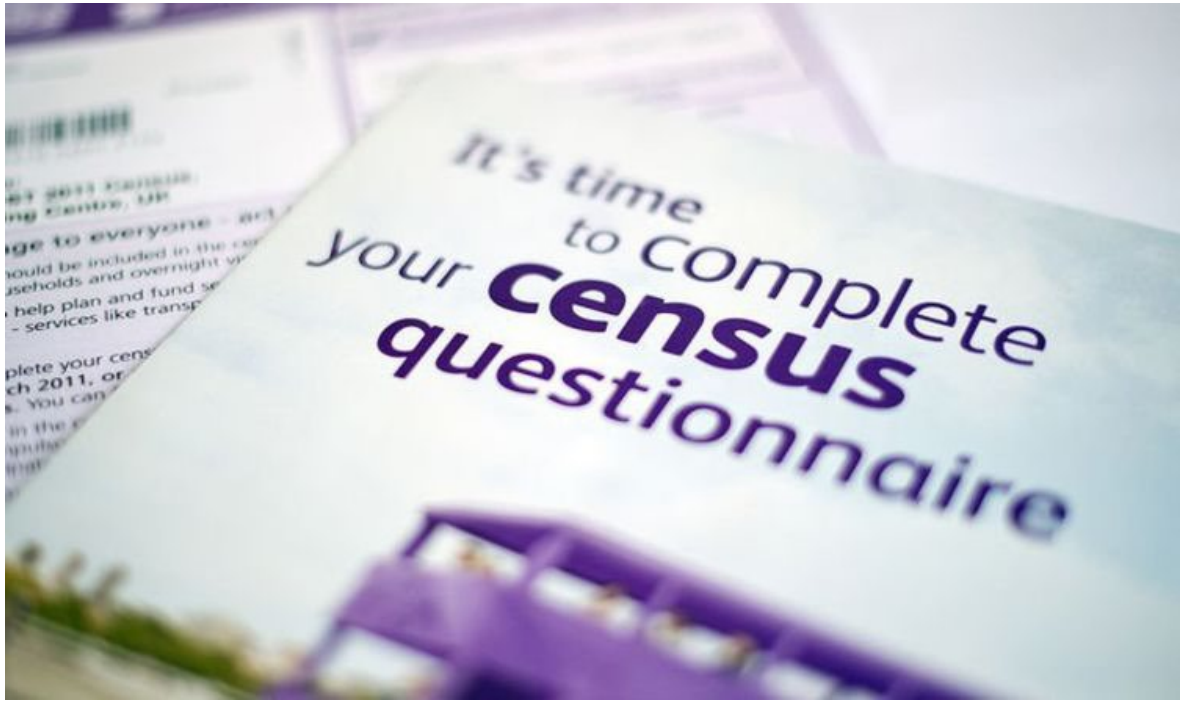
UK Data Service

Census 2021

Planning and designing your research



Nigel de Noronha, UK Data Service

19 July 2022

UKRI Economic and Social Research Council

Copyright © year Institution. Created by Organisation, UK Data Service.

What we will cover

- Research focus
- Geographical scale
- Developing your analytical skills
- Some more advanced techniques

UK Data Service

Research focus

Units of analysis

- People (individual)
- Households
- Places

ONS Planned census data release

1

- Population estimates by age and sex at local authority level
- Topic summaries
- Area profiles

28th June 2022 with topic summaries and area profiles following in autumn winter 2022.

(NI first release 24th May 2022 with slower schedule for release. Scotland census a year later)

2

- Multivariate data
- Flexible table builder (ONS site)
- Small populations (short-term residents)

Winter 2022 to early 2023

3

- Alternative population bases
- Small populations
- Flow data
- Microdata

From spring 2023

UK wide census release to follow all data being available

Phase 1 – topic summaries

Mainly univariate data released every two to three weeks throughout autumn and winter 2022

- demography and migration
- ethnic group, national identity, language, and religion
- health, disability, and unpaid care
- housing
- labour market and travel to work
- sexual orientation and gender identity
- education
- UK armed forces veterans

Geographical data

Boundary data		
Administrative e.g. Local authority, Health	Electoral e.g. constituency, ward	Statistical areas built on output areas

Will provide administrative and electoral boundaries at time of release, output areas in July 2022, lower-level super output areas in August 2022 and mid-level super output areas in September 2022

Output area geography

Output area minimum 40 households and 100 residents, average target 125 households, maximum 250 households and 625 people

