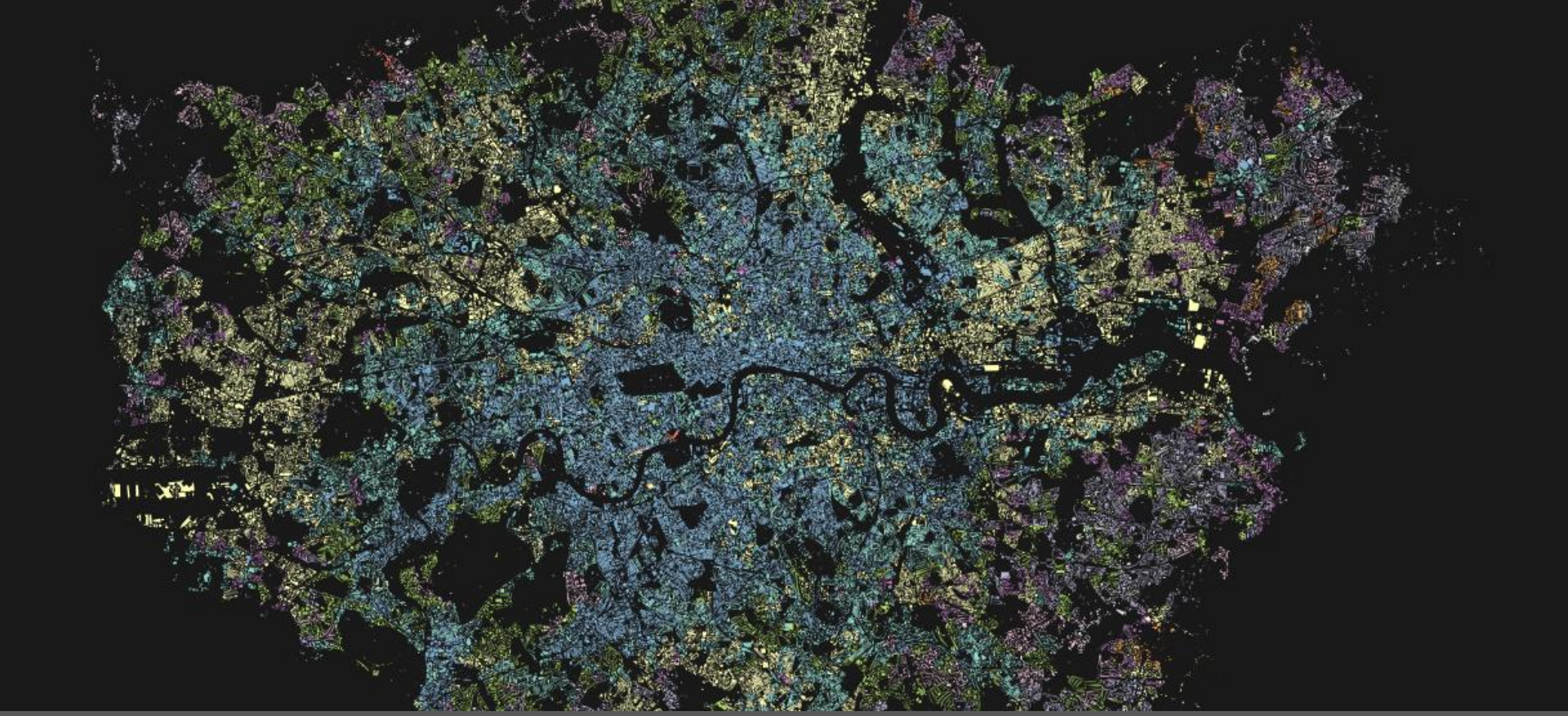


Identifying Regional Digital Cultures:

Developing a national classification of Internet use and engagement

UNIVERSITY OF LIVERPOOL

Dean Riddlesden

Geographic Data Science Laboratory

LIFE CHANGING | World Shaping

Outline

- Background
- Why?
- The three fundamentals of engagement:
  - Infrastructure
  - Behaviors
  - Demographics
- Developing a national classification
- Assessment and impact
- Conclusions

Background

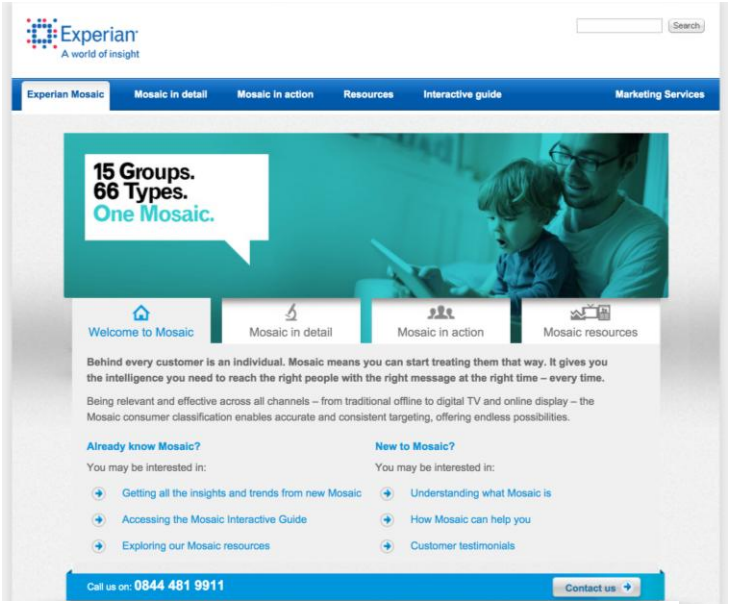
Geodemographics links the sciences of geography and demography

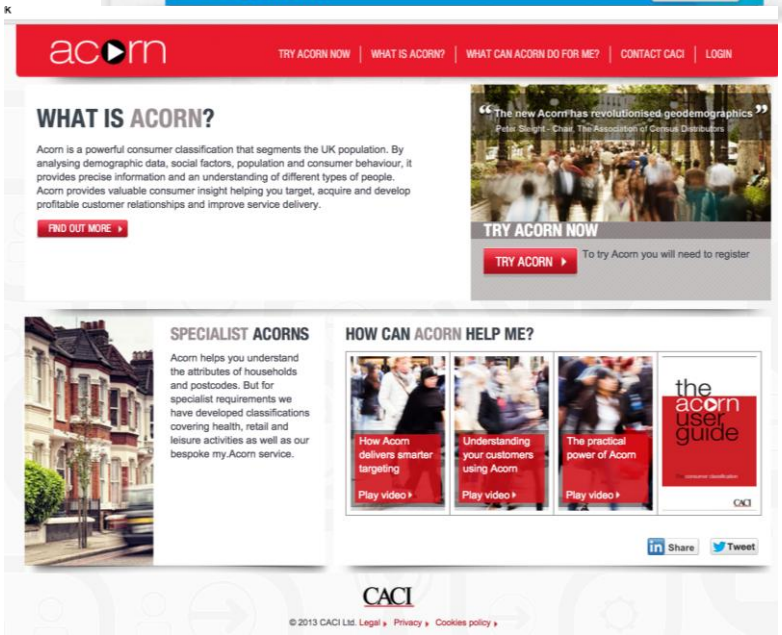
Commercial geodemographic classifications began to emerge in the late 1970's

- PRIZM (Claritas) USA
- Acorn (CACI) UK
- Mosaic (Experian)

Predominantly used in the targeting of 'ideal' populations for products and services

- Some public sector significance
- Recent shift from 'black-box' to open source



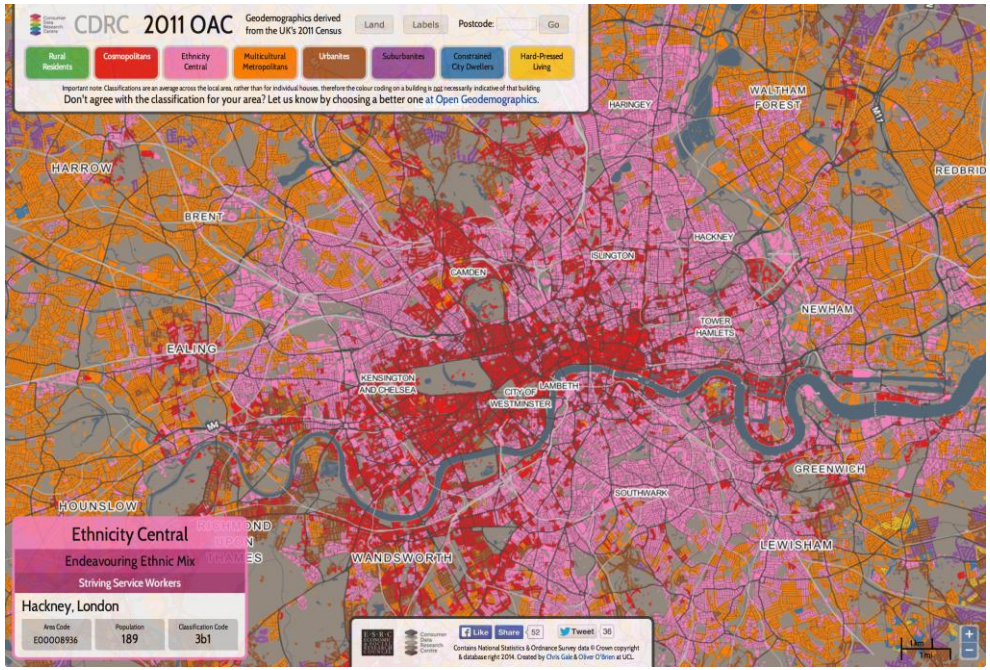


Background cont.

