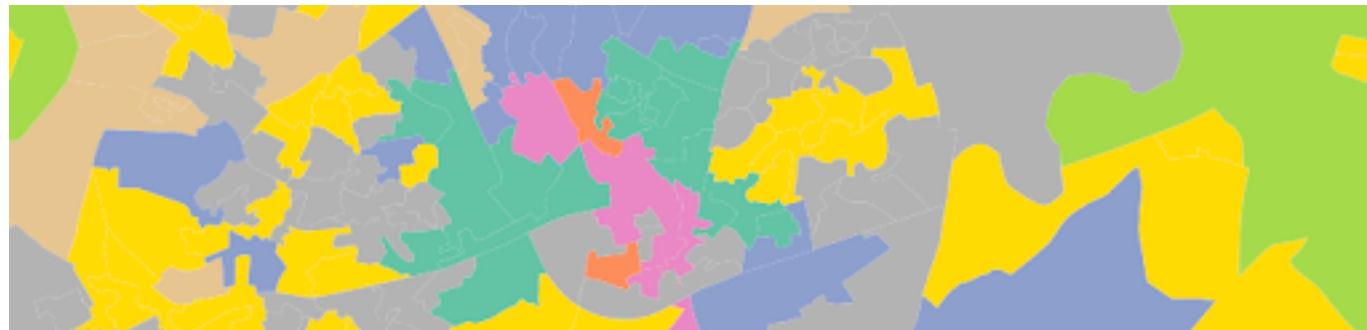


Principles of Spatial Analysis

SHORT LECTURE 01, WEEK 09: GEODEMOGRAPHICS



geodemographics

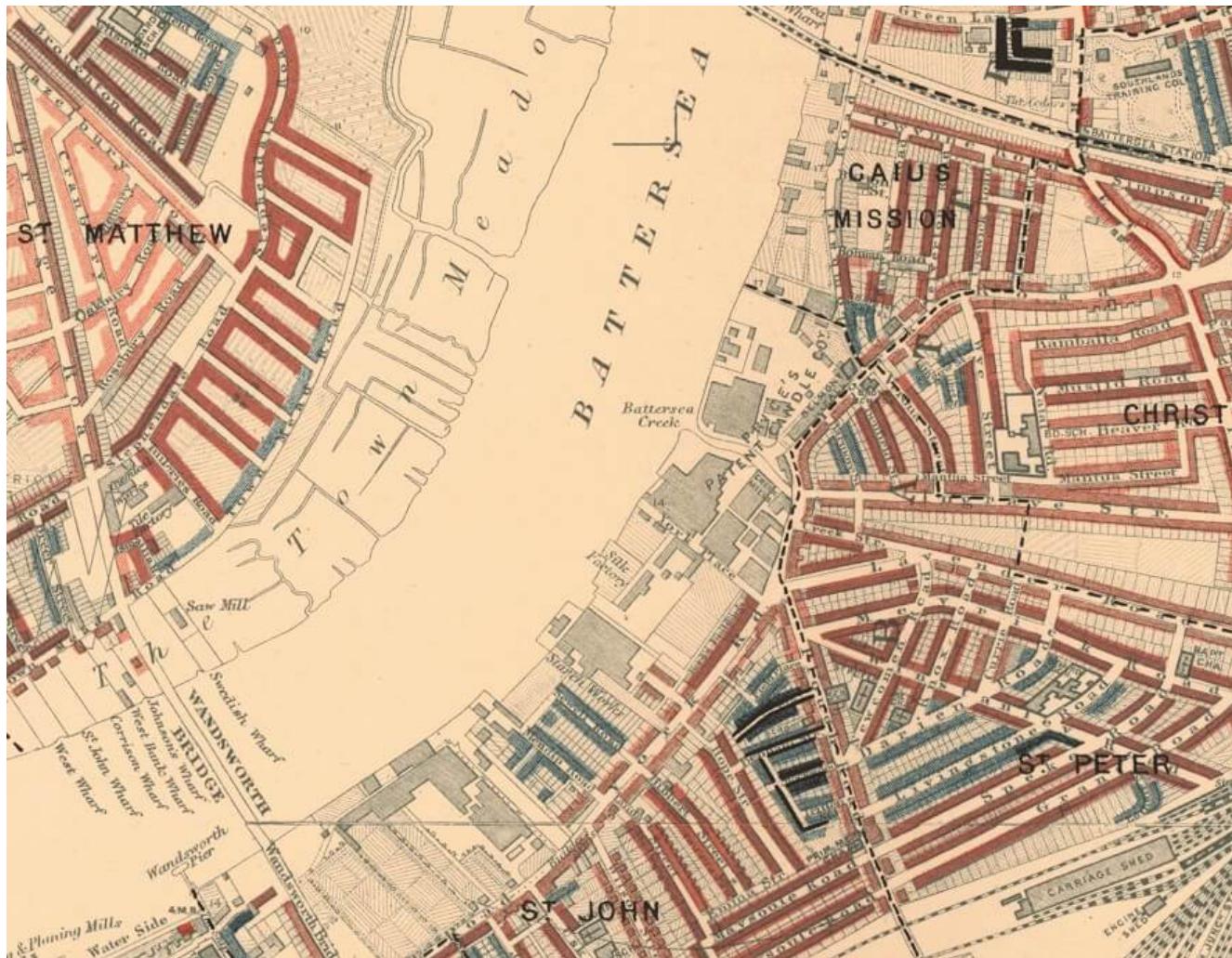
- analysis of people (*demographics*) by where they live (*geo*)
- used to identify similar neighbourhoods or administrative areas
- means of multivariate data reduction for the differentiation of areas
- been around for many many years, dating back to the late 1800s