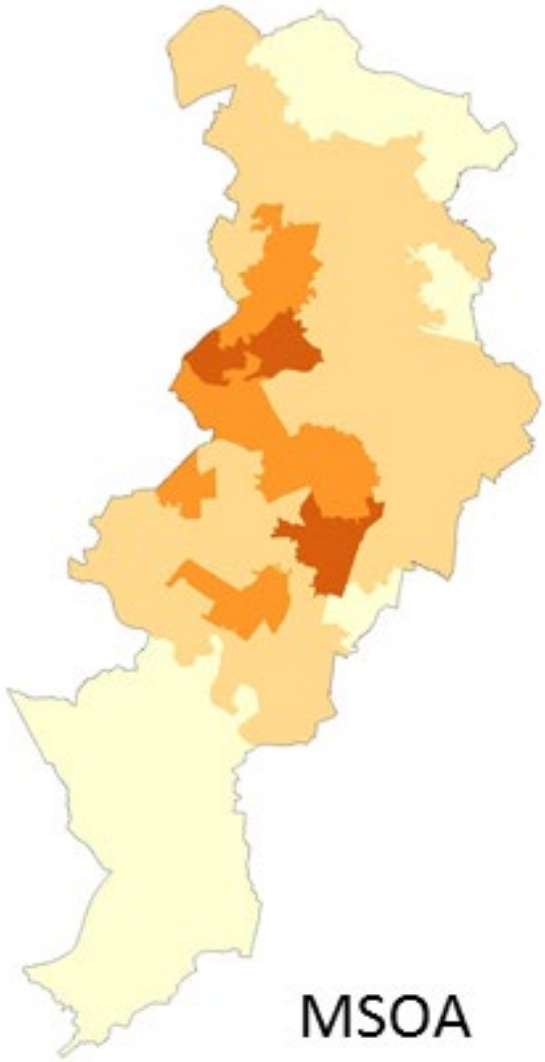
- introduced in 2001
- aim for more homogeneous population through matching characteristics of population
- but try to minimise changes

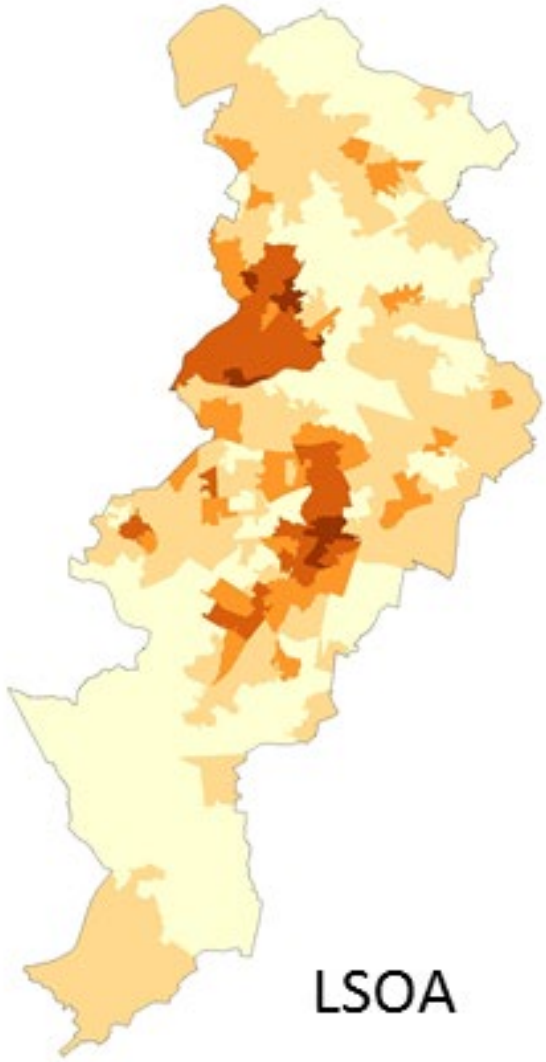
Built into larger super output areas

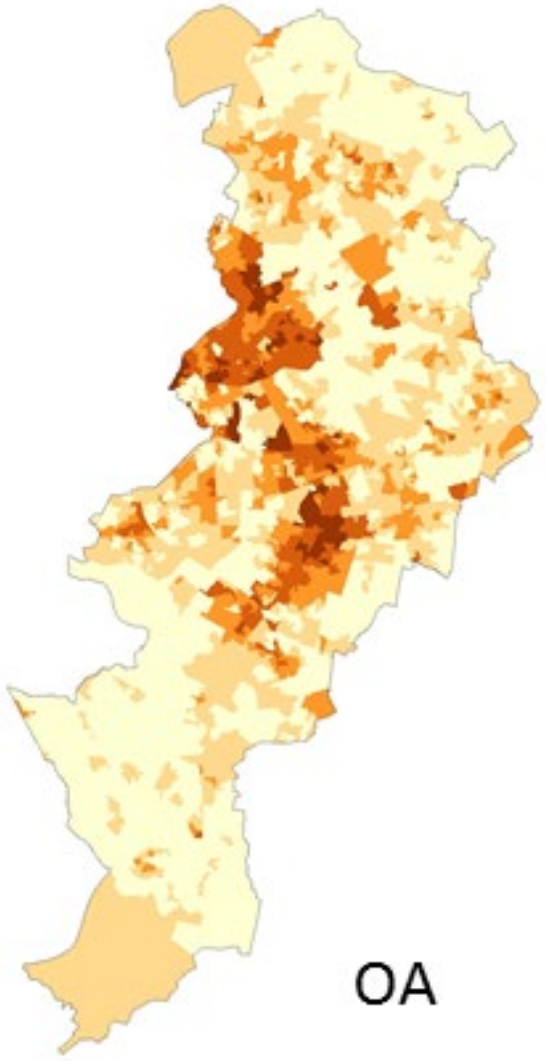
- lower level super-output areas (LSOAs) minimum 400 households and 1,000 people, maximum 1,200 households and 3,000 people – used in published statistics e.g. IMD, recorded crime
- Mid-level super output areas (MSOAs) minimum 2,000 households and 5,000 people, maximum 6,000 households and 15,000 people – used in published statistics e.g. educational attainment, Covid cases