Open source geodemographics

- Output Area Classification (OAC)

- Initially 2001
- Updated 2011
- In line with Census releases from ONS



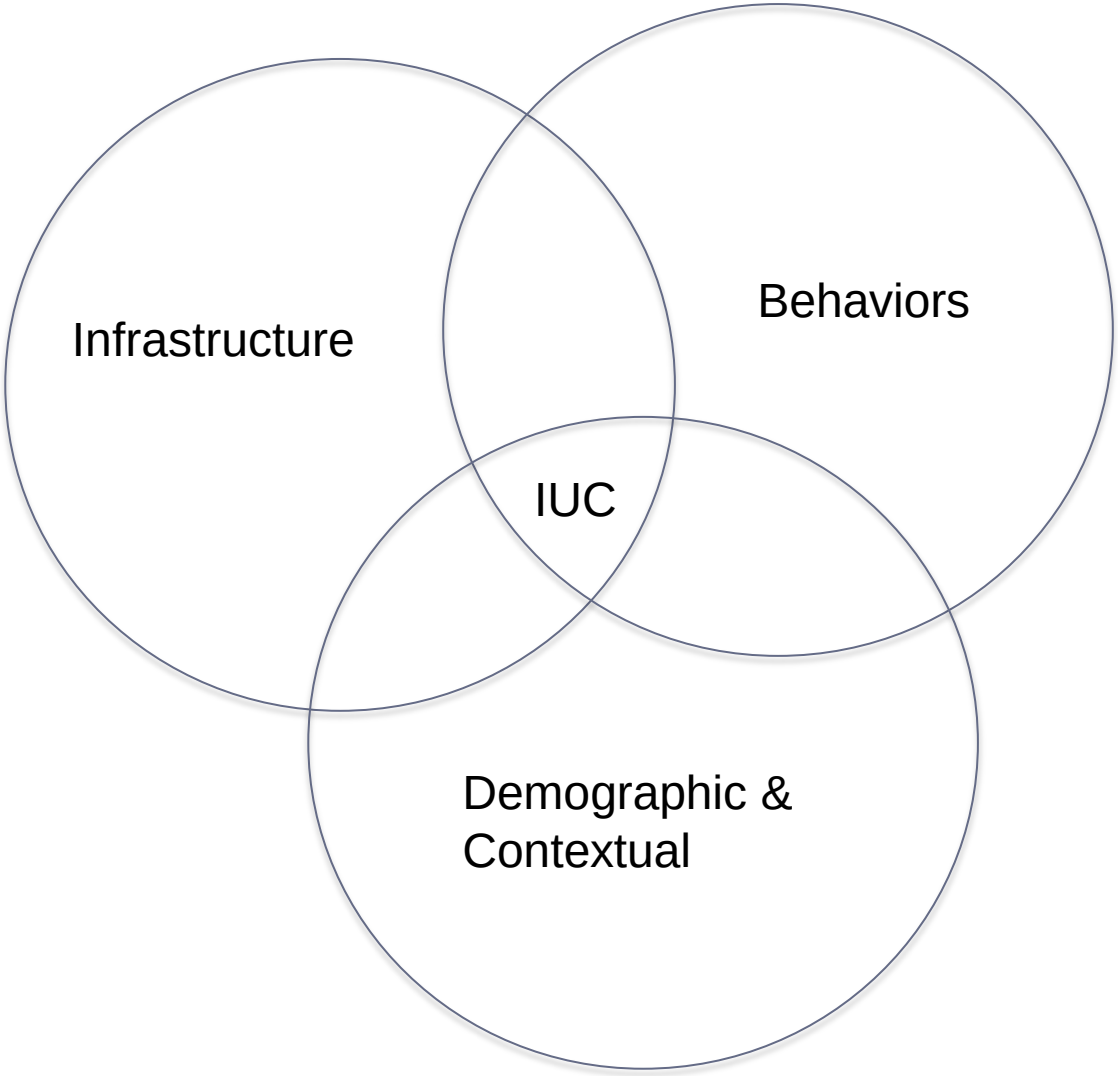
- E-society Classification

- 2007
- Predominantly based on lifestyle and consumer survey data/ some census

- Internet User Classification (IUC)

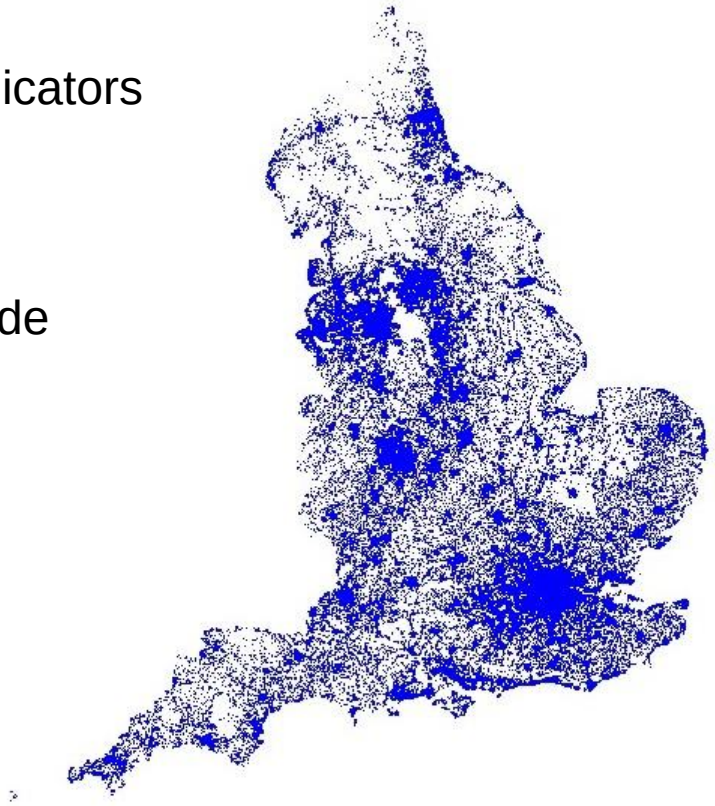
- Survey/ Crowdsourced data/ Census/ Infrastructure

The Three Fundamental Requirements



1. Infrastructure

- Comprehensive studies of English broadband infrastructure (fixed-line and mobile)
- Performance and access evaluated by indicators of socio-spatial structure (OAC and open geodemographic profiling)
- A geography of supply to consider alongside demand
- What are the apparent disparities?



Applied Geography
Volume 52, August 2014, Pages 25–33

Broadband speed equity: A new digital divide?

Dean Riddlesden, Alex D. Singleton

Show more

DOI: 10.1016/j.apgeog.2014.04.008

Get rights and content

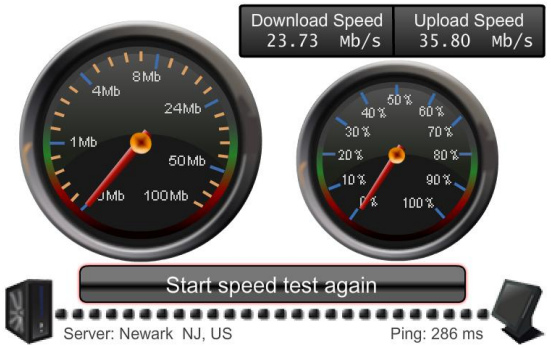
Open Access funded by Economic and Social Research Council

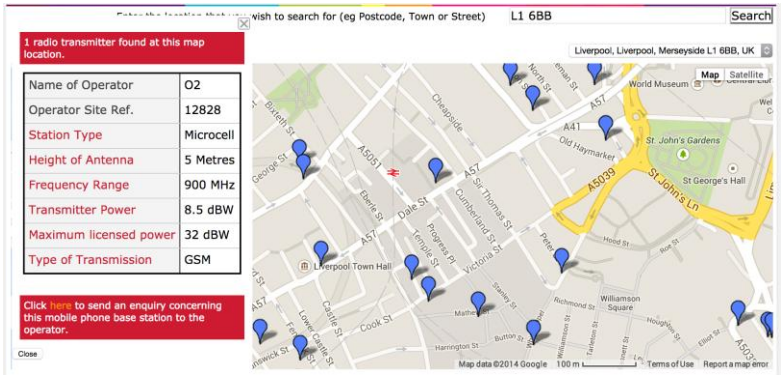
Infrastructure Data

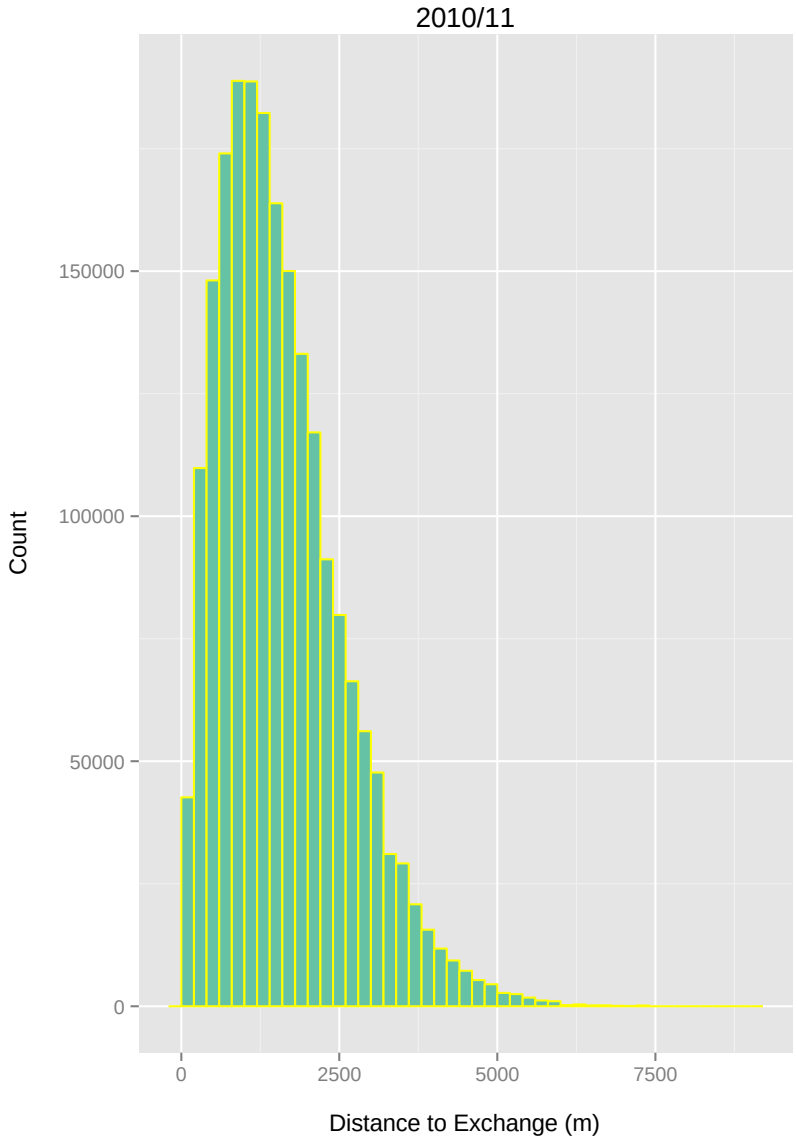
- Hard to find, but essential to the project
- No public data/ Incumbent suppliers don't like to share
- The OfCom Sitefinder fiasco/ had to get creative