Charles Booth

- created the first geodemographic-style classification
- shipping business owner and philanthropist
- survey: **Life and Labour of People in London**
- mostly qualitative analysis by walking through areas
- books and books and books with notes

Charles Booth

- noticed that there is a geographical pattern in the distribution of different social categories: people who live in a particular neighbourhood and share similar living conditions, are of similar characteristics and social status



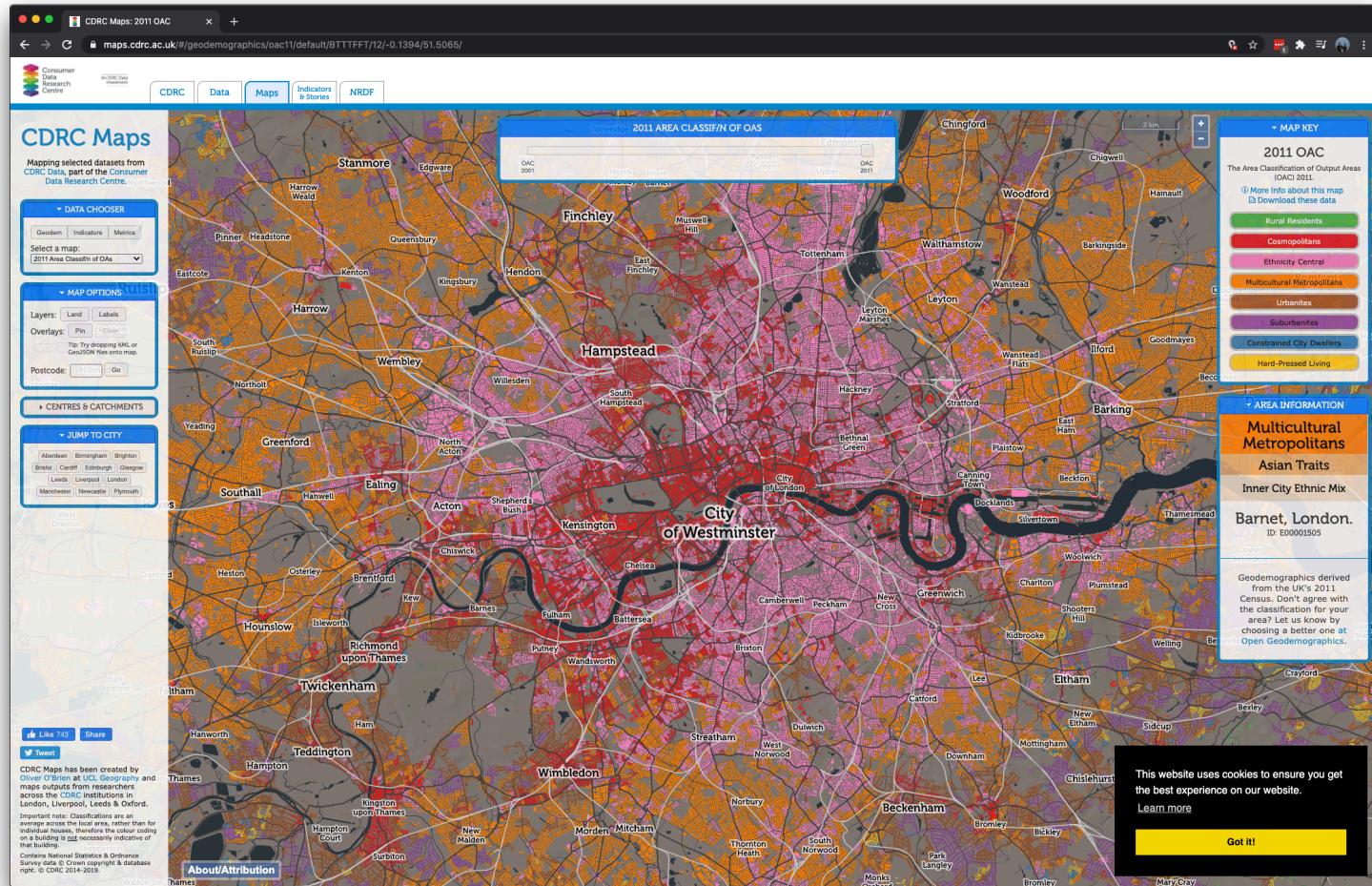
Maps Descriptive of London Poverty 1889. Charles Booth's *Inquiry into Life and Labour in London* (1886-1903).