All boundaries constrained within local authorities

Output area specificity







Private rented sector in Manchester
Source: 2011 census

Phase 2 (winter 2022 to early 2023)

Short-term population (intend to stay in the UK less than 12 months)

Defined sets of tables including two or more variables at different geographical scales through ONS interface

Flexible table builder (incorporating statistical disclosure control)

The variable categories will vary between univariate and multivariate tables based on statistical disclosure control. This is likely to lead researchers to need to balance geographical scale and the level of detail available.

Breakout session

- *What aspect(/s) of peoples' lives are you interested in?*
- *What area(/s) are you interested in and at what scale?*

UK Data Service

Skills analysis

Helping to define the skills you need for your research

Skills matrix

Task	Skills needed
Identifying sources of data	Excel and understanding of variables
Downloading data	Web interface
Preparing data	Excel (formulae)
Analysing data	Excel / statistical software
Presenting data	Excel for tables and charts
	GIS for mapping
Spatial analysis	GeoDA (free software)
	R, STATA or GIS
Multivariate analysis	Statistical software
Combining other data	Data Management skills in Excel or statistical package

UK Data Service

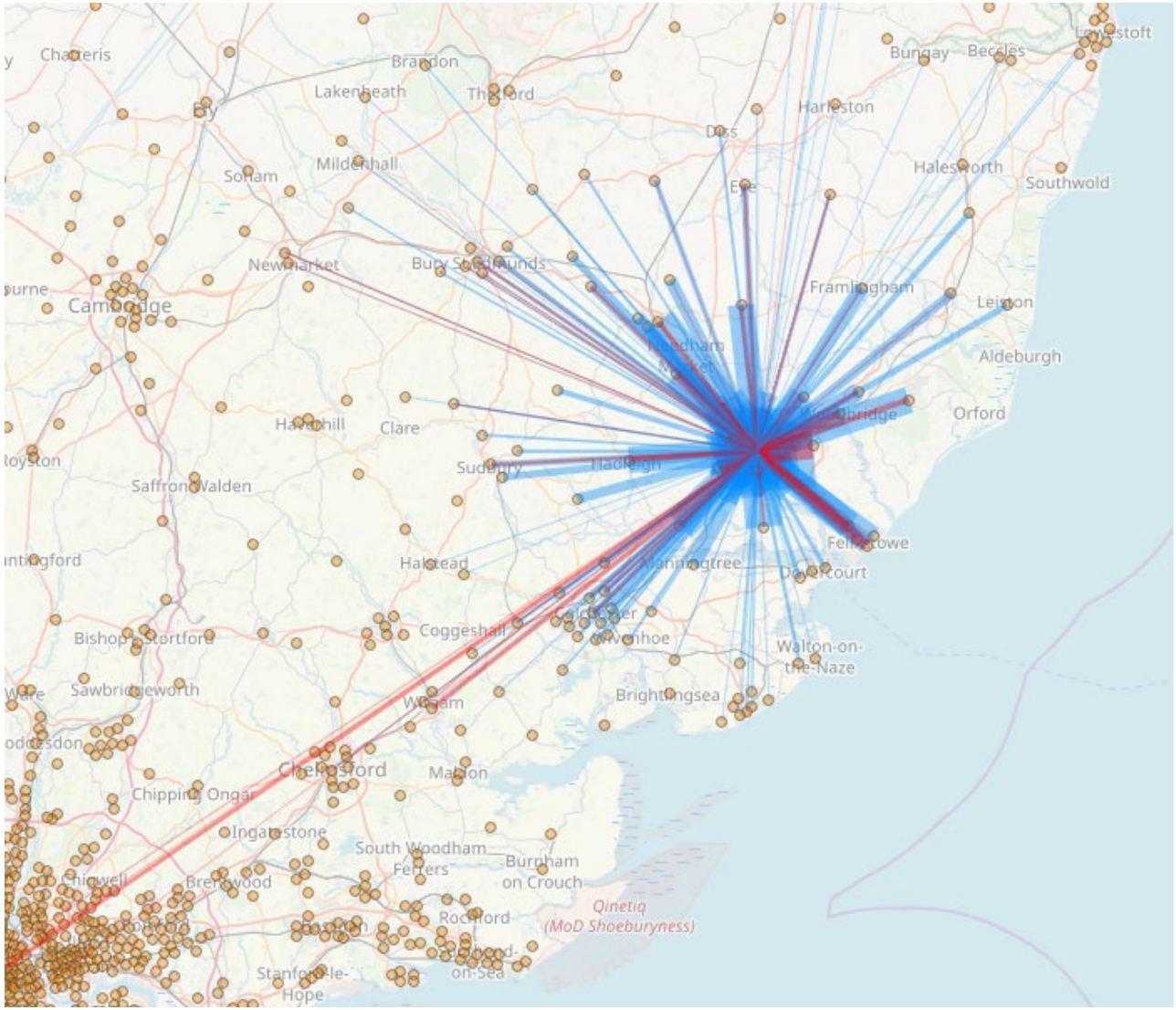
… and there is more

Phase 3 releases and some potential applications

Phase 3 (from spring 2023)

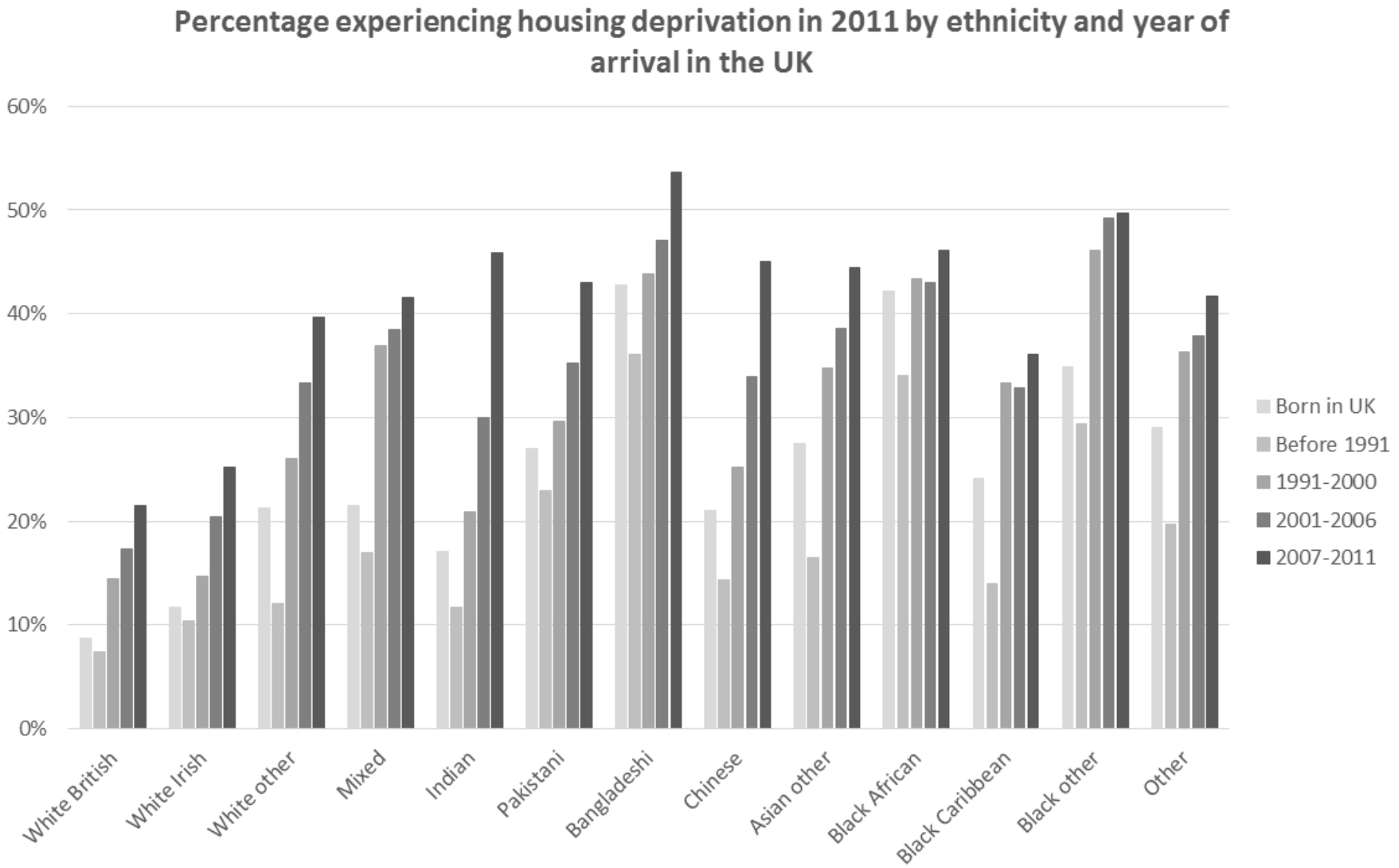
- Alternative population bases (Workplace, Workday, Out-of-term, Second address
- Small populations
- Detailed migration data
- Flow data
- Microdata

