us know what you think, using the drop down boxes below.

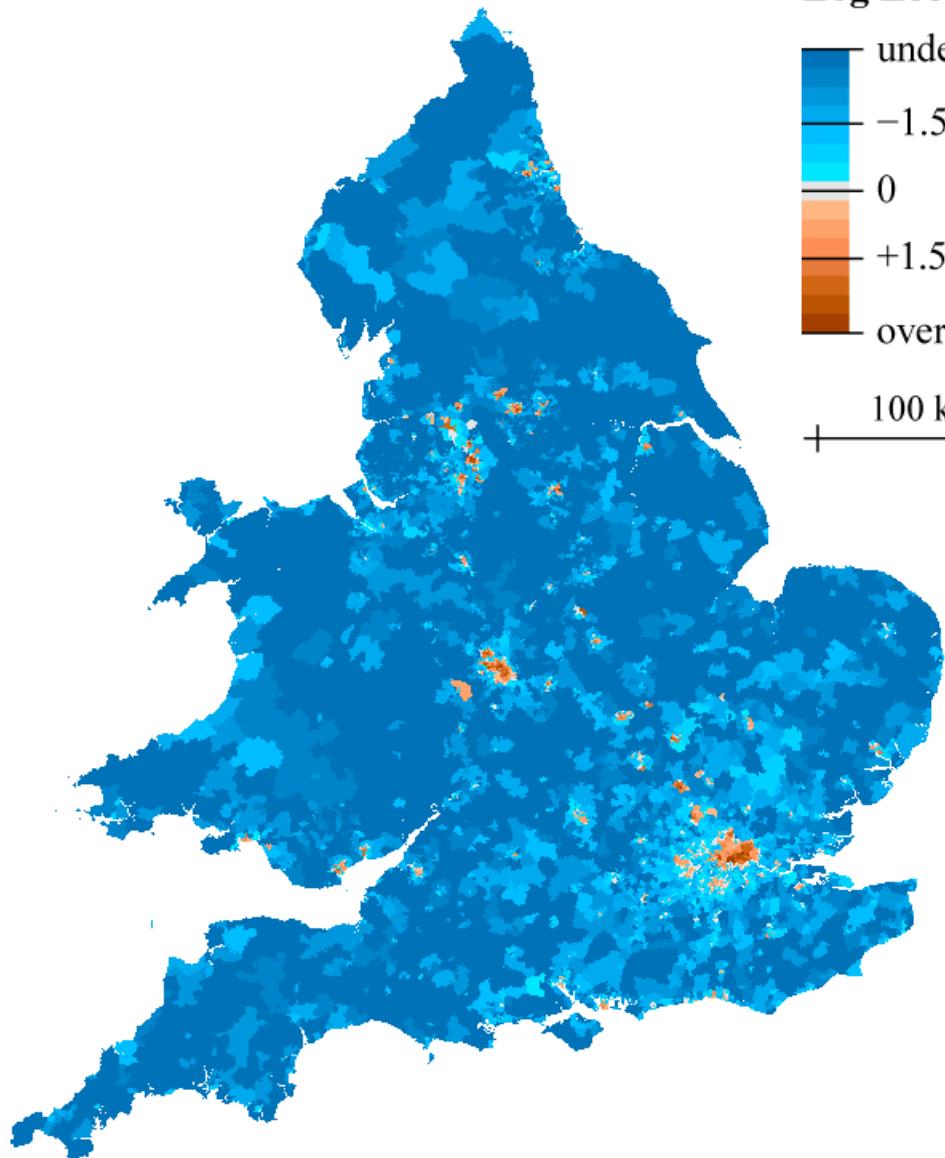
Do we got it right?

I consider myself to be:

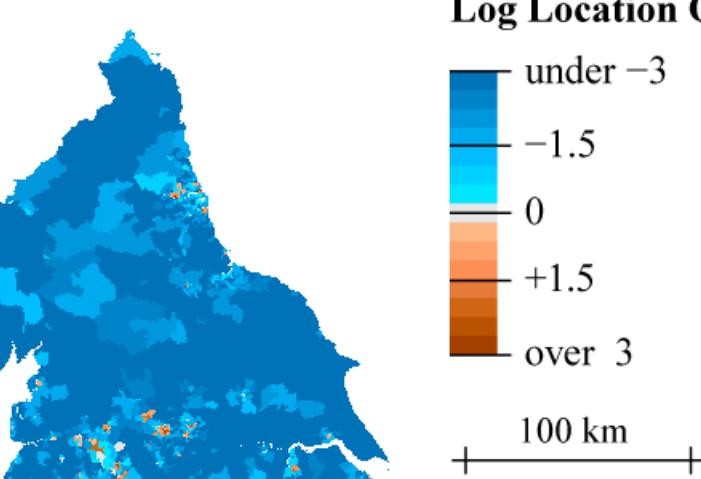
Other feedback (max 500 characters including spaces):

# SPATIAL DISTRIBUTION OF ETHNIC GROUPS

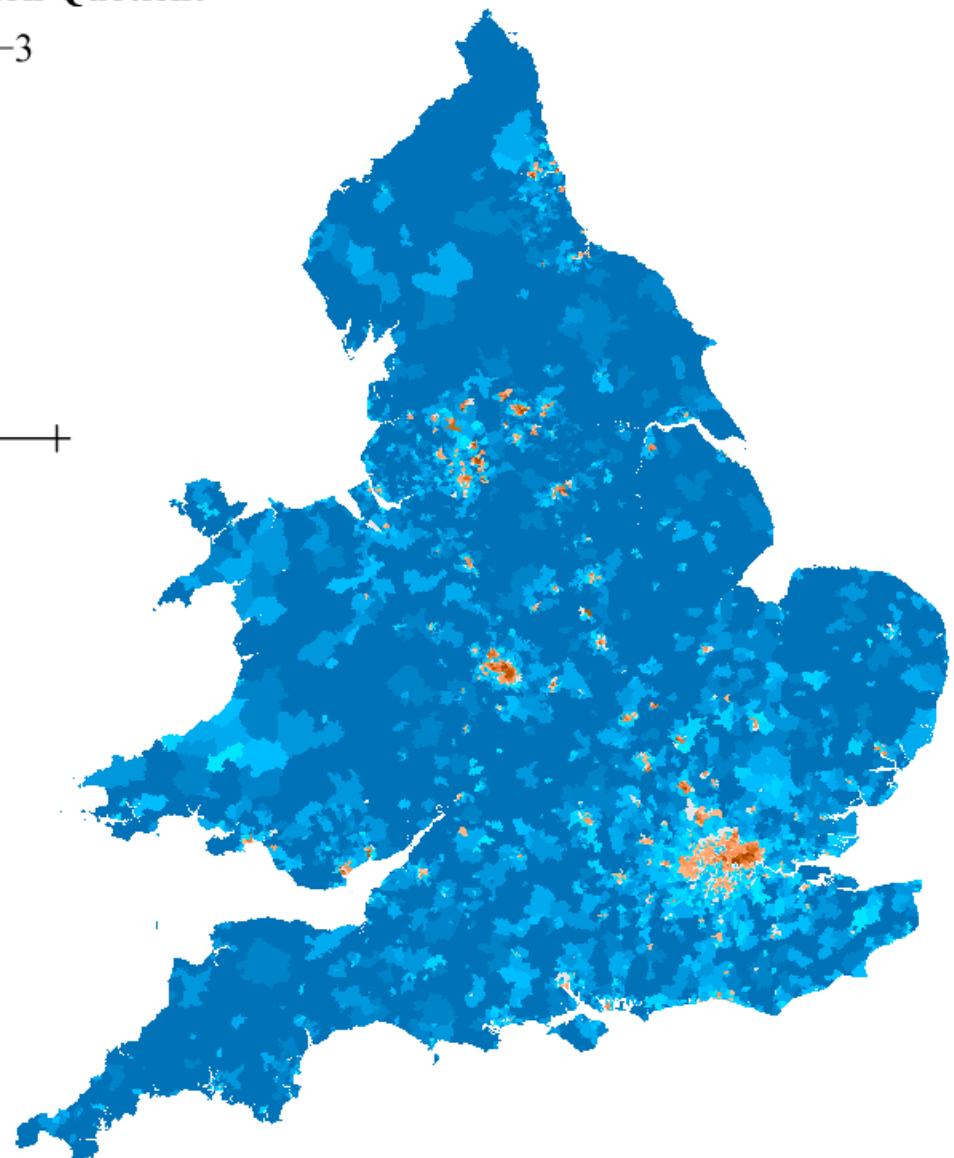
Bangladeshi (Census)