Classification	Colour	
Lowest class. Vicious, semi-criminal.	Black	
Very poor, casual. Chronic want.	Dark blue	
Poor. 18s. to 21s. a week for a moderate family.	Light blue	
Mixed. Some comfortable others poor.	Purple	
Fairly comfortable. Good ordinary earnings.	Pink	
Middle class. Well-to-do.	Red	
Upper-middle and upper classes. Wealthy.	Yellow	

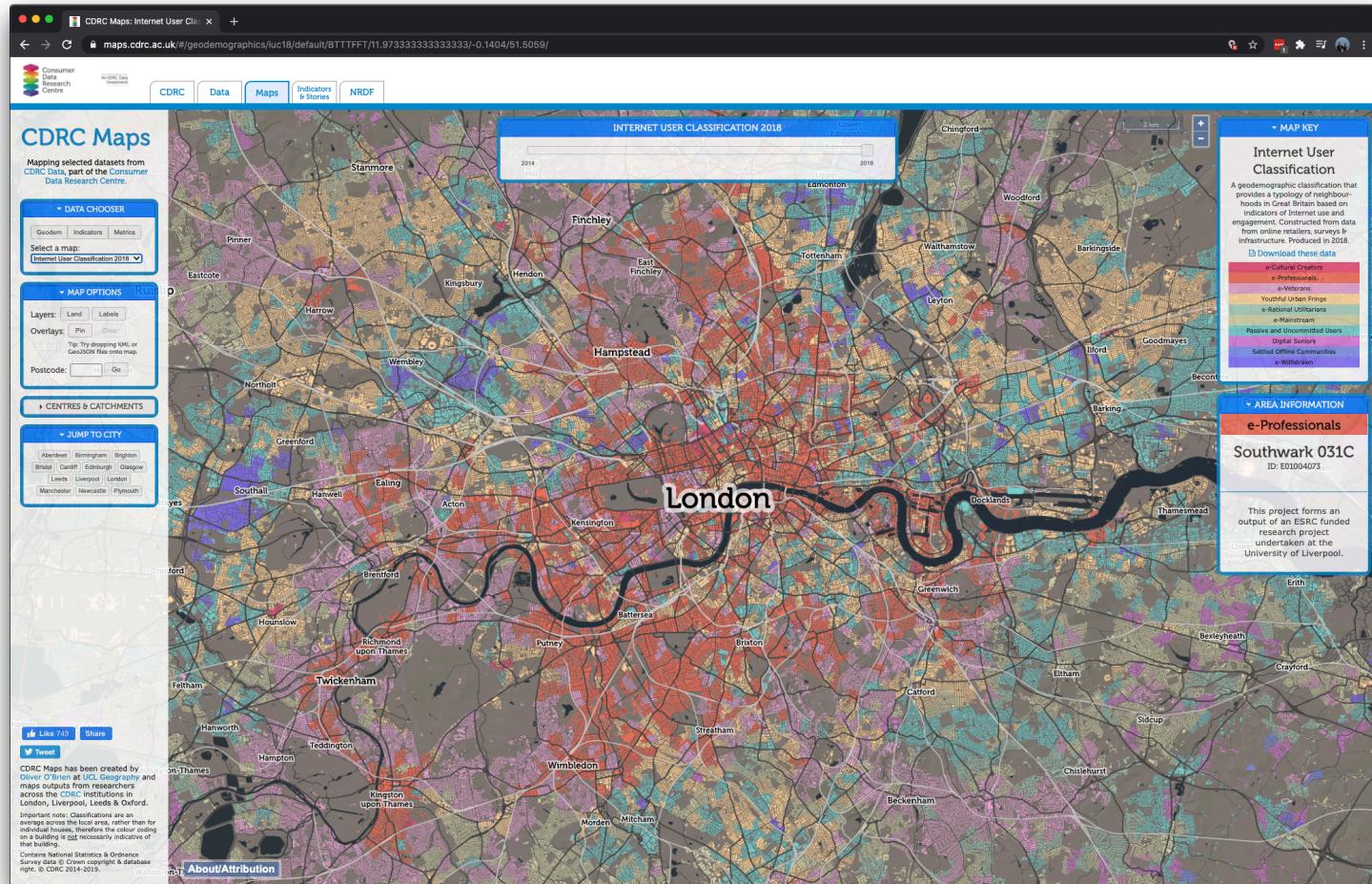
Maps Descriptive of London Poverty 1889. Charles Booth's *Inquiry into Life and Labour in London* (1886-1903).

geodemographics

- further developed in 1970s to target urban deprivation funding
- commercial sector also got involved (CACI ACORN / Experian MOSAIC)
- Office for National Statistics' Output Area Classification (2001, 2011, 2021)
- ONS' Output Area Classification completely open using Census data
- all classifications have a similar structure, typically hierarchical



2011 OAC on maps.cdrc.ac.uk



Internet User Classification on maps.cdrc.ac.uk

geodemographic classification

"(...) is created by **assembling a wide range of measures** that describe the characteristics of areas and/or those people living within them, and then, through the implementation of unsupervised learning (clustering), **identifies groups of areas that share common characteristics**. Emerging clusters may be divided or aggregated to create a hierarchy, and it is typical that these be accompanied by labels, descriptions, photographs, diagrams and graphs"