Origin destination maps



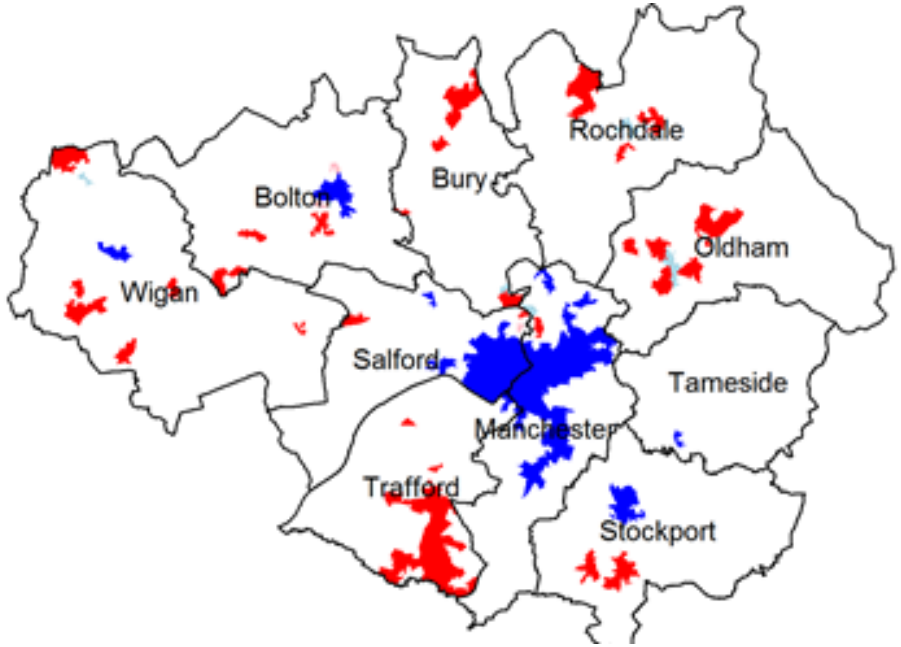
Commuting to Ipswich
Source: Healthy Suffolk website

Multivariate analysis

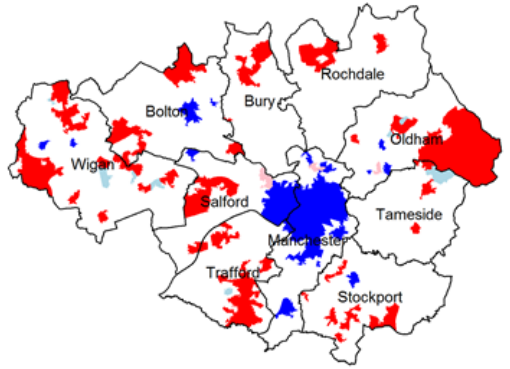


Source: census microdata 2011

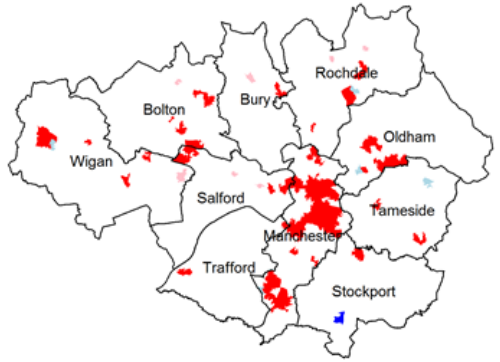
Spatial clusters (couples with dependent children)