Log Location Quotient

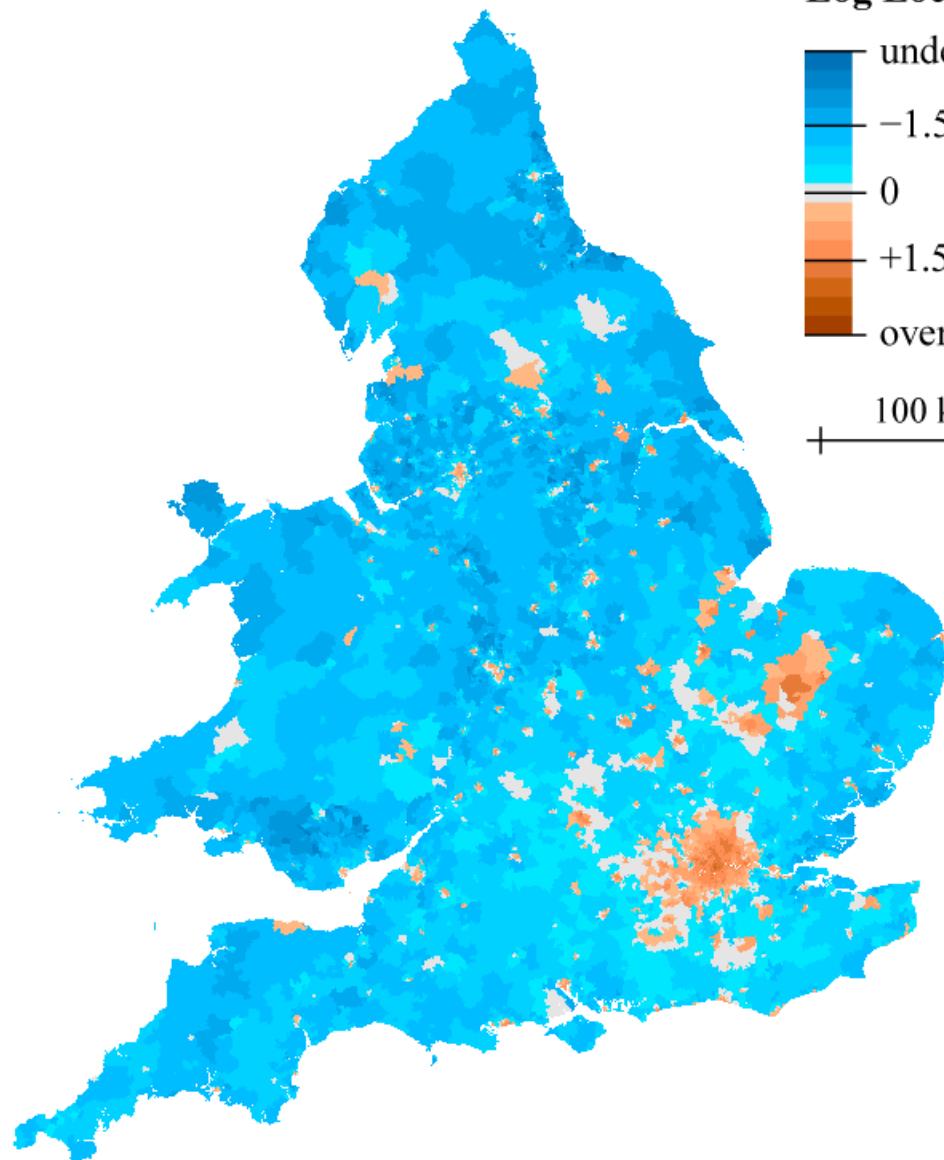


Bangladeshi (Consumer Reg.)

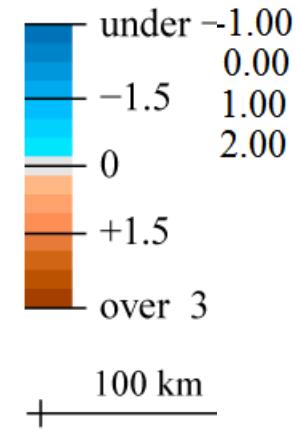


# SPATIAL DISTRIBUTION OF ETHNIC GROUPS

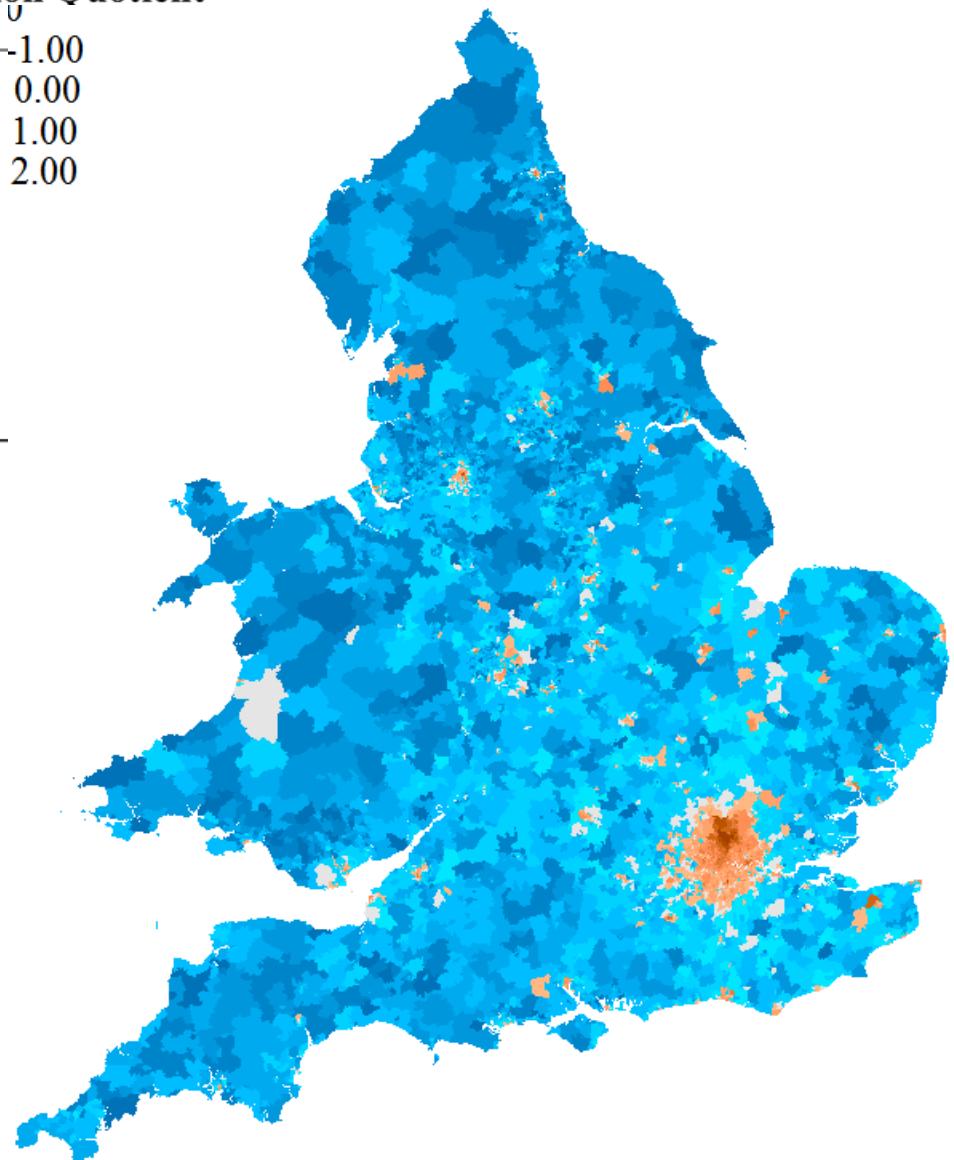
Other White (Census)