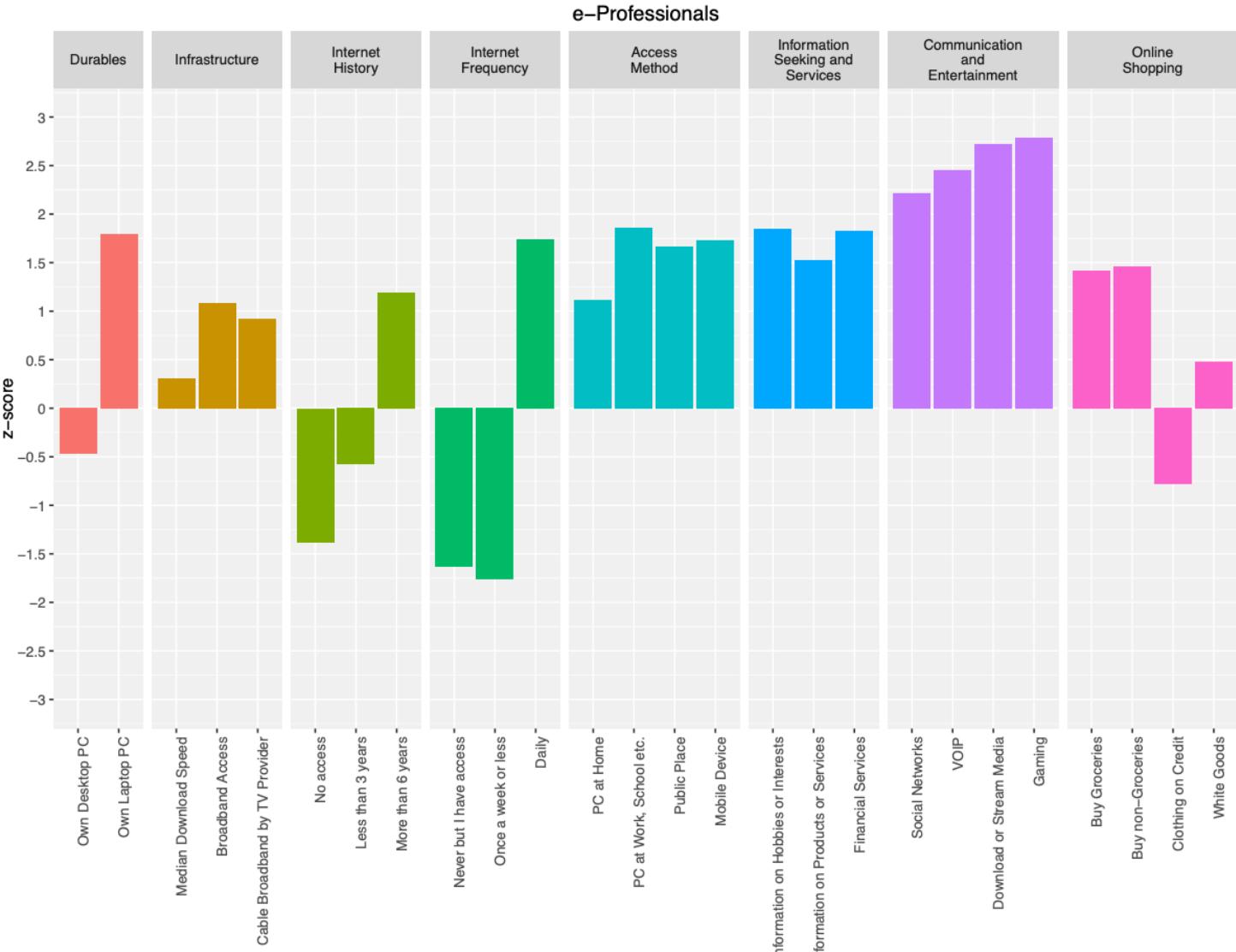
(Singleton *et al.* 2020)

Internet User Classification

- bespoke classification by CDRC researchers
- how do populations interact with the internet
- 'profiles of internet use and engagement'
- built from a range of consumer data, survey data, and open data
- open data available through CDRC
- Singleton *et al.* 2020

Internet User Classification

- British Population Survey: internet access, frequency of internet usage, access to PC, type of internet use
- transactional (consumer) data on online shopping
- average broadband speed
- census variables such as age, ethnicity
- National Statistics Socio-economic classification (NS-SEC)



Internet User Classification mean attributes of the *e-Professionals*

e-Professionals

"This Group has high levels of Internet engagement, particularly regarding social networks, communication, streaming and gaming, but relatively low levels of online shopping, besides groceries. They are new but very active users, with a very high proportion of the population engaging on a daily basis. (...)