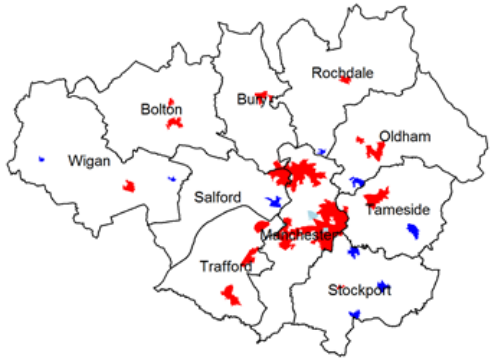
Owned



Social Rented

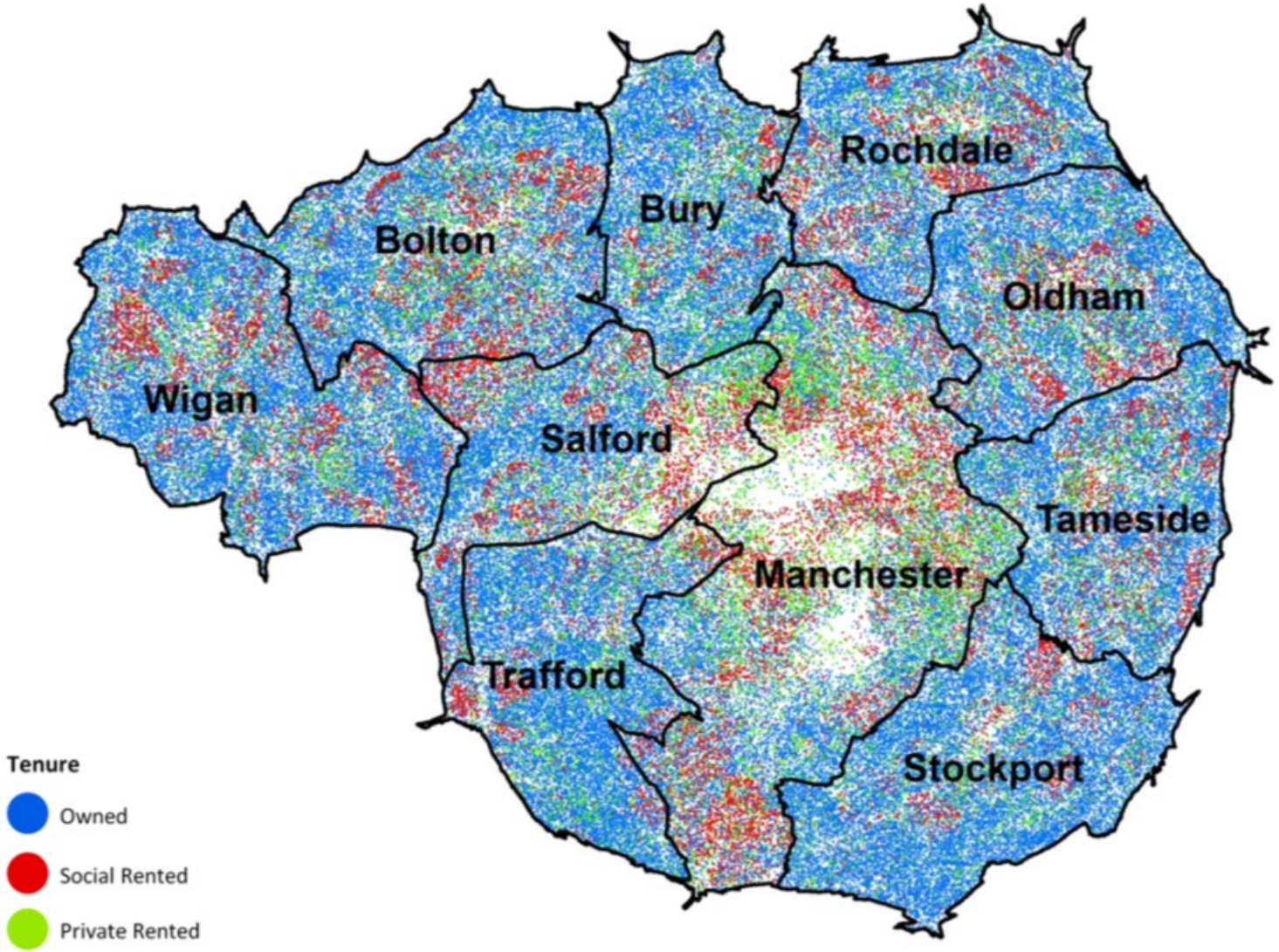


Private Rented



Greater Manchester

Summarised by a dot-density map



Cartogram of couples with dependent children by tenure in Greater Manchester

Sources: ONS boundary data; ONS lookup tables; Census 2011 standard table LC4101EW contains National Statistics data © Crown copyright and database right 2012 and Ordnance Survey data © Crown copyright and database right 2012