Log Location Quotient



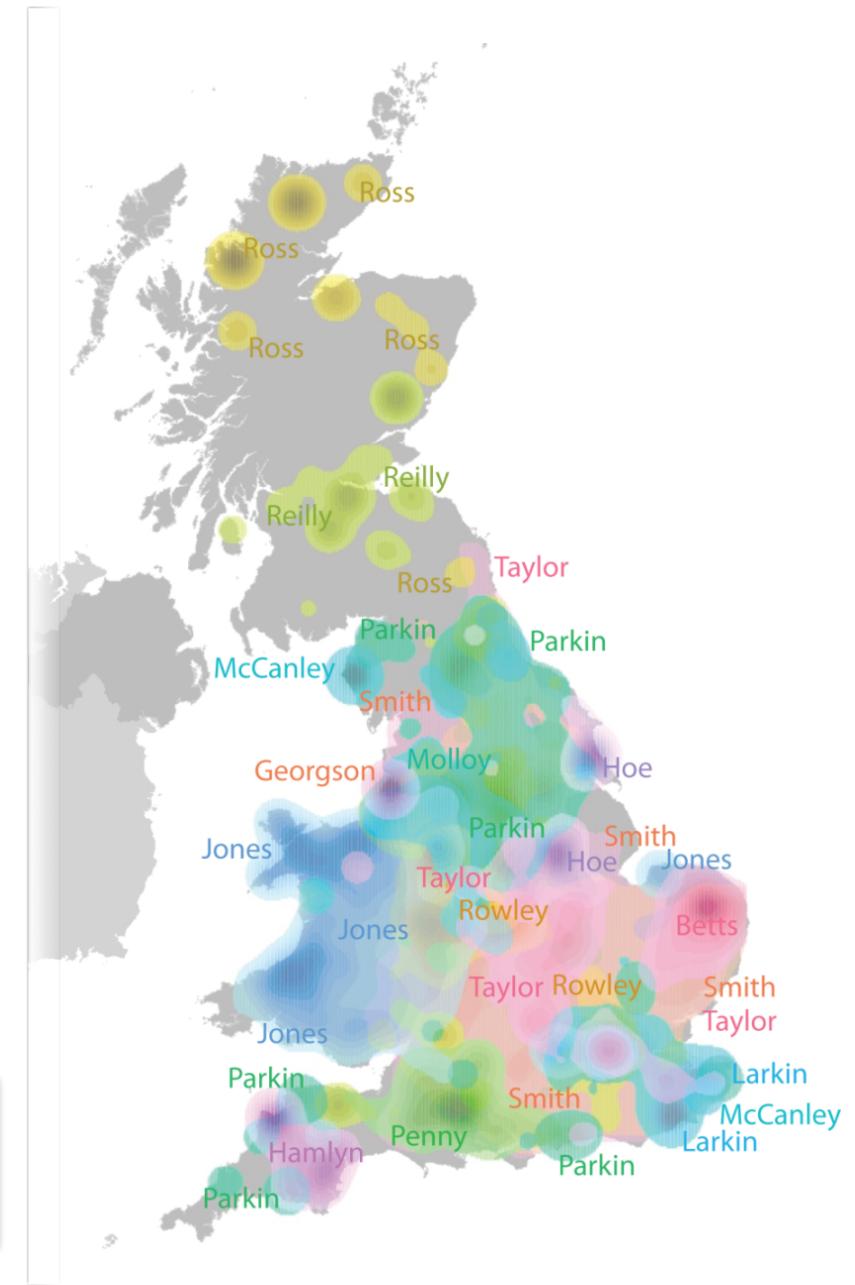
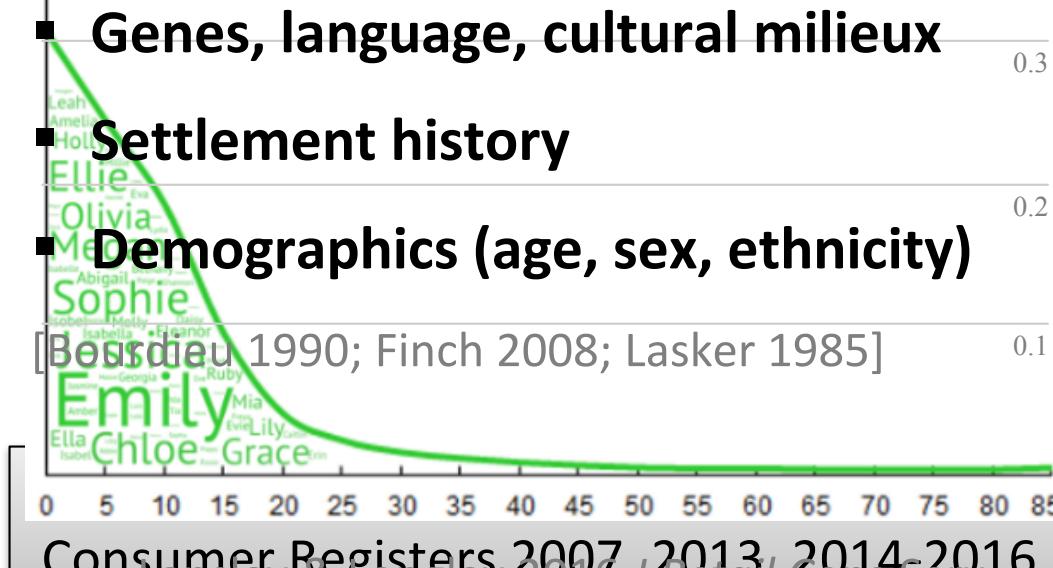
Greek (Consumer Reg.)