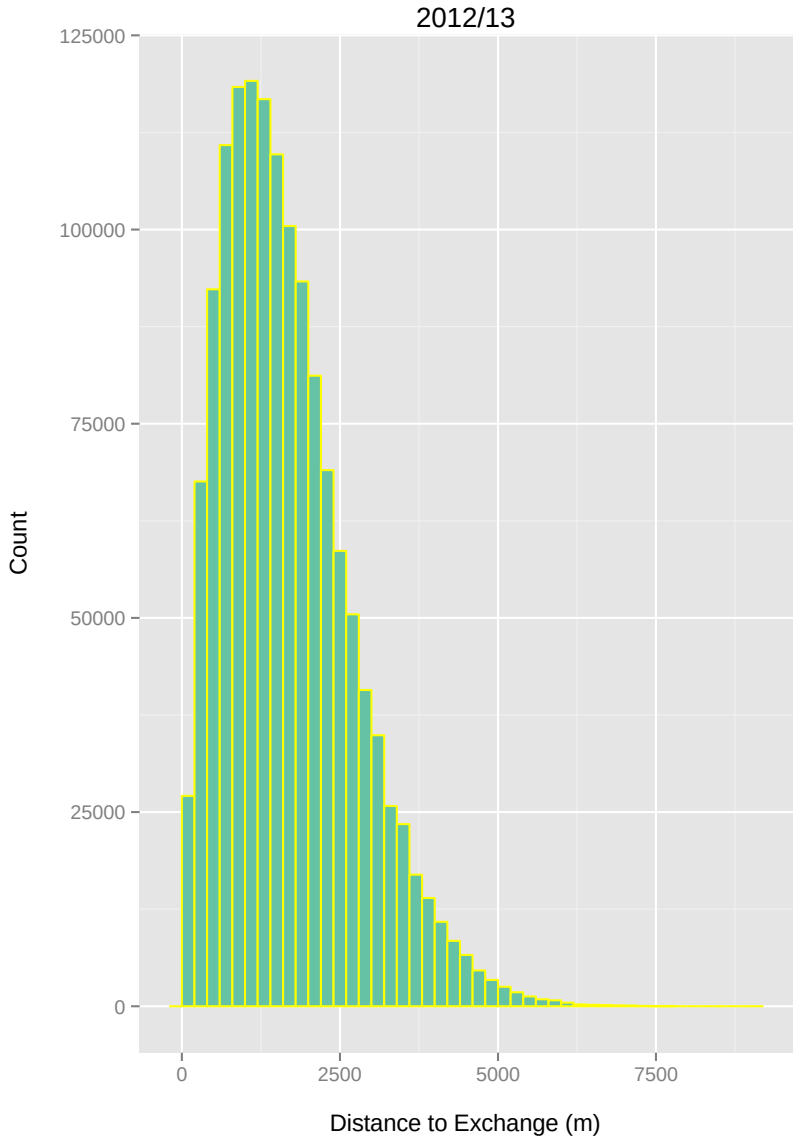
Data:

- 7.6m geo-tagged speed test estimates
- Exchange location database (web scraped)
- OfCom Sitefinder database (eventually)
- Laborious 'Janitor' work