UK Data Service

Change over time

Combining census data from multiple years

Tables

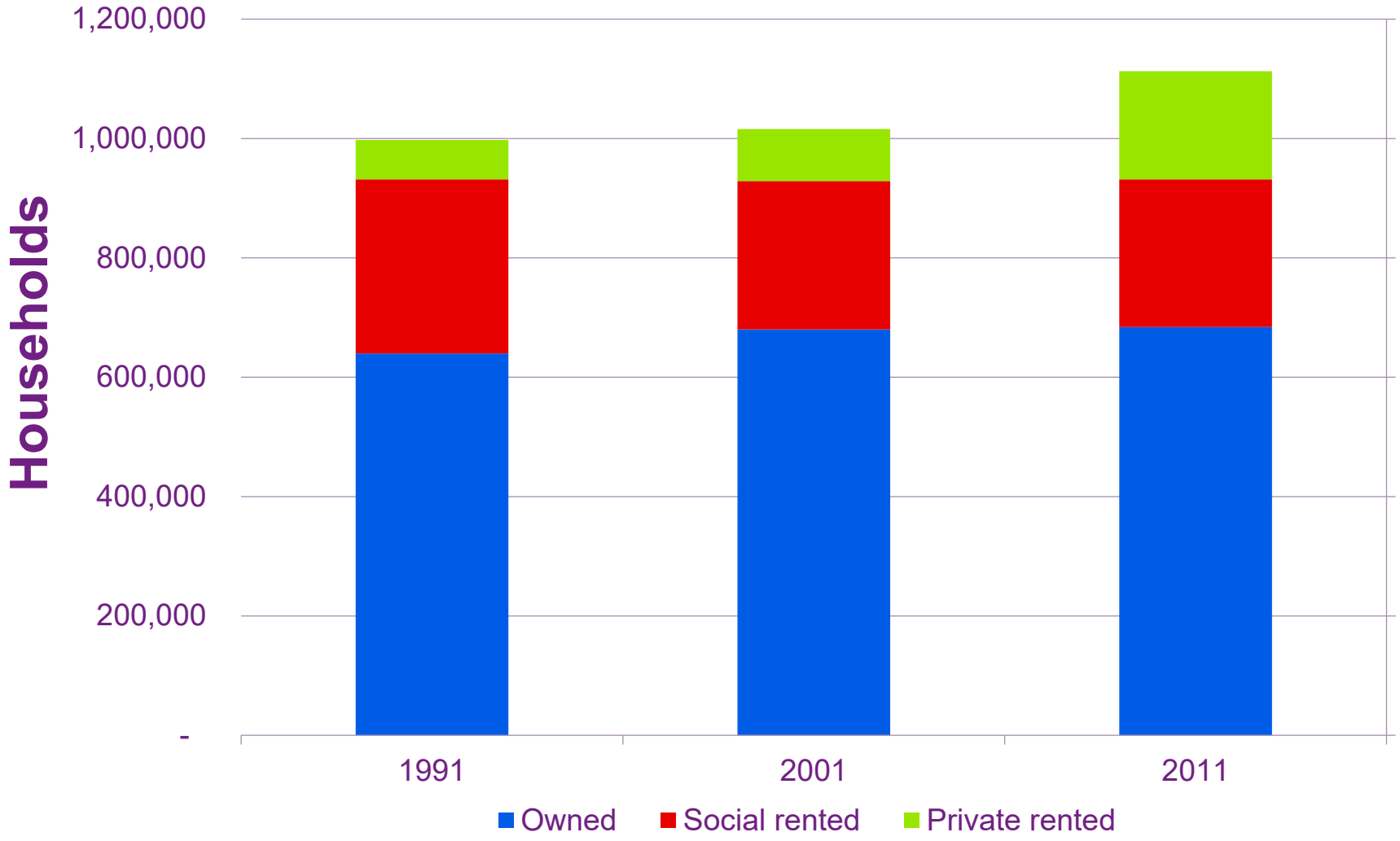
Table 1: changes in the numbers and proportions of household by tenure and region from 2001 to 2011 in descending order of the number of PRS households in the region in 2011

Households ('000s)	2001 (proportion by tenure)			2011 (proportion by tenure)			Change (%age 2001 to 2011)		
Region	Owned	Social	PRS	Owned	Social	PRS	Owned	Social	PRS
England	14,054 (69%)	3,940 (19%)	2,456 (12%)	14,149 (64%)	3,904 (18%)	4,011 (18%)	95 (1%)	-37 (-1%)	1,555 (63%)
London	1,705 (57%)	790 (26%)	521 (17%)	1,618 (50%)	786 (24%)	862 (26%)	-86 (-5%)	-4 (-1%)	341 (65%)
South East	2,431 (74%)	459 (14%)	397 (12%)	2,444 (69%)	487 (14%)	624 (18%)	12 (0%)	29 (6%)	227 (57%)
North West	1,948 (69%)	565 (20%)	300 (11%)	1,957 (65%)	550 (18%)	502 (17%)	9 (0%)	-14 (-2%)	202 (67%)
South West	1,524 (73%)	282 (14%)	280 (13%)	1,544 (68%)	302 (13%)	419 (18%)	20 (1%)	19 (7%)	139 (50%)
East	1,622 (73%)	369 (17%)	240 (11%)	1,656 (68%)	380 (16%)	387 (16%)	33 (2%)	12 (3%)	147 (61%)
Yorks and Humber	1,396 (68%)	434 (21%)	234 (11%)	1,435 (65%)	402 (18%)	386 (17%)	39 (3%)	-31 (-7%)	152 (65%)
West Midlands	1,498 (70%)	444 (21%)	212 (10%)	1,504 (66%)	435 (19%)	355 (15%)	6 (0%)	-8 (-2%)	143 (67%)
East Midlands	1,251 (72%)	303 (17%)	178 (10%)	1,287 (68%)	300 (16%)	308 (16%)	37 (3%)	-3 (-1%)	129 (72%)
North East	678 (64%)	295 (28%)	93 (9%)	703 (62%)	260 (23%)	168 (15%)	24 (4%)	-35 (-12%)	75 (81%)