# Population characteristics revealed by names

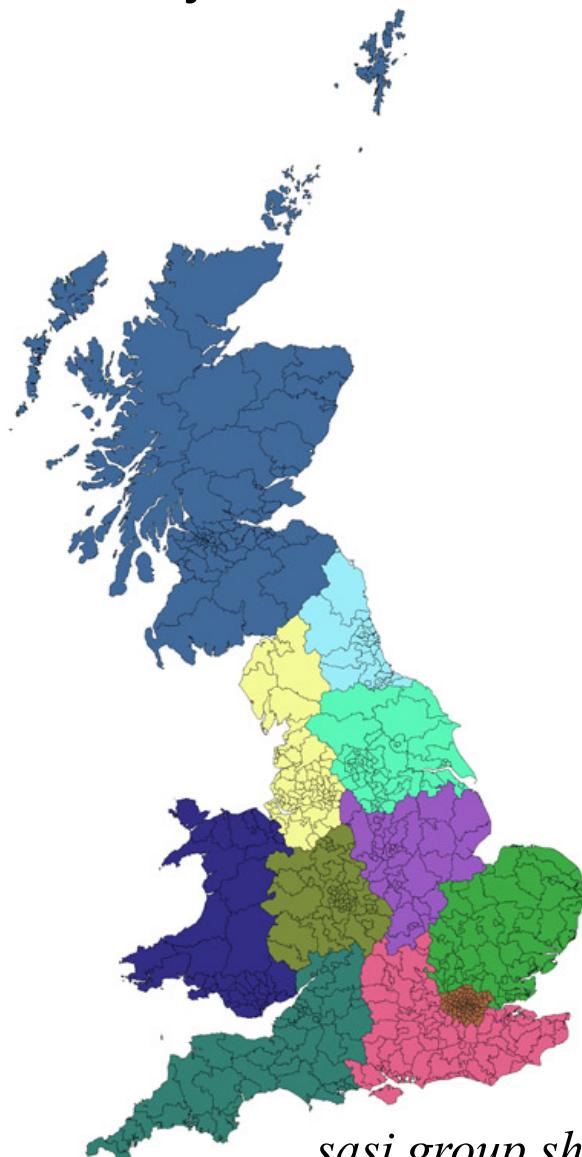
- Inheritance of surnames
- Regional name origins
- Periodical forename fashions
- Widely collected 'Big Data'

Markers of

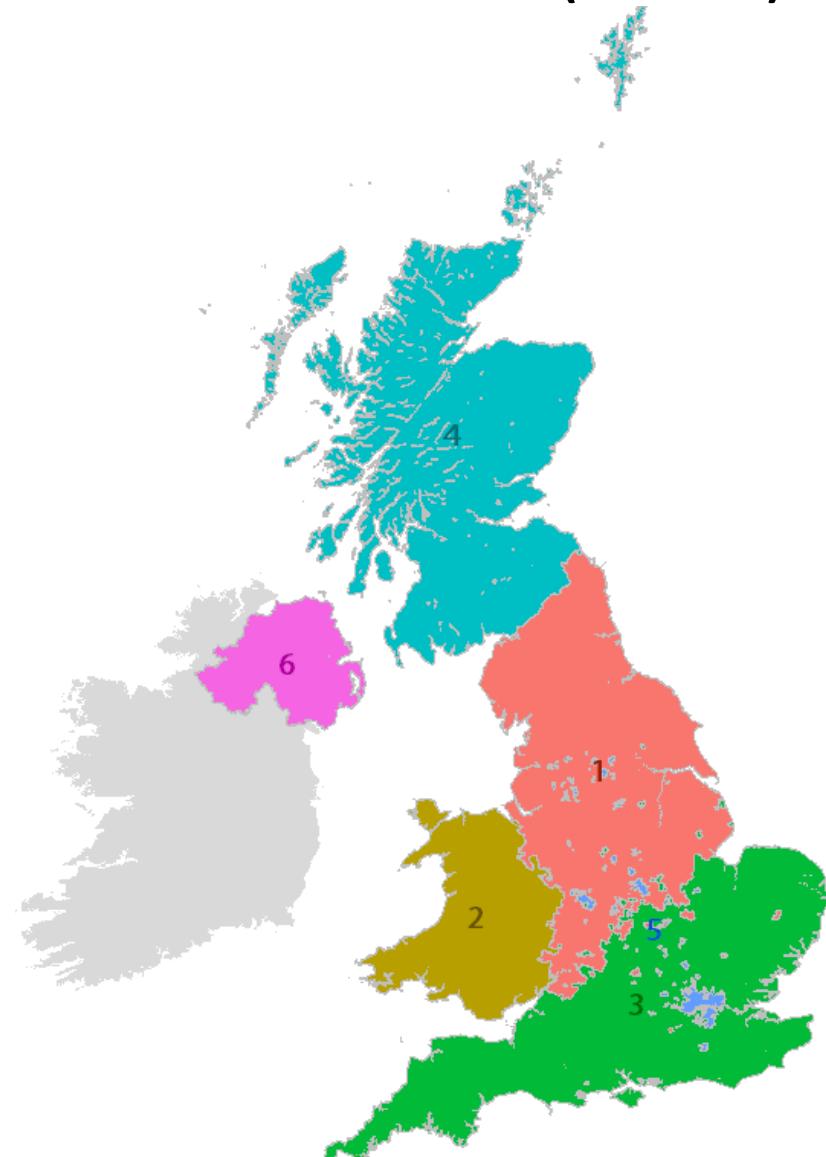


# REGIONALISATION

based on jurisdictions

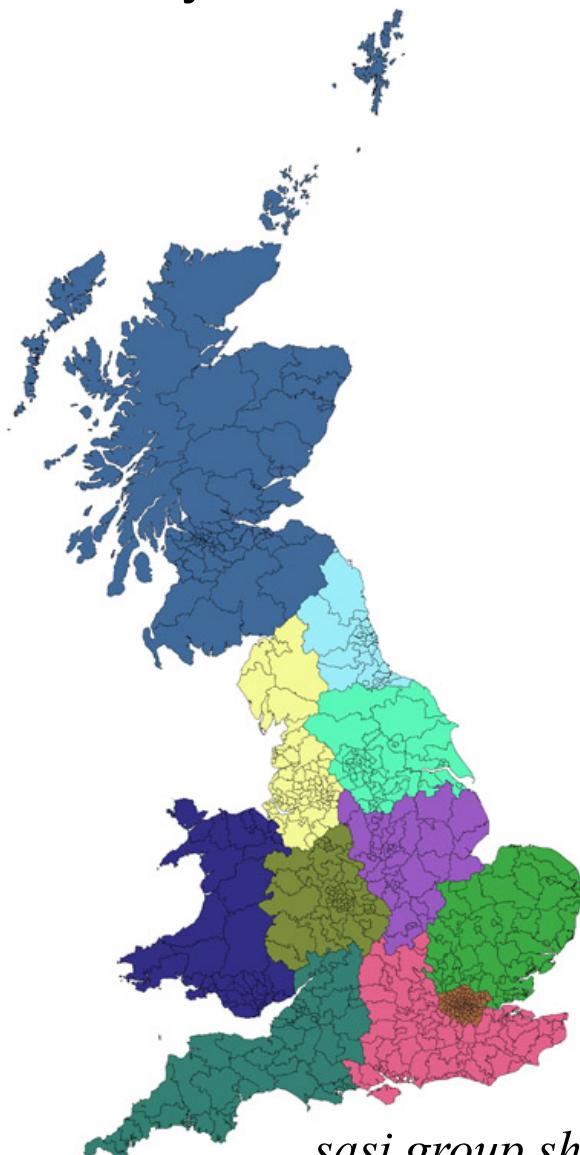


based on surnames (CR 2011)