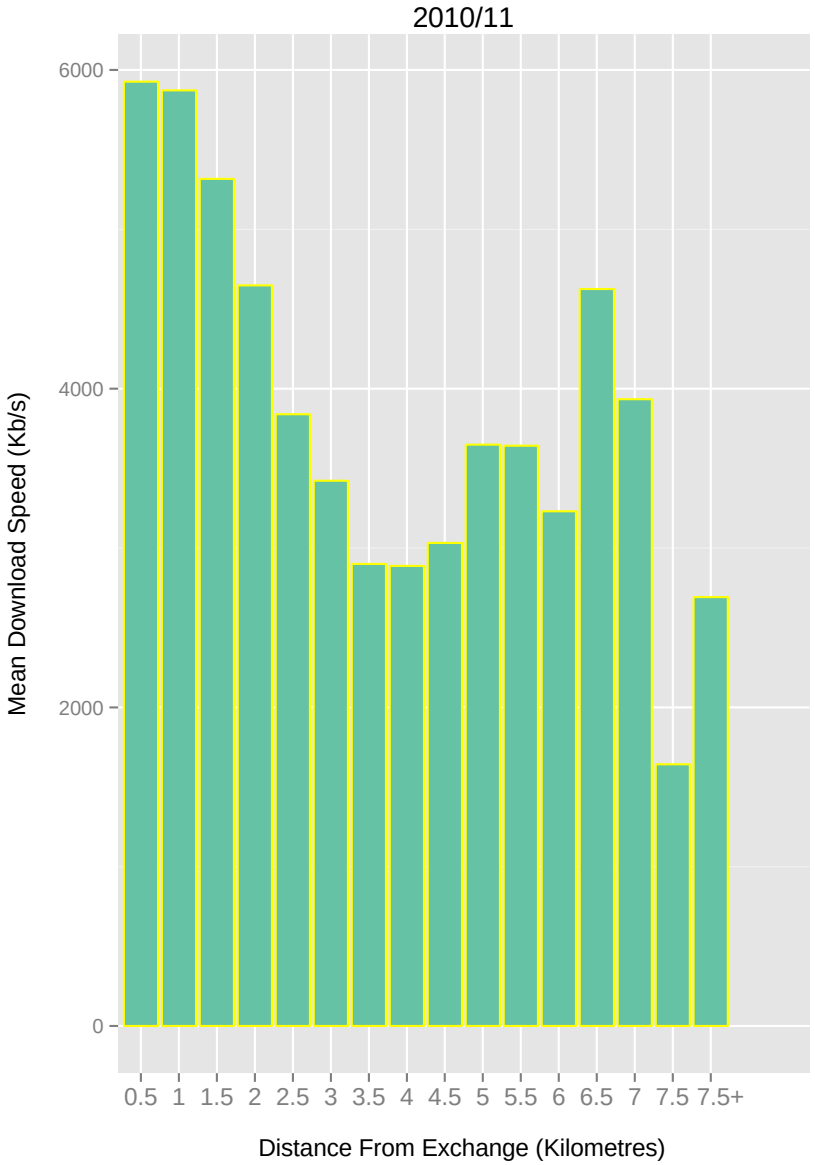


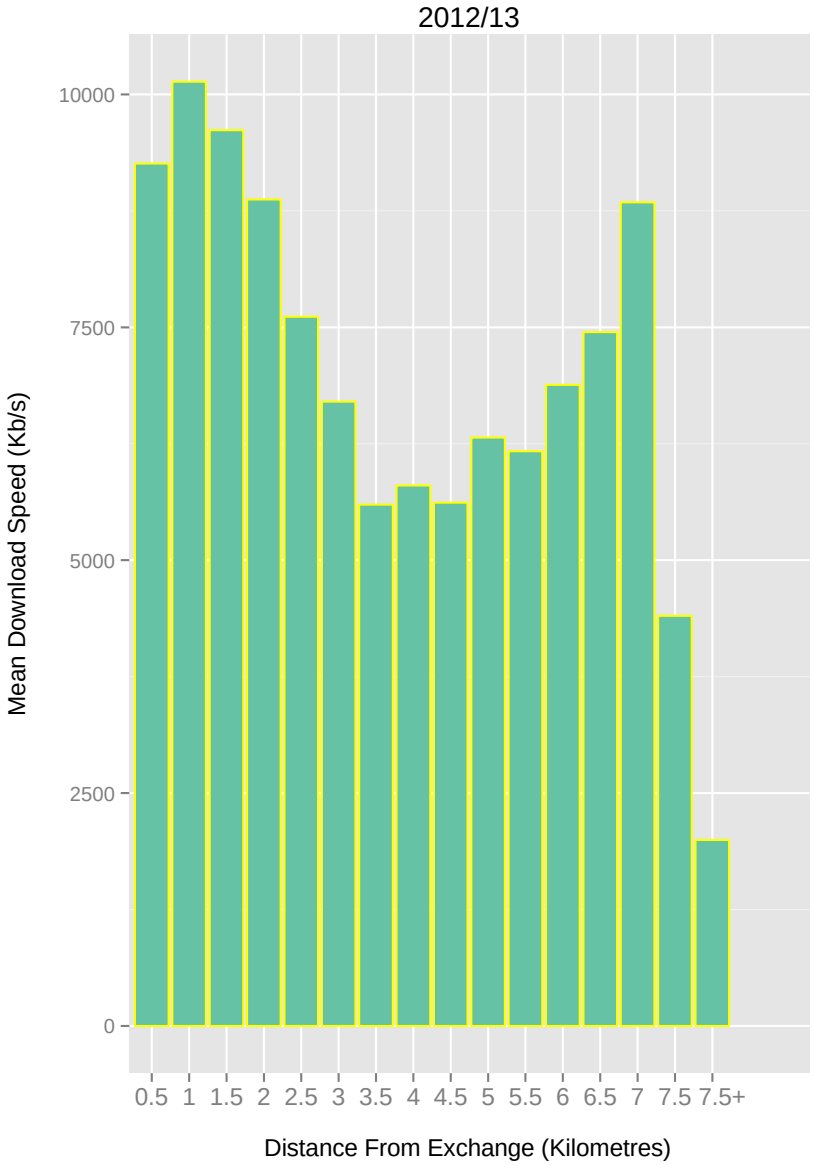




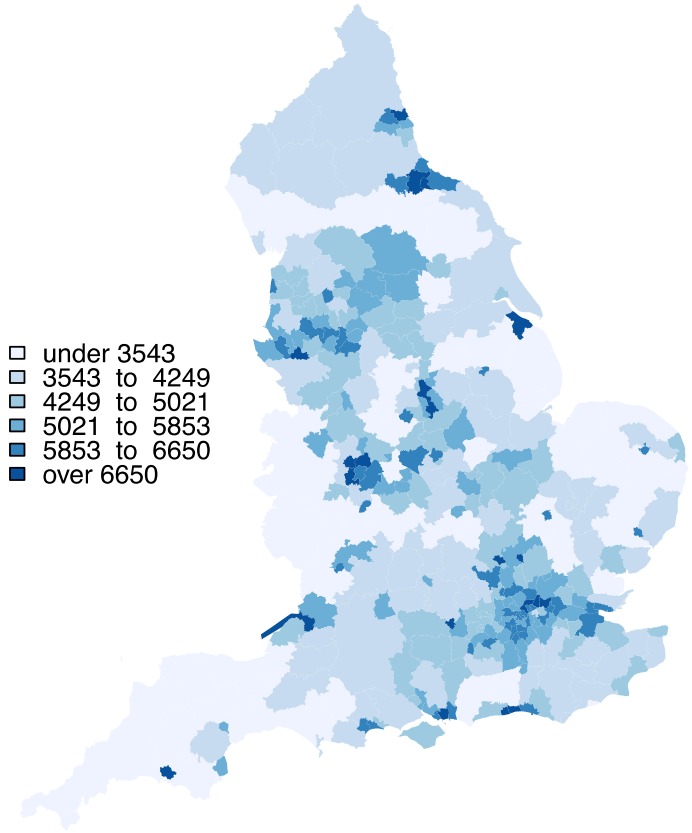
UNIVERSITY OF LIVERPOOL

School of Environmental Sciences

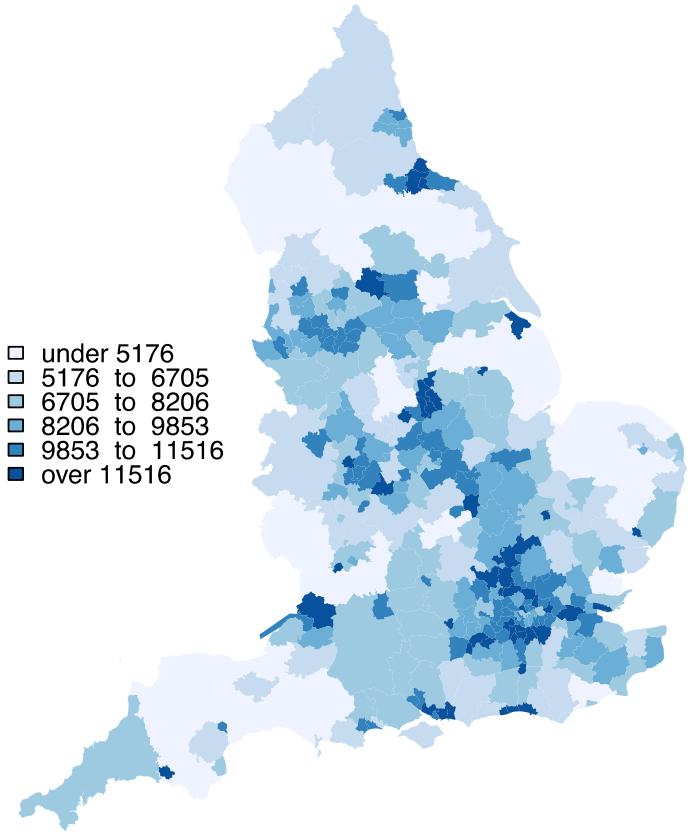

2010/11
Mean Download Speed (Kb/s)
6000
4000
2000
0
0.5 1 1.5 2 2.5 3 3.5 4 4.5 5 5.5 6 6.5 7 7.5 7.5+
Distance From Exchange (Kilometres)


2012/13
Mean Download Speed (Kb/s)
10000
7500
5000
2500
0
0.5 1 1.5 2 2.5 3 3.5 4 4.5 5 5.5 6 6.5 7 7.5 7.5+
Distance From Exchange (Kilometres)