Geographically, this Group is mainly located close to the city centre or within the proximity of Higher Education Institutes, where infrastructure accessibility, such as cable broadband, is sufficient"

Internet **User** Classification

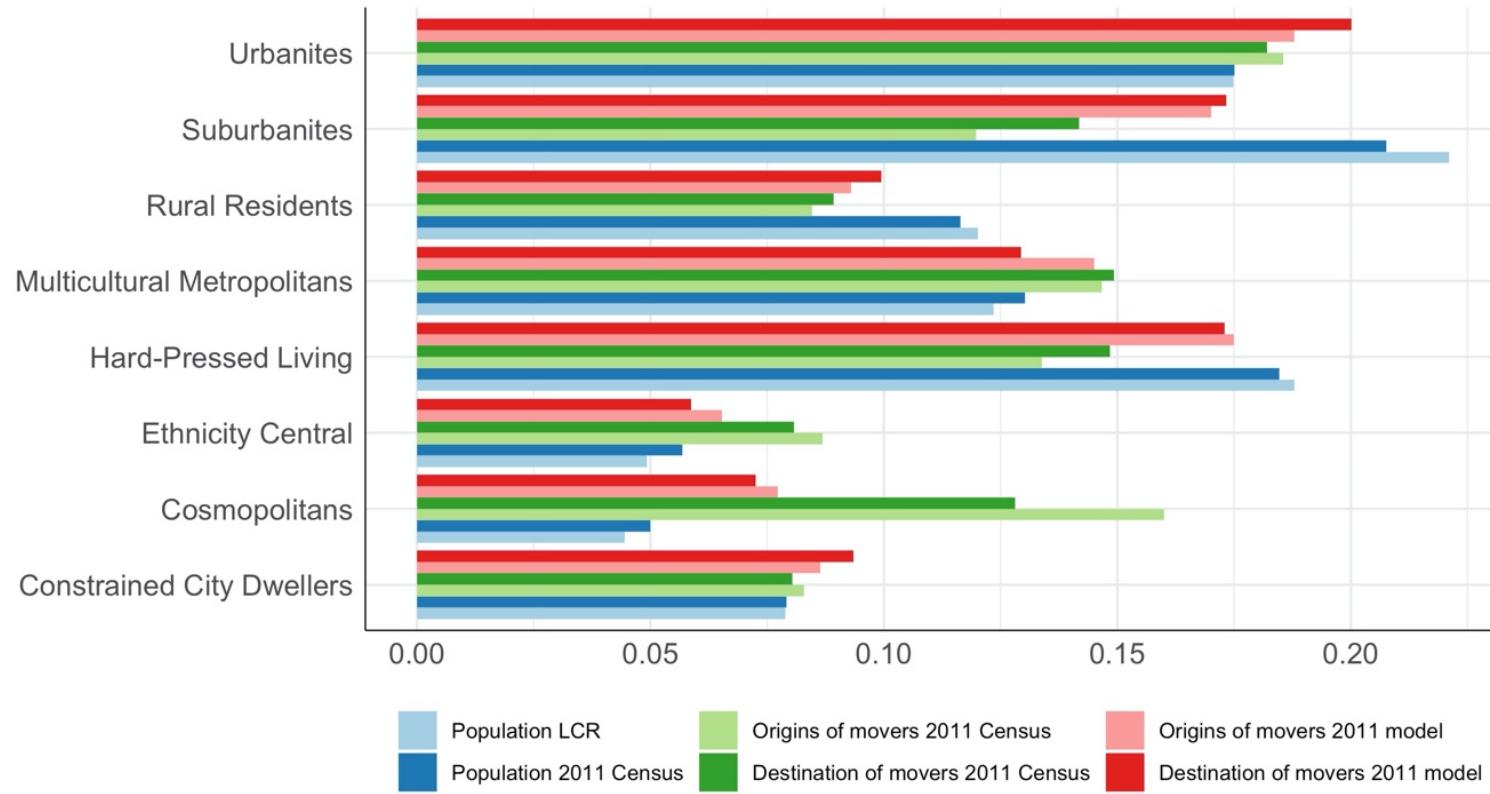
- measures of access to and use of internet
- identification of areas to target potential interventions
- analysis of areas where people are likely to work from home (COVID!)

limitations of geodemographics

- highly dependent on the input data (complete data necessary!)
- input data can get old very quickly (depending on the topic)
- inherent biases within the input data

further applications of geodemographics

- using the geodemographic classification as input for further analysis
- Harris *et al.* 2007: differences in school choice between social groups
- Brunsdon *et al.* 2011: participation in higher education
- Martin *et al.* 2018: analysis of travel to work flows
- Goodman *et al.* 2011: socio-economic inequalities in exposure to air pollution



let's put it into practice