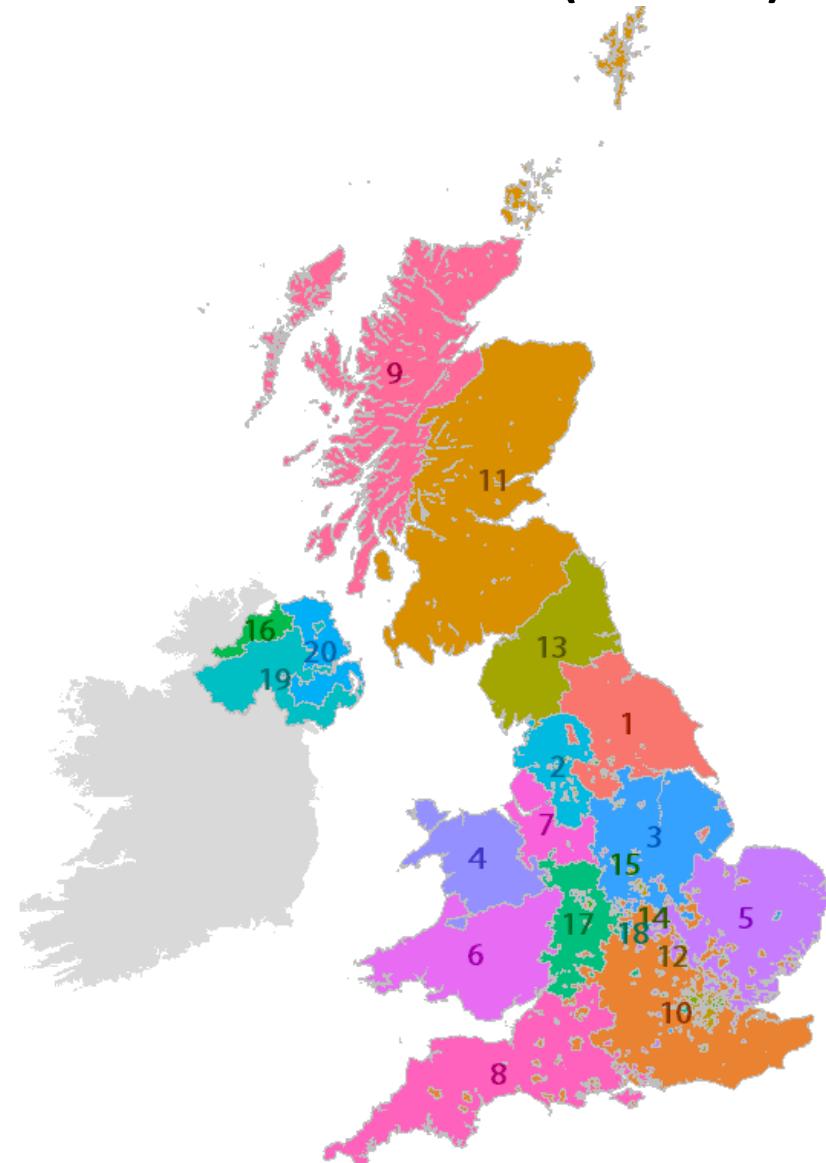


# REGIONALISATION

based on jurisdictions

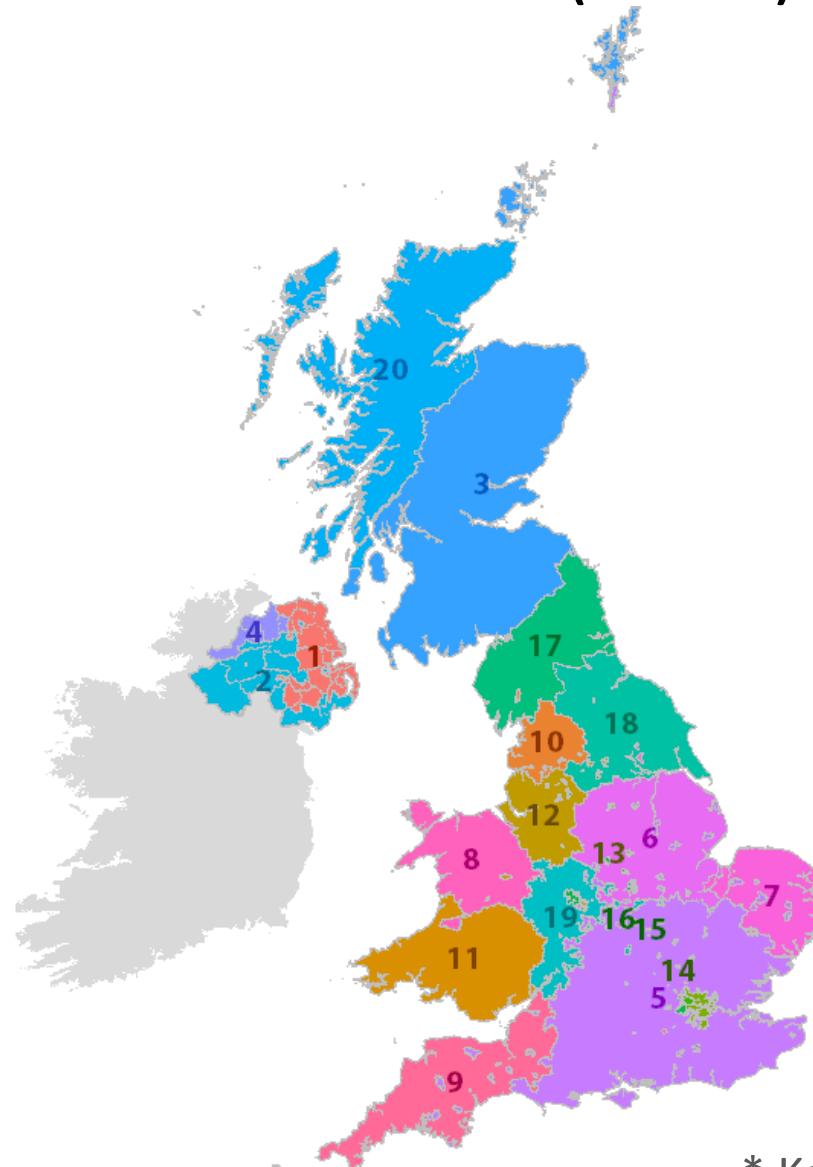


based on surnames (CR 2011)