Mean Download Speeds (Kbps) by English District 2010/11

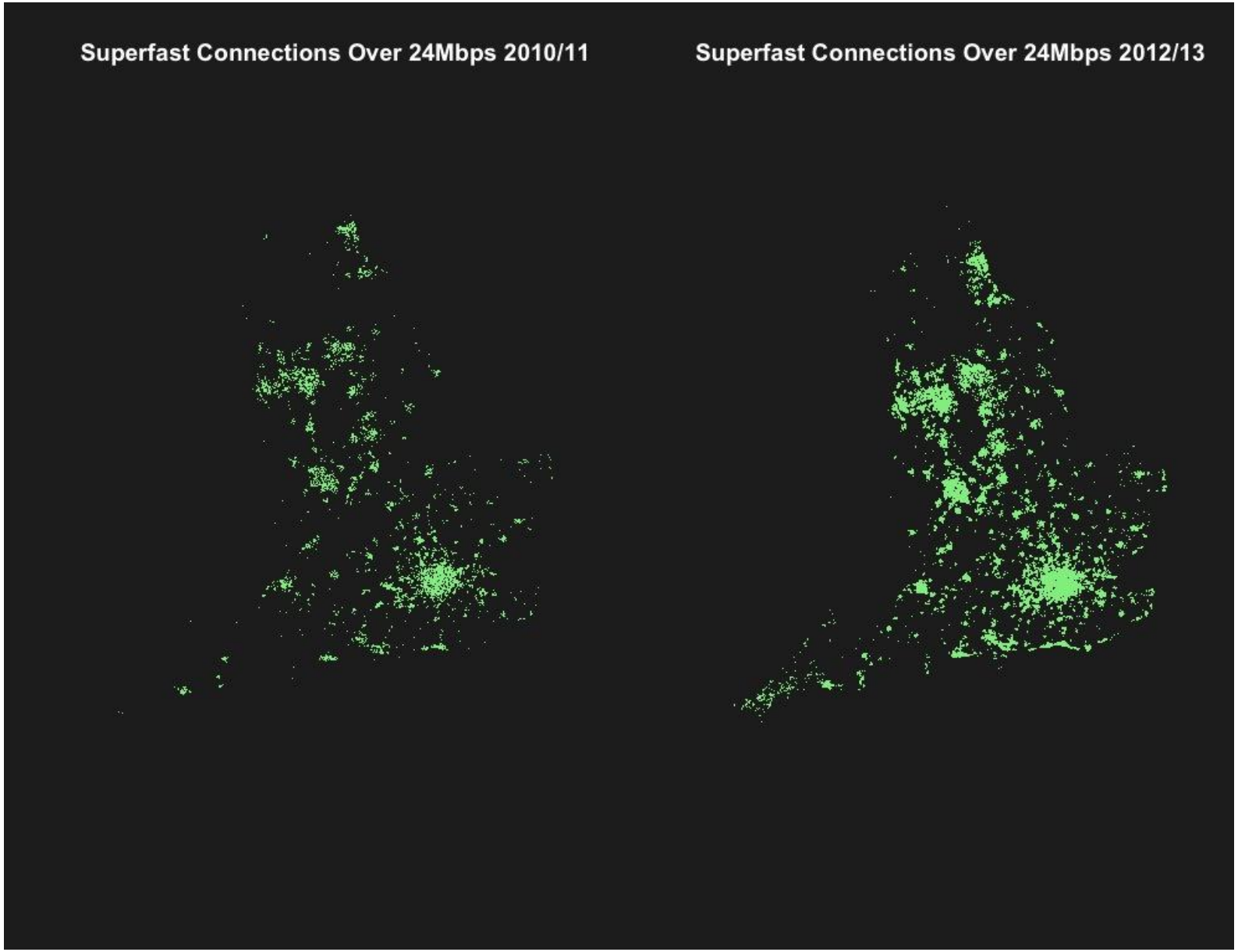


Mean Download Speeds (Kbps) by English District 2012/13

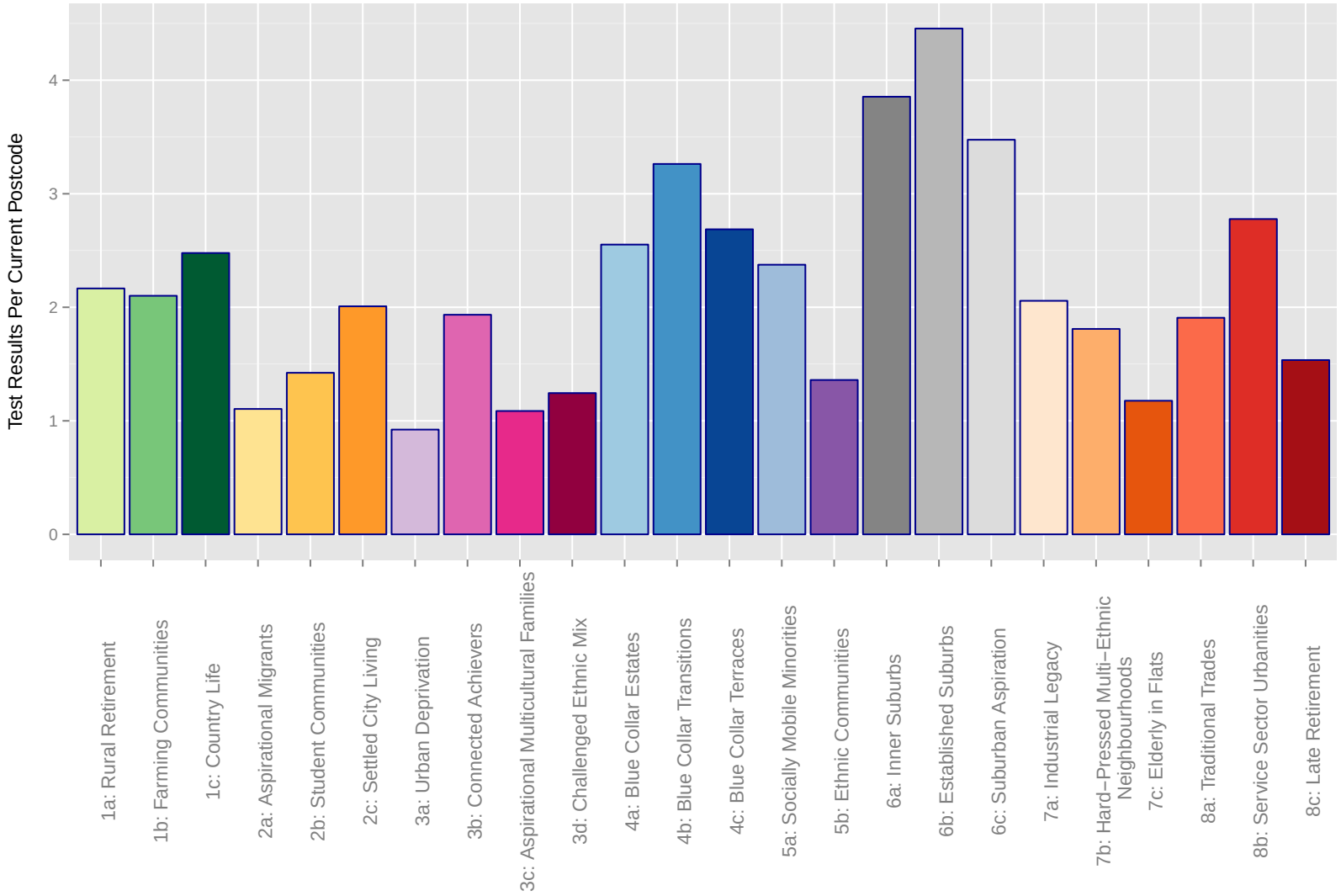


UNIVERSITY OF LIVERPOOL

School of Environmental Sciences

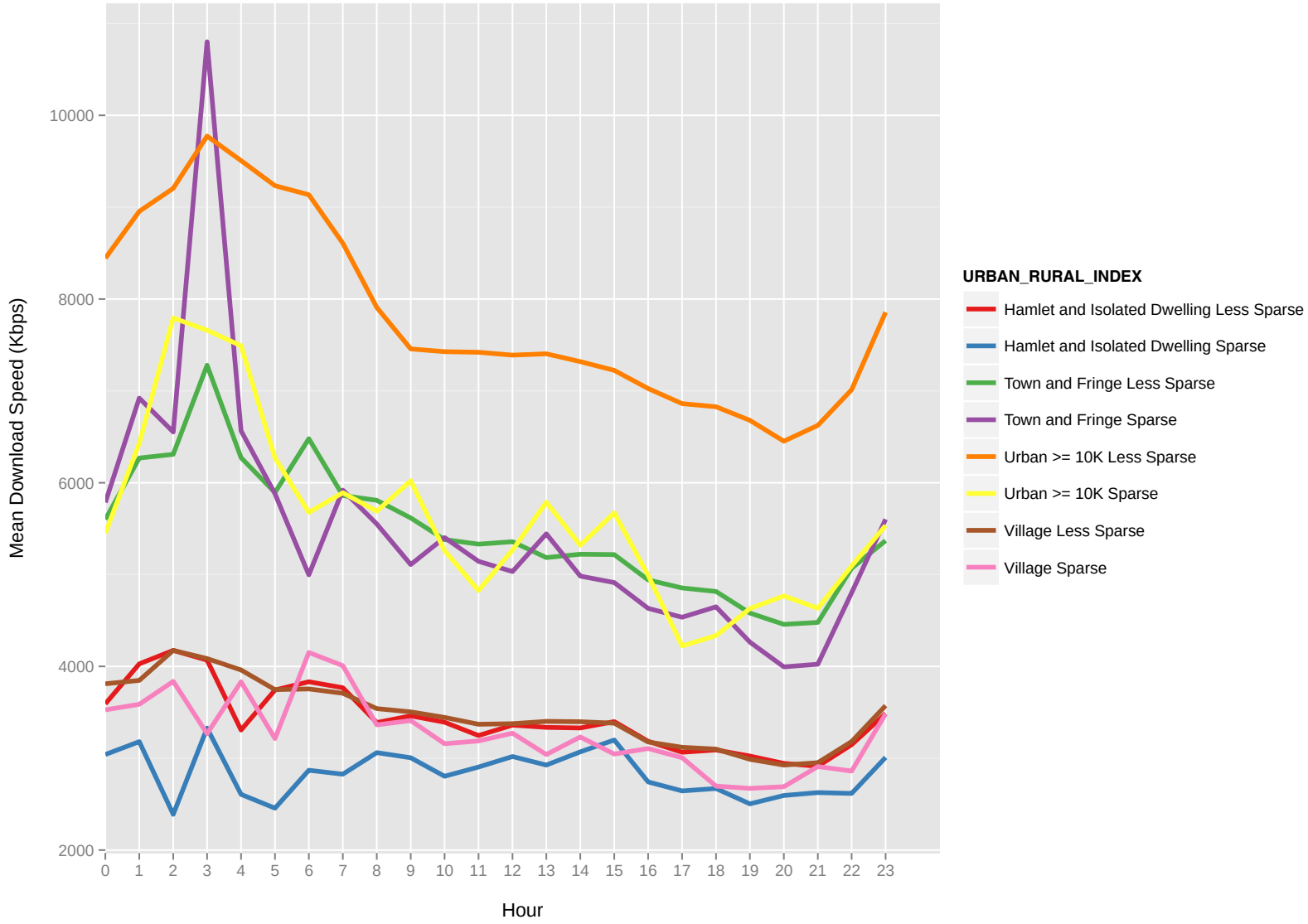

Superfast Connections Over 24Mbps 2010/11
Superfast Connections Over 24Mbps 2012/13

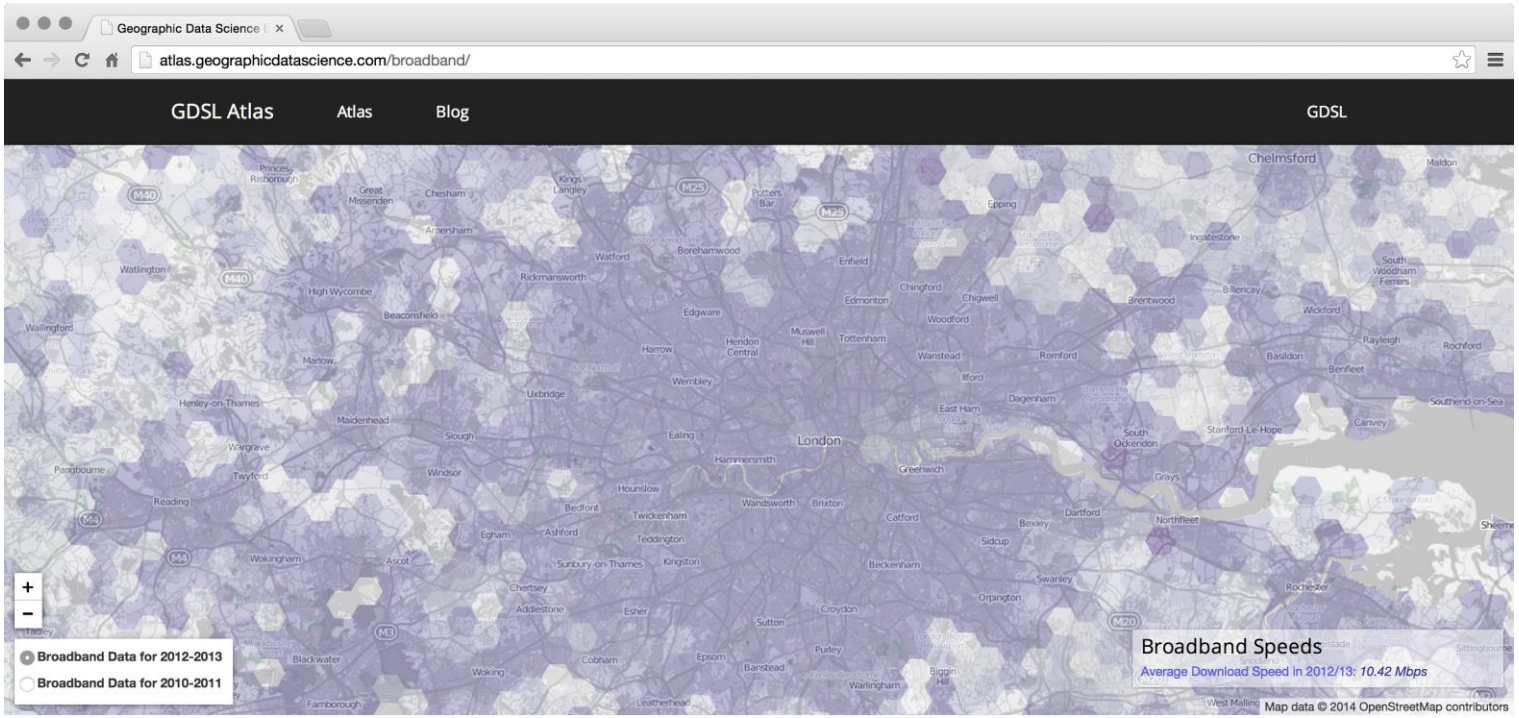
Who was supplying the data?



Preliminary 2011 Output Area Classification

Considering Spatio-temporal variation





Broadband Speed Map

Background

This map is the output of a short project working with Broadband Speedchecker who provide an online test of broadband upload and download speeds. Tests are run within a web browser by uploading and downloading of a small file of known size. The time taken to run the tests are recorded and and a megabytes per second score calculated and returned. You can run a test yourself below:

Map Authors

This GDSL map from the University of Liverpool was created by:

Liam Bratley - @LiamBratley

Dean Riddlesden - @deanriddlesden

Alex Singleton - @alexsingleton

Full UK Map

http://geographicdatascience.com/

2. Behaviors

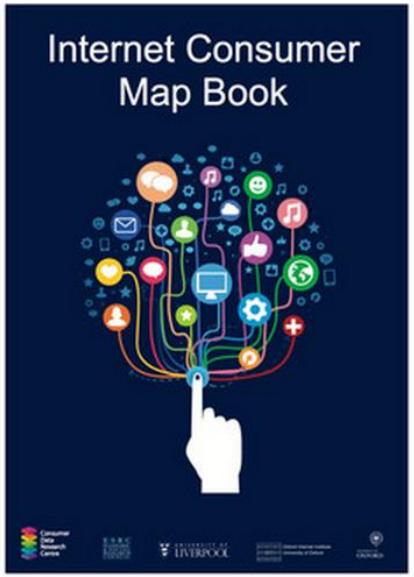
Key aims:

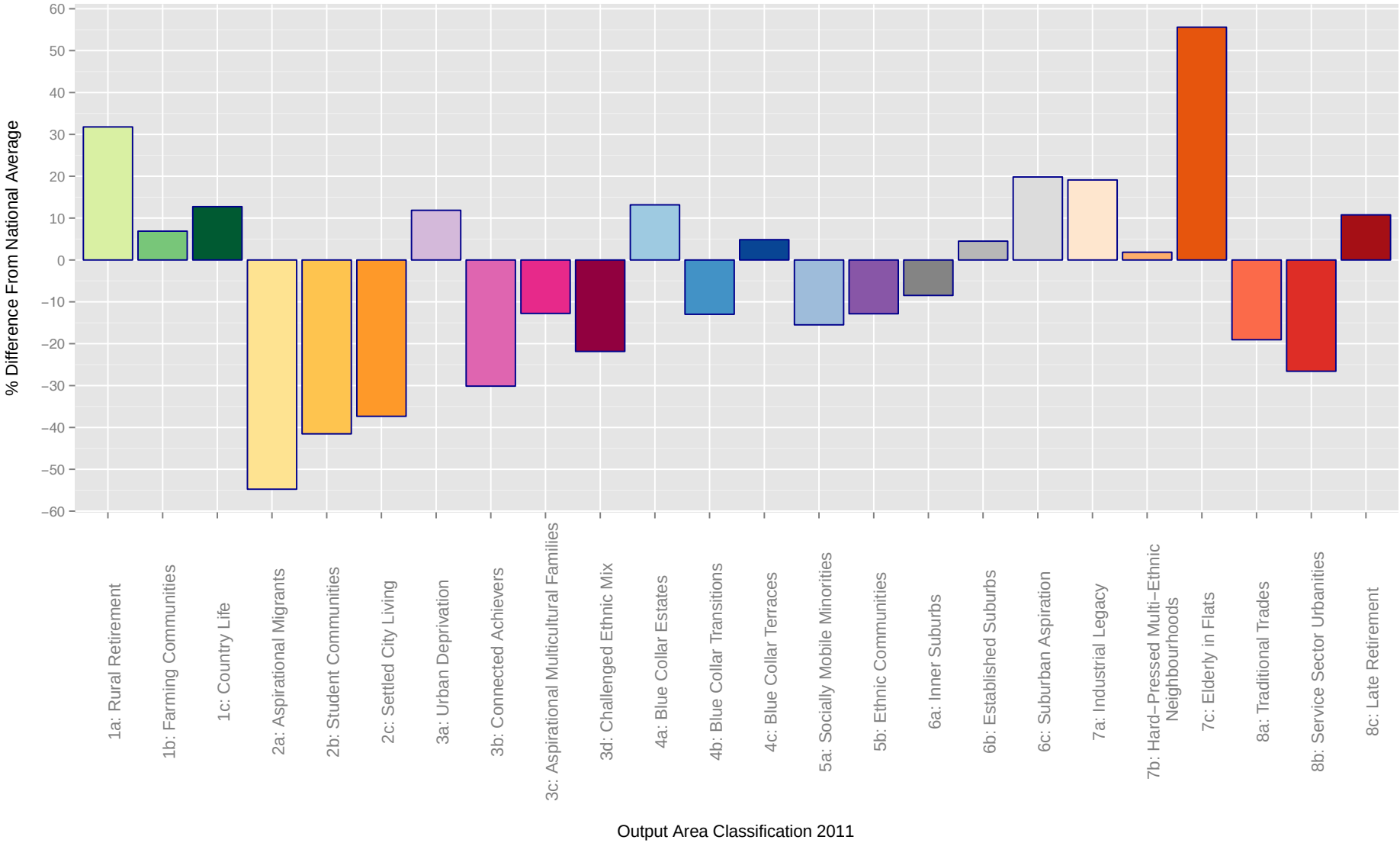
- What are the geographies of demand for various Internet applications?
- What can we learn about preferences and perceptions?
- Can we start to identify user groups?
- How do disparities in supply match with disparities in demand?
- How can we model this nationally?

MK41 8RG

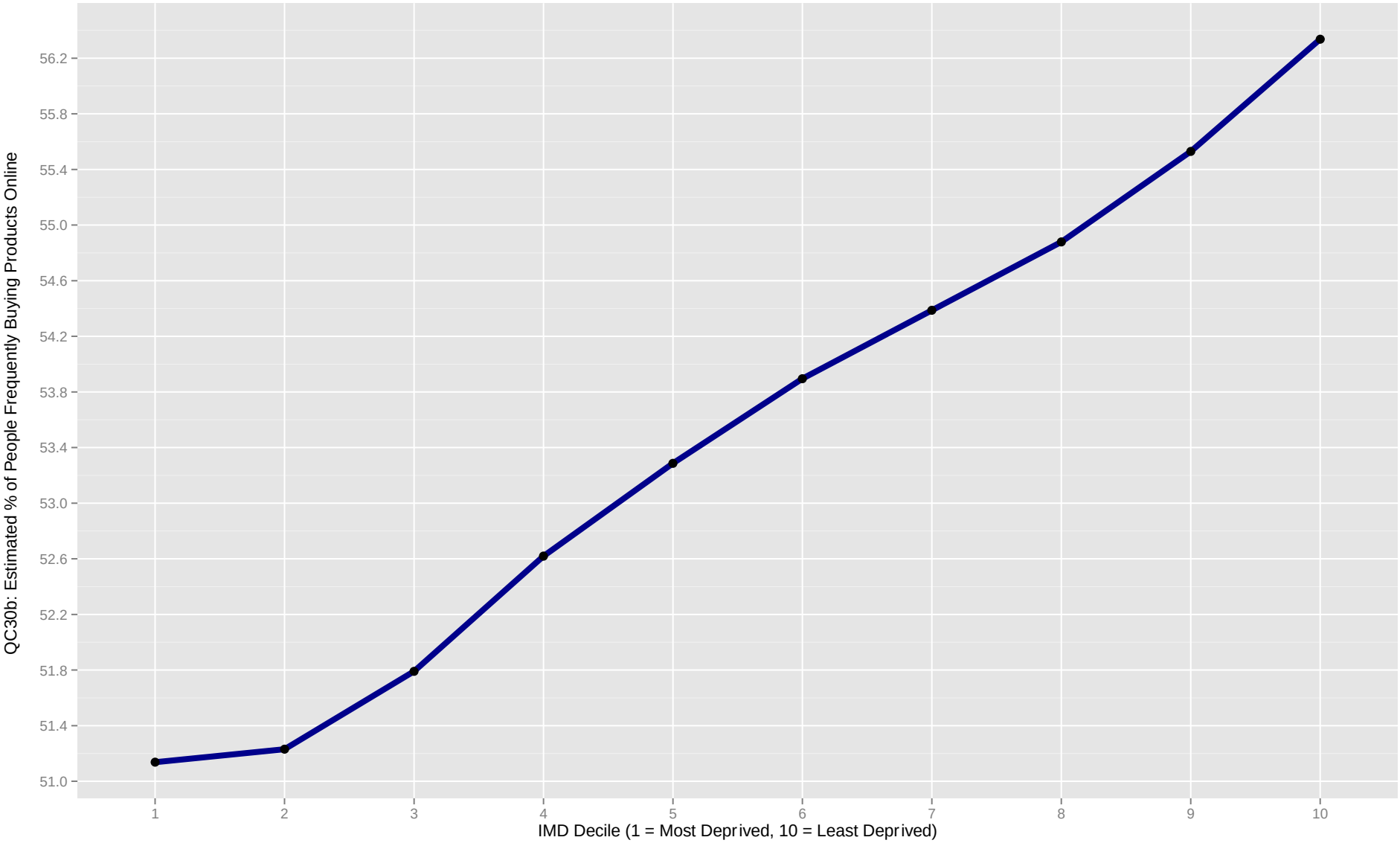
Behavioral Data

- Again, not easy to find
- ESRC Understanding Society – Issues
- OXIS / OII offered the most comprehensive survey data (vast dataset)
- Partnership between GDSL and OII established
- Data exchanged for the purpose of creating small area estimates of Internet engagement
- Output: Internet Consumer Map Book