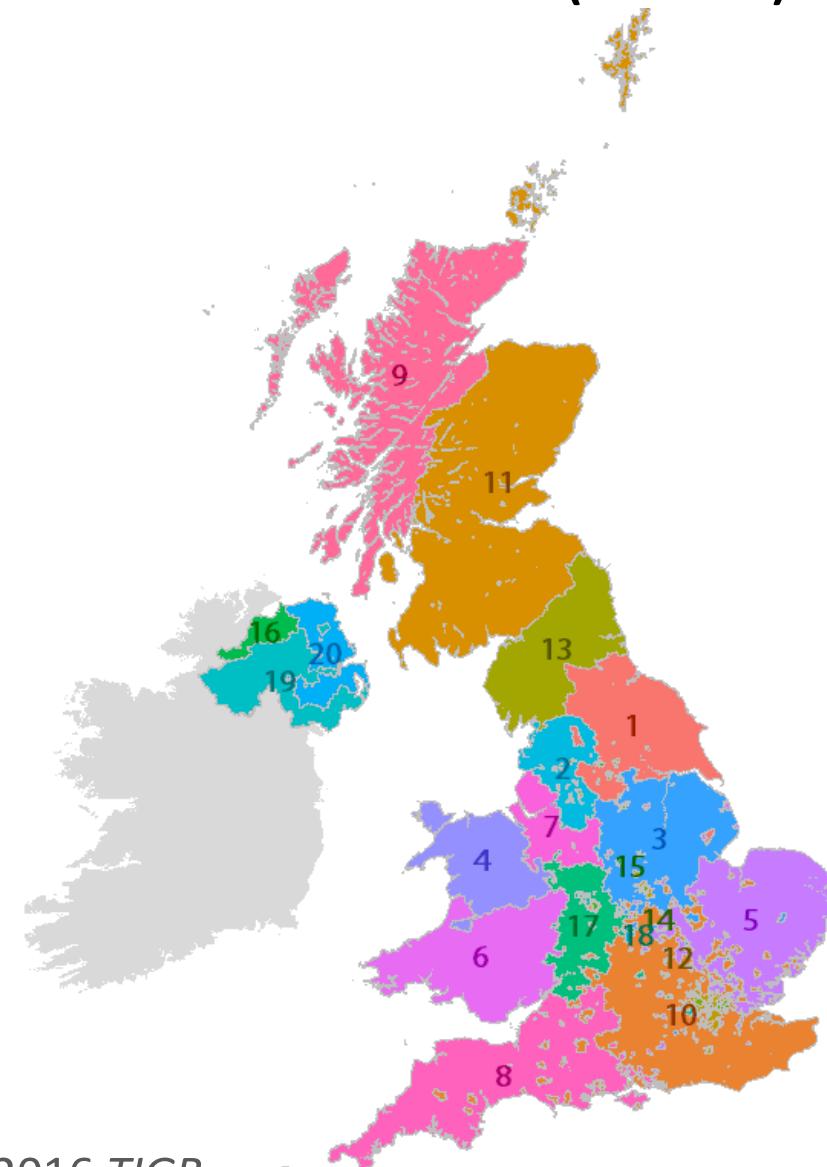


# 'TEMPORAL' REGIONAL INTEGRITY\*

based on surnames (CR 2007)



based on surnames (CR 2011)

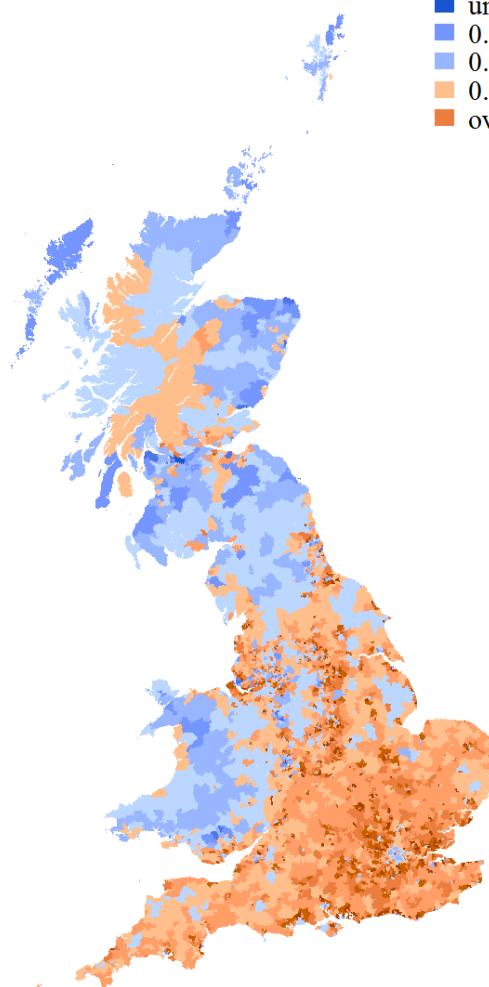


\* Kandt et al 2016 *TIGB*

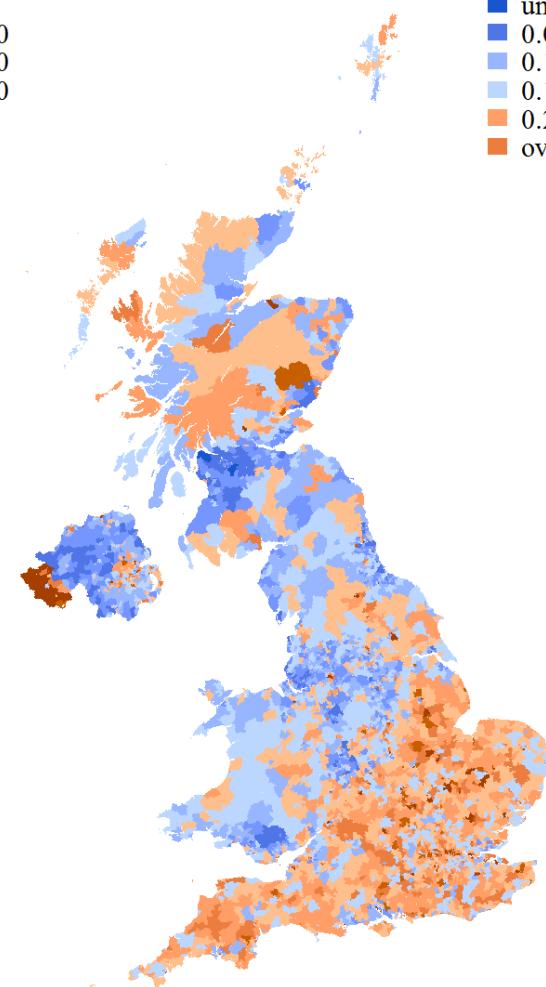


# PROPORTION OF NEWLY IMPORTED SURNAMES

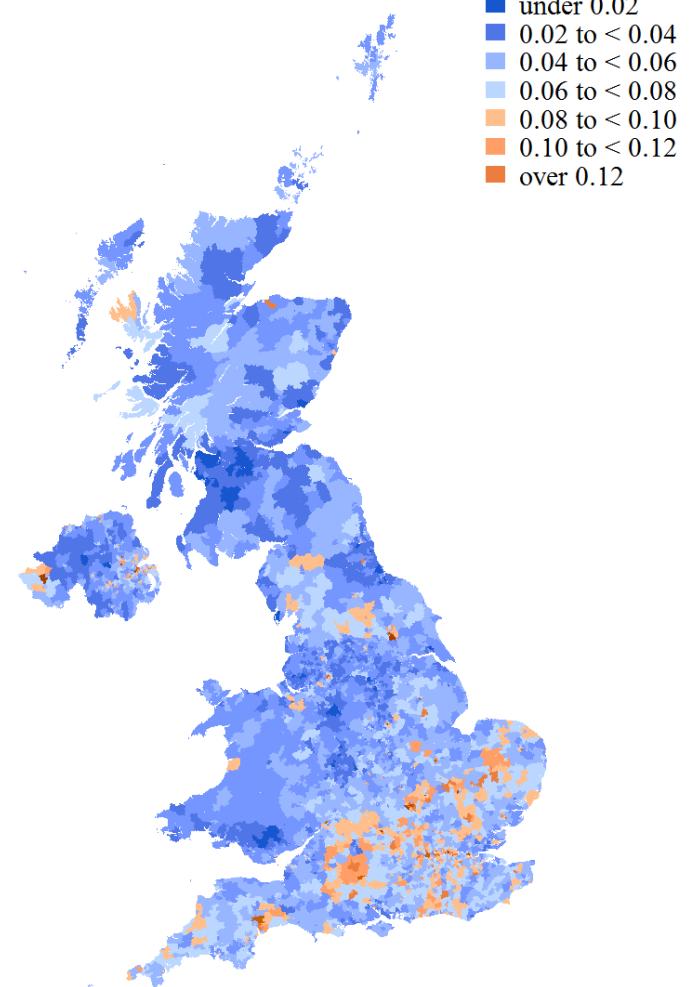
1881-2007

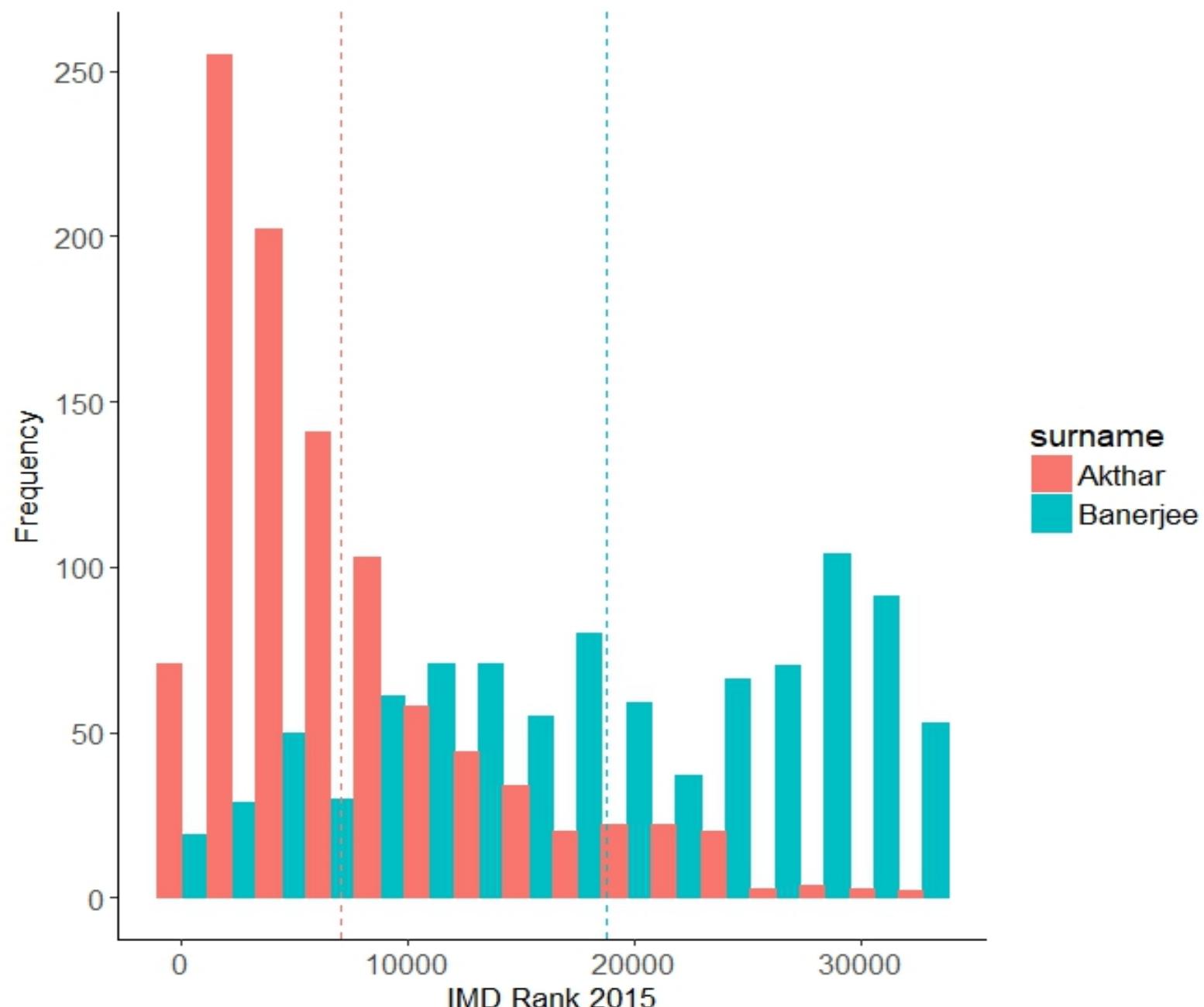


2007-2013



2013-2016





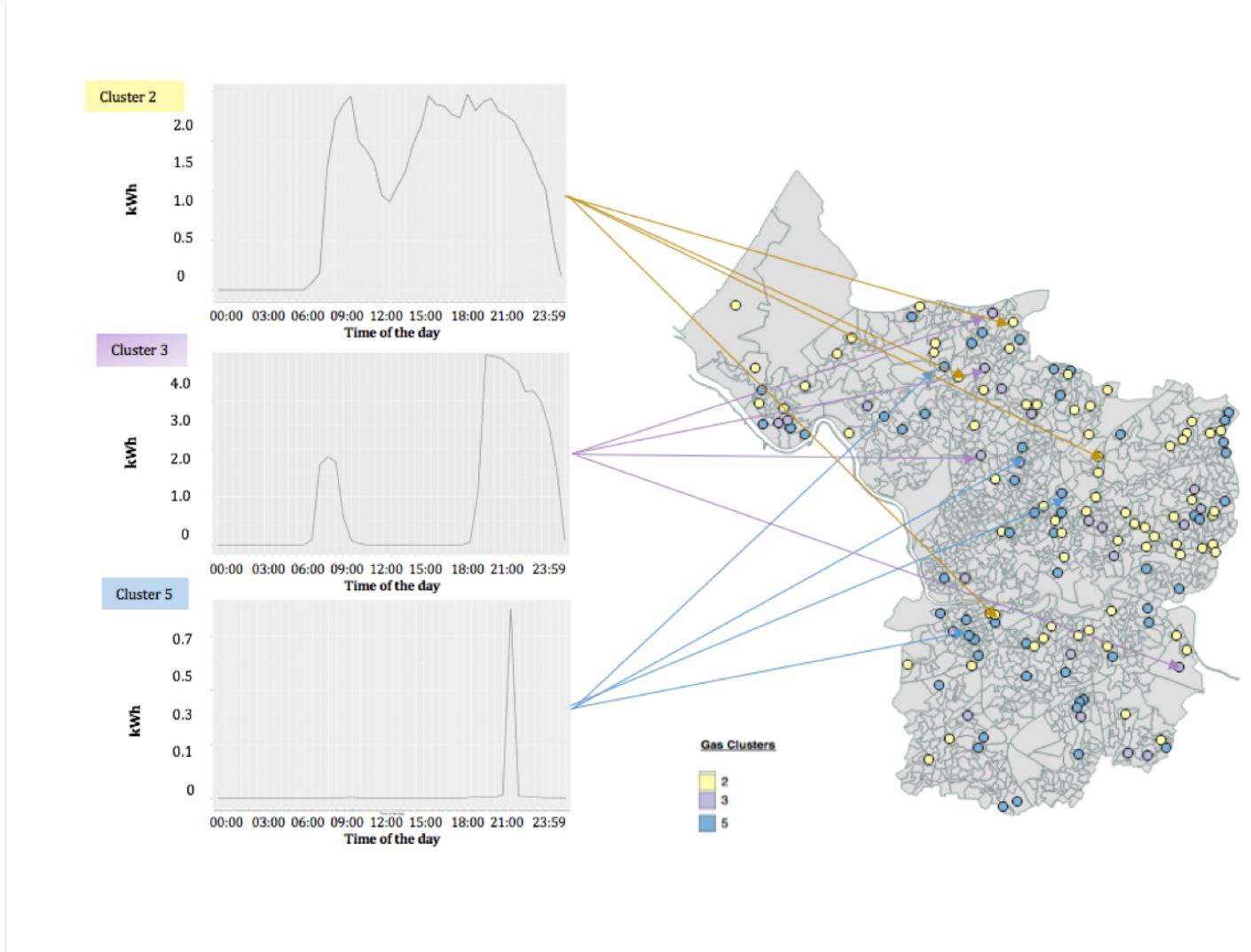


Consumer  
Data  
Research  
Centre

An ESRC Data  
Investment