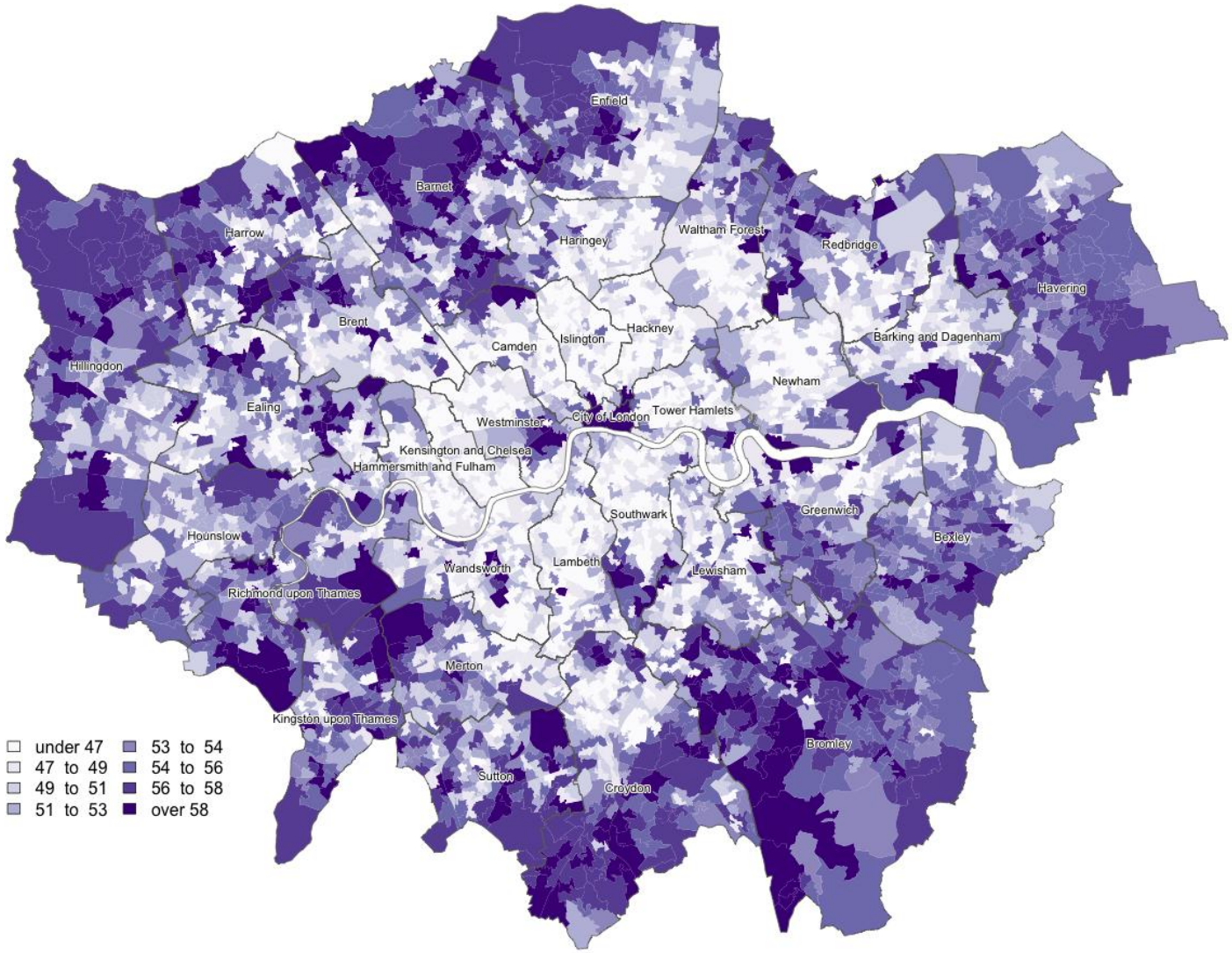
Internet non-use by OAC



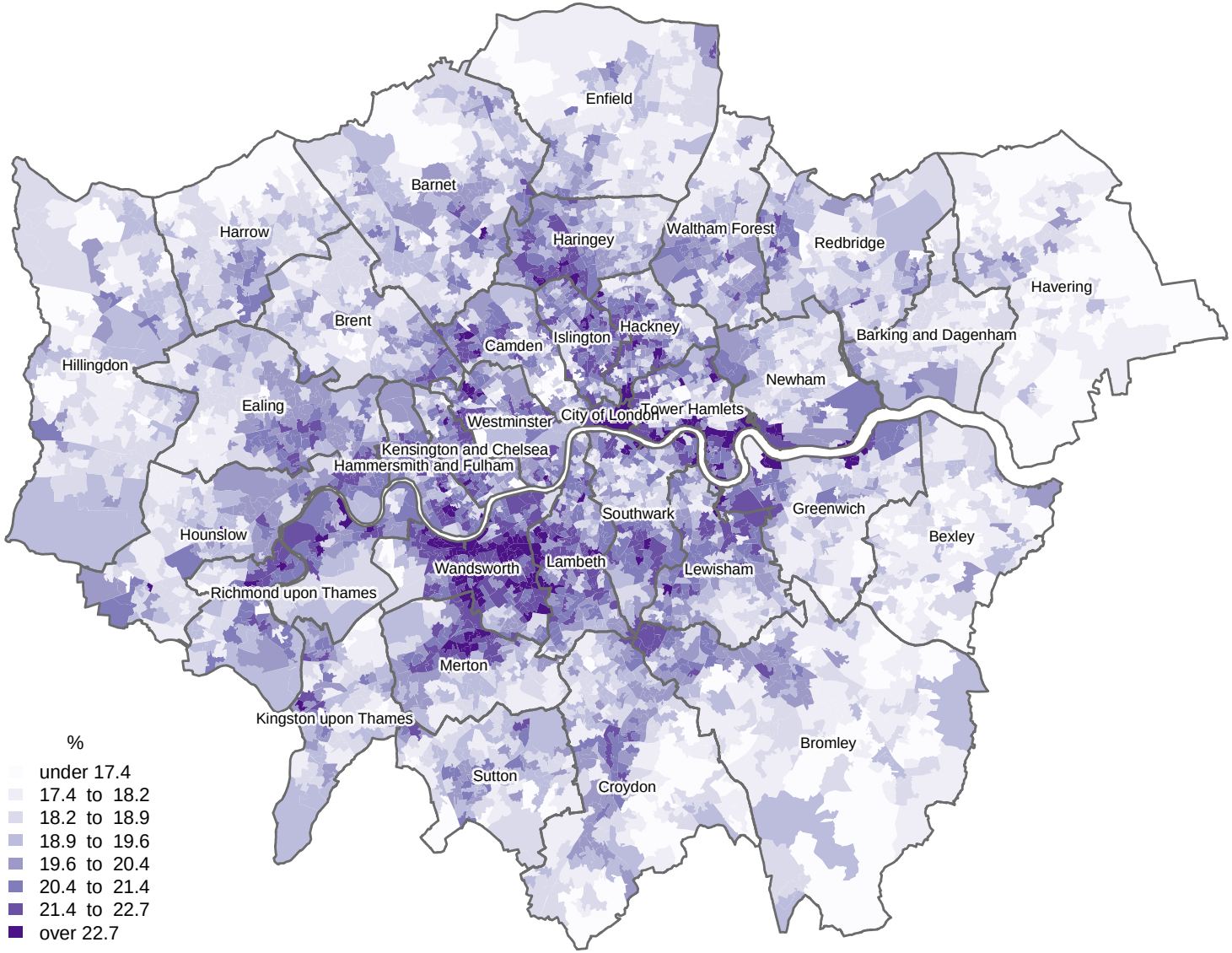
Frequent online shopping by Indices of Multiple Deprivation

UNIVERSITY OF LIVERPOOL

School of Environmental Sciences



% Frequently buying products online - London



% Frequently ordering food or groceries online - London

3. Demographic and Contextual Data

- 2011 Census data
- Area level statistics for multiple domains
- Assists in cluster formation and unpicking results
- Data selected based on previous research
- 31 contextual indicators:

Age
Level of qualification
Employment sector
Full time students
Population density

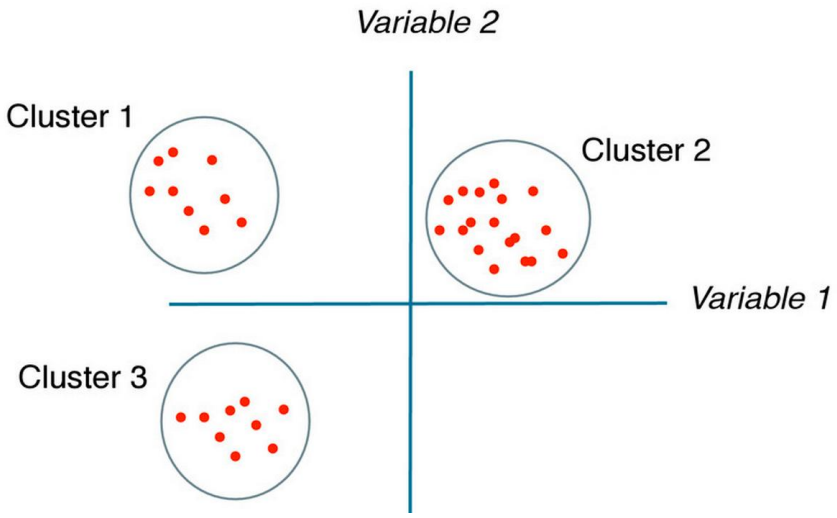
Building a geodemographic classification

Assessment of input variables:

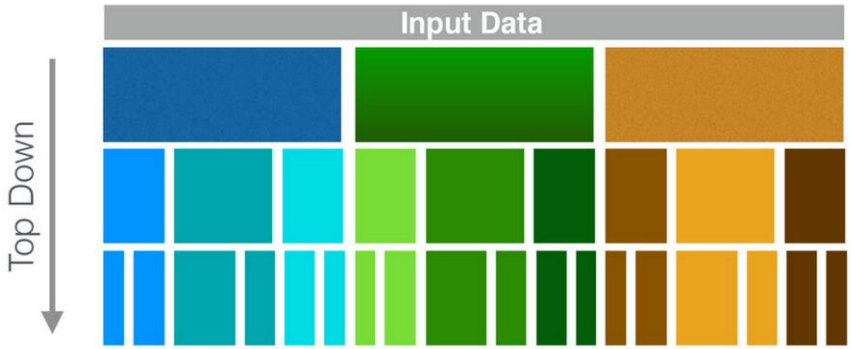
- Distributions
- Skew
- Normalize?
- Standardize

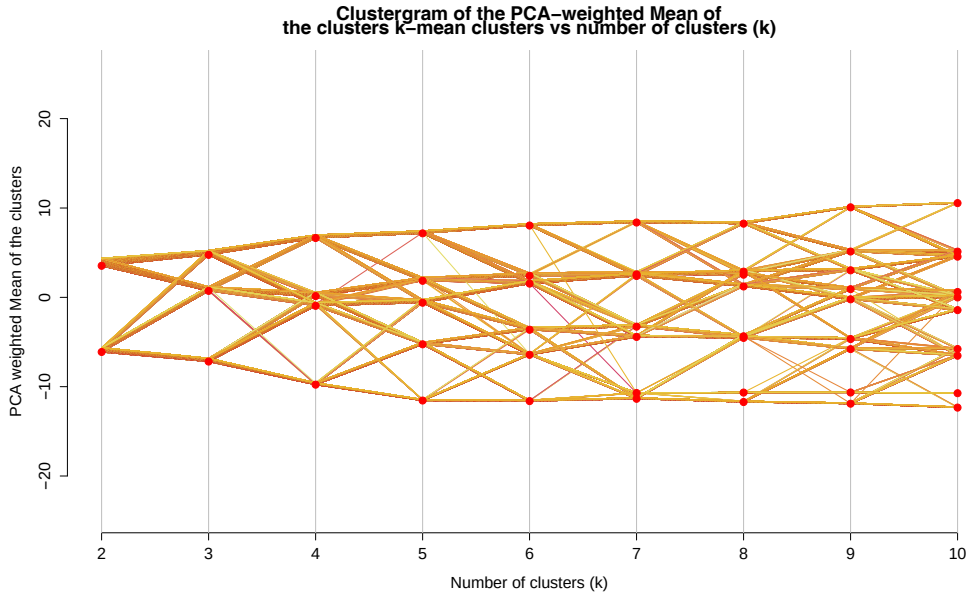
Cluster:

- K-means
- 10,000 iterations
- Supergroup and Group level hierarchy



Hierarchy





Supergroup	Group
1: e-unengaged	1a: e-unengaged elderley
	1b: e-marginals: not a necessity
	1c: e-marginals: the choose nots
2: e-professionals	2a: next generation users
	2b: e-experts
3: e-baseline	3a: e-mix
	3b: young and mobile
4: e-students	4a: e-students
5: e-rural and fringe	5a: e-fringe
	5b: constrained by infrastructure
	5c: low density high connectivity