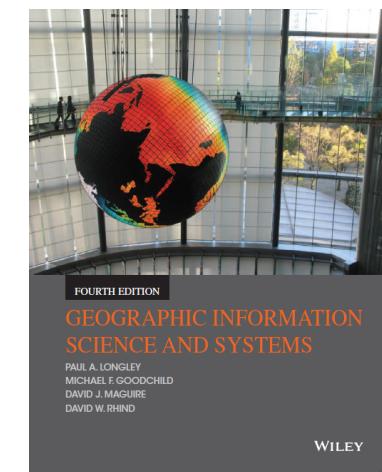
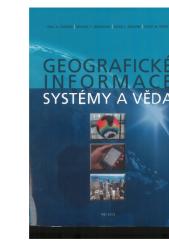
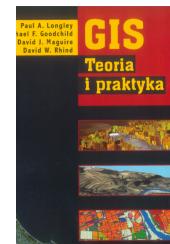
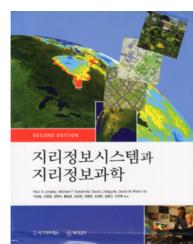
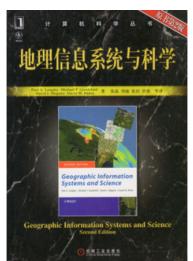
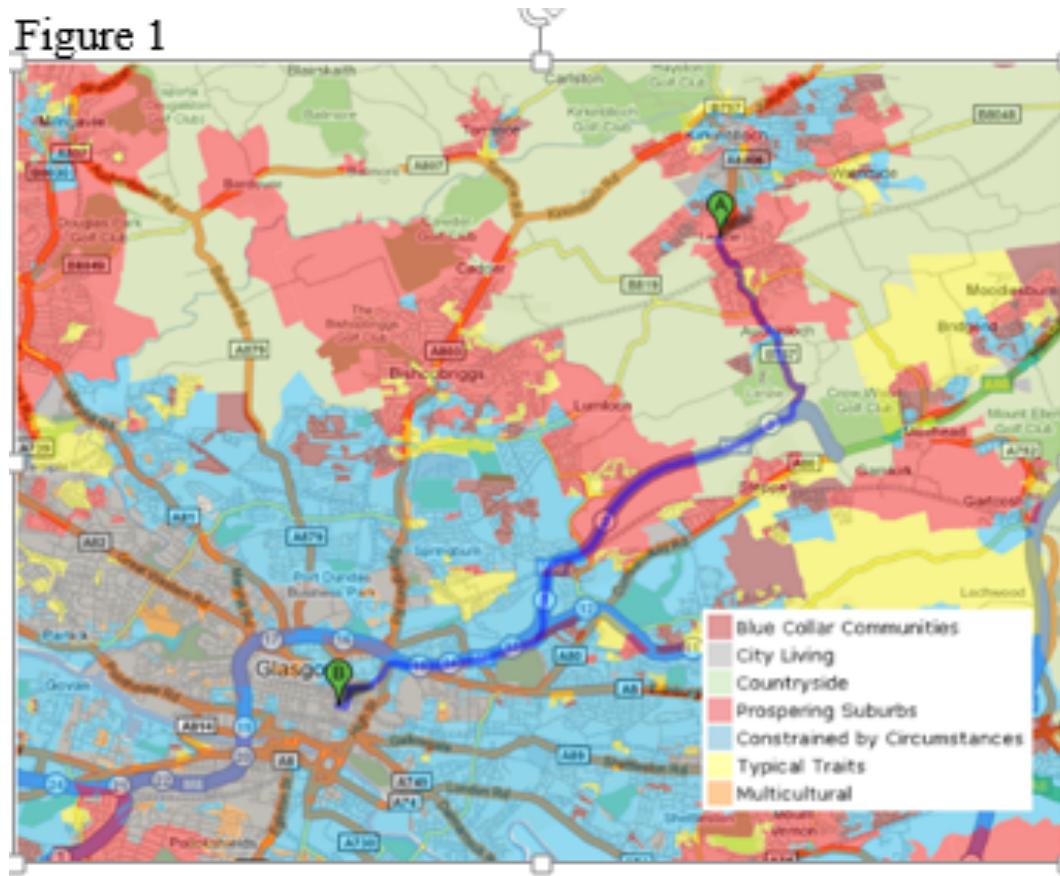
# Profiling Customers

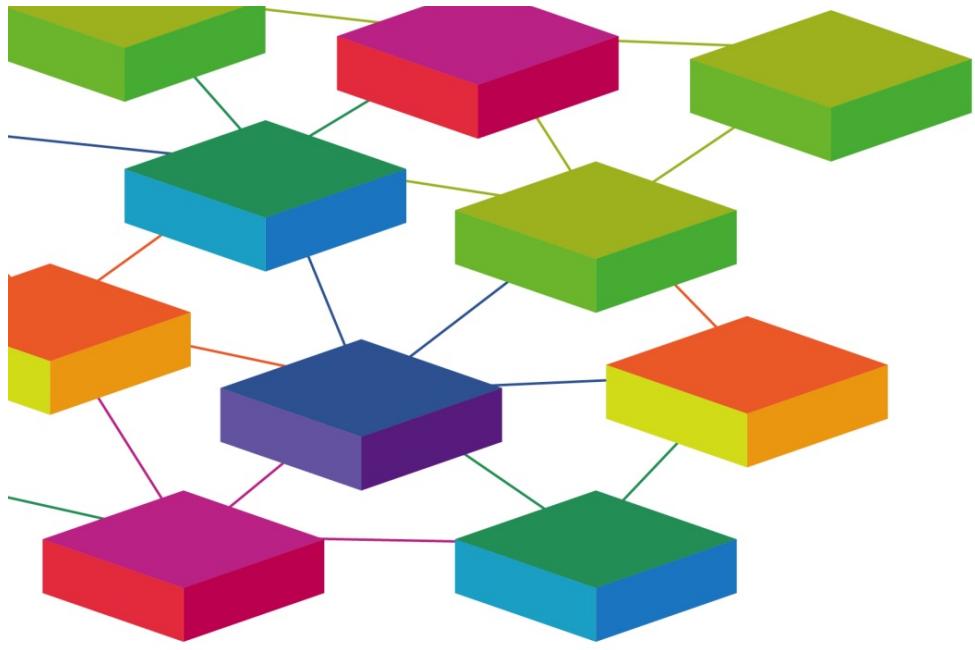
## City of Bristol





**Figure 1**





[www.cdrc.ac.uk](http://www.cdrc.ac.uk)