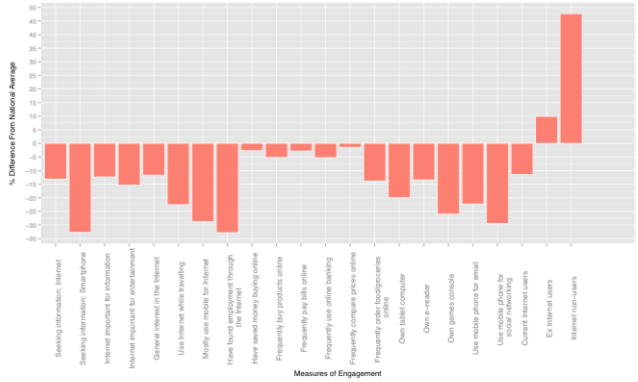
Internet User Classification (IUC)
User Guide

Dean Riddlesden
University of Liverpool
email: d.riddlesden@liverpool.ac.uk

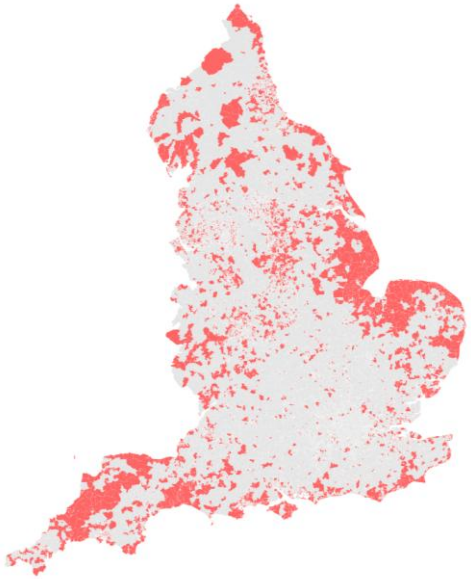
October 2014

3.1 Group 1a: e-unengaged elderly

The e-unengaged elderly group are characterised by large elderly populations who show little or no engagement with the Internet across all applications. Age structure within this group is heavily skewed towards the elderly with the highest proportion of residents aged 75 plus of any group in the IUC. As a result, Internet enabled device ownership is lower than the supergroup average and the lowest of any group in the IUC. Internet non-use is higher than the supergroup average and far above the national average. Geographically this group tends to cluster around coastal and lower density rural areas that attract elderly populations. Infrastructure provision and performance is typically slightly below the national average. The e-unengaged elderly group accounts for 4% of all Lower Super Output Areas nationally.

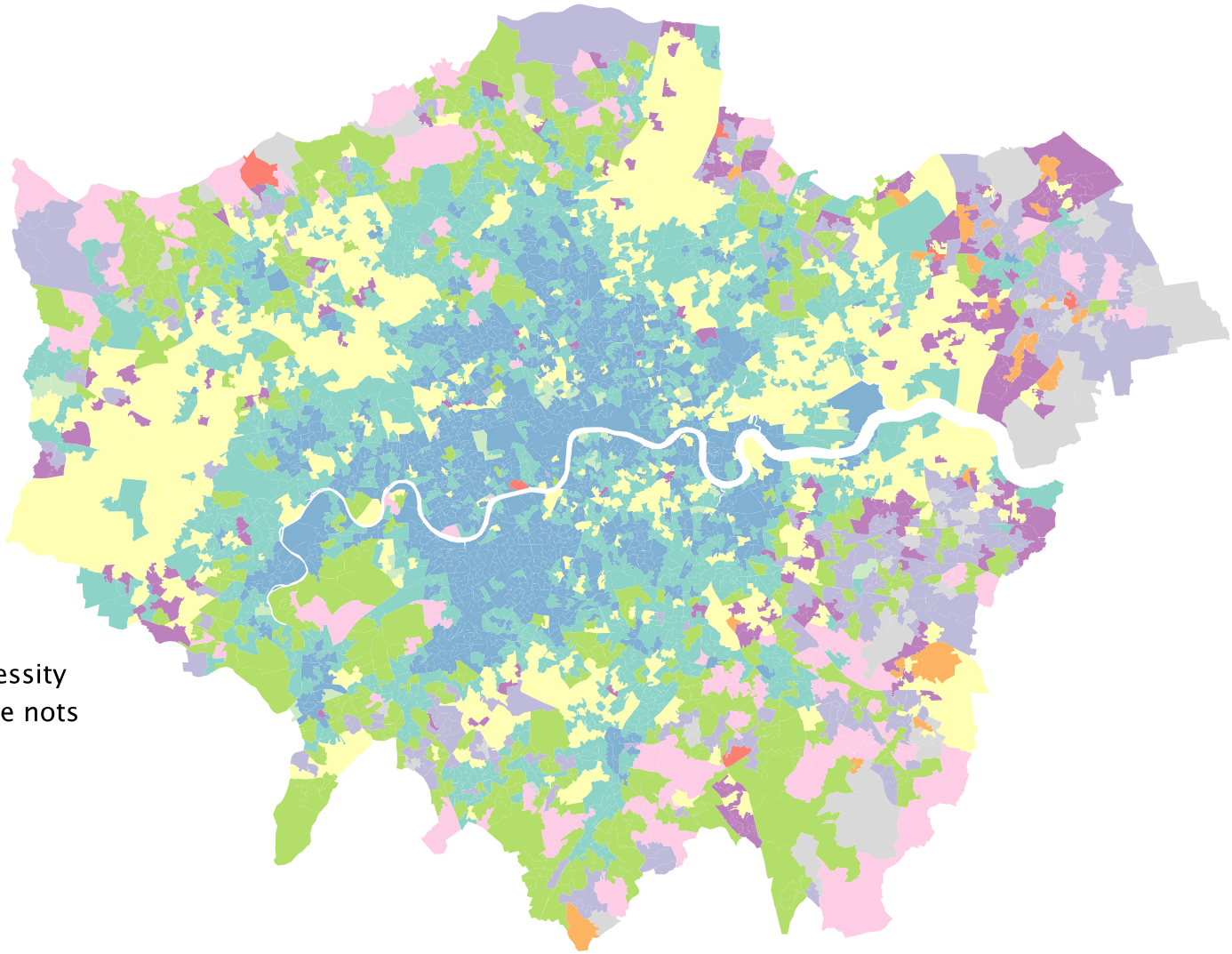


Group 1a: e-unengaged elderly: Engagement Characteristics



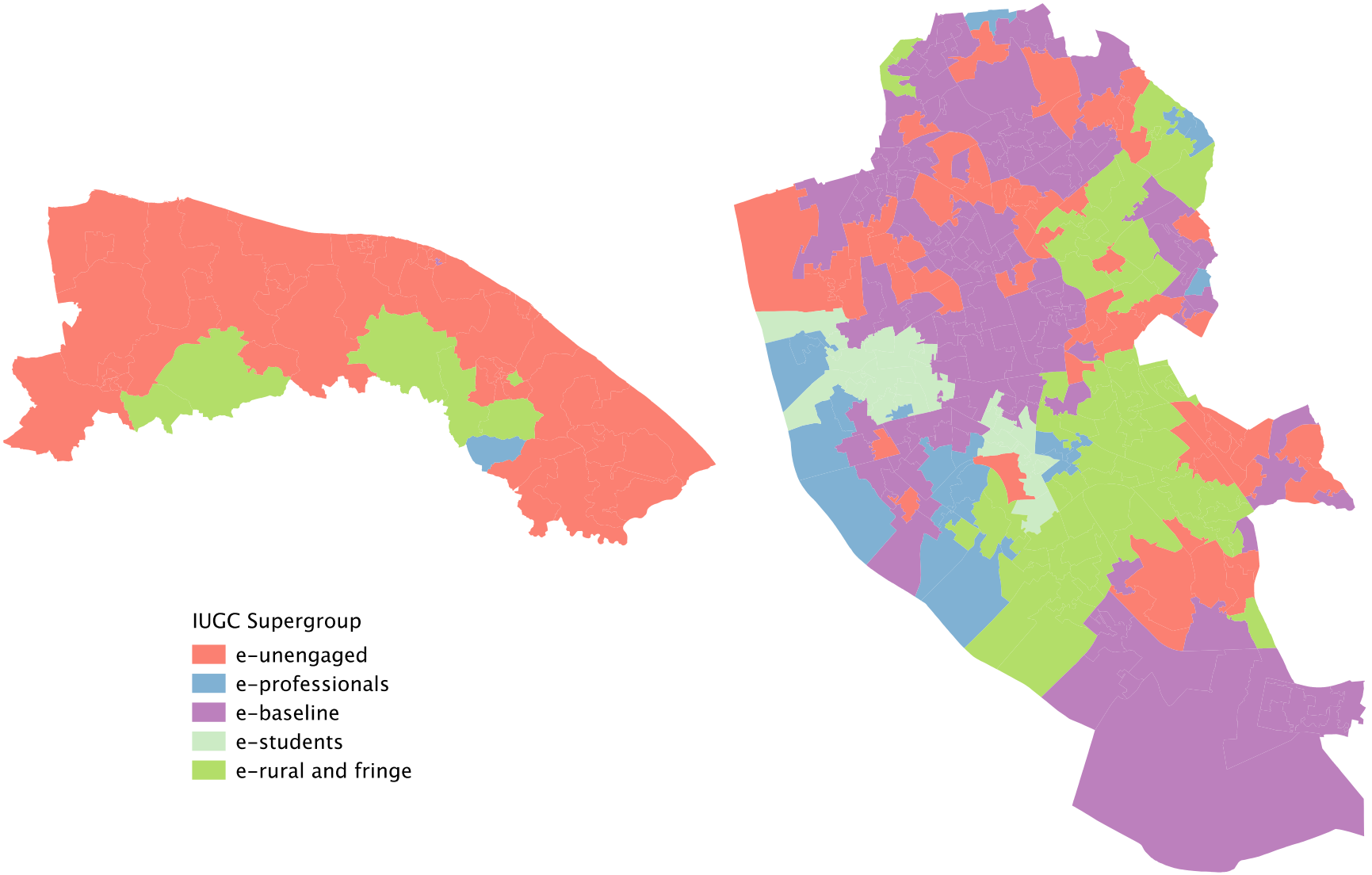
UNIVERSITY OF LIVERPOOL

School of Environmental Sciences



UNIVERSITY OF LIVERPOOL

School of Environmental Sciences


IUGC Supergroup
e-unengaged
e-professionals
e-baseline
e-students
e-rural and fringe

Conclusions

- Geodemographic classifications can be a powerful tool if built carefully
- Input data is crucially important – E-society failures
- Data acquisition is time consuming and difficult
- Interesting patterns of user groups at national, regional and sub-regional levels
- Wide variety of potential uses in further research/ policy/ industry
- IUC success will depend on uptake – time will tell

Questions?