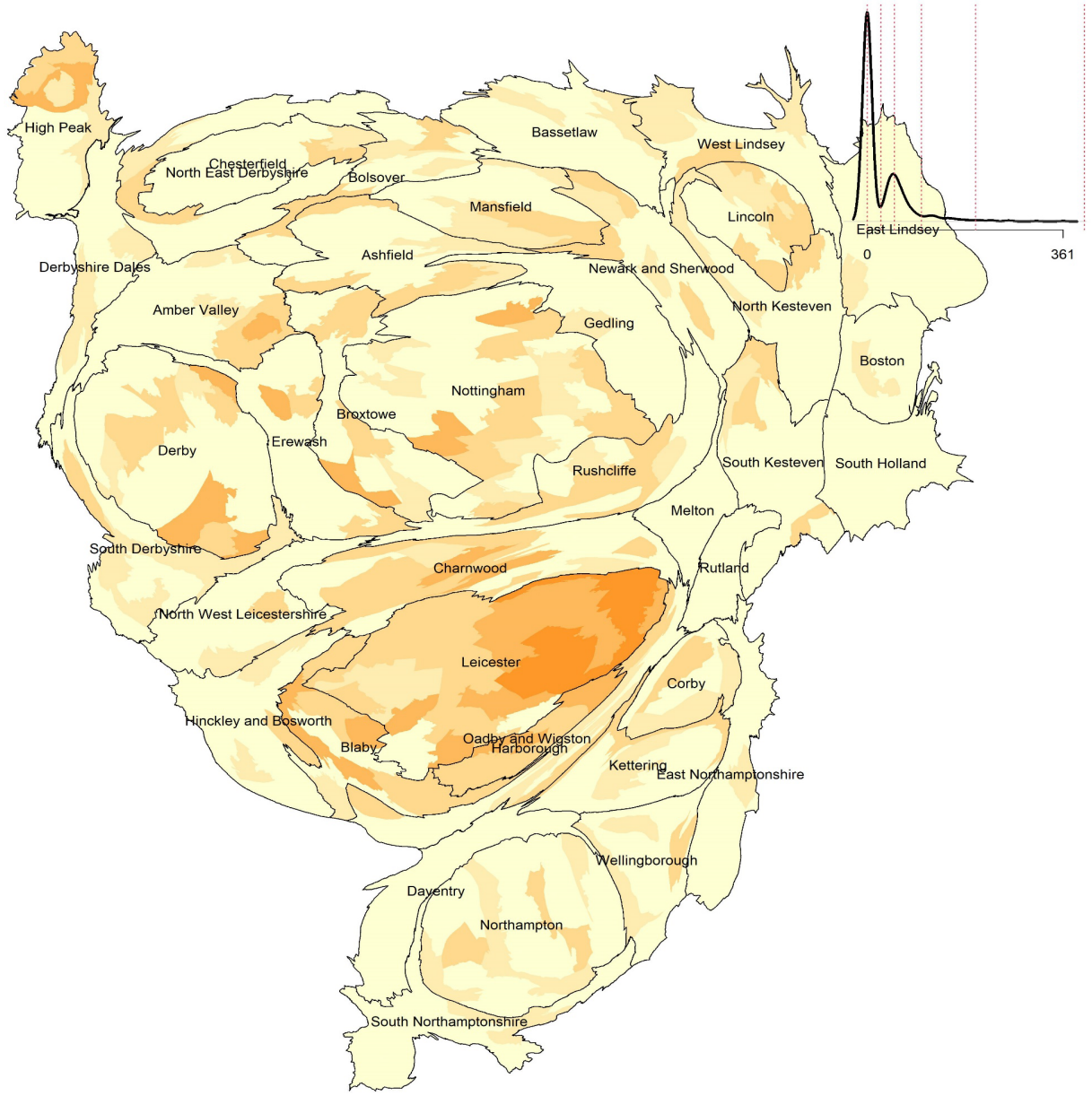
Sources: Census 2001 standard table KS18; Census 2011 standard table KS402EW

Charts

Figure: changes in household tenure in Greater Manchester from 1991, 2001 and 2011 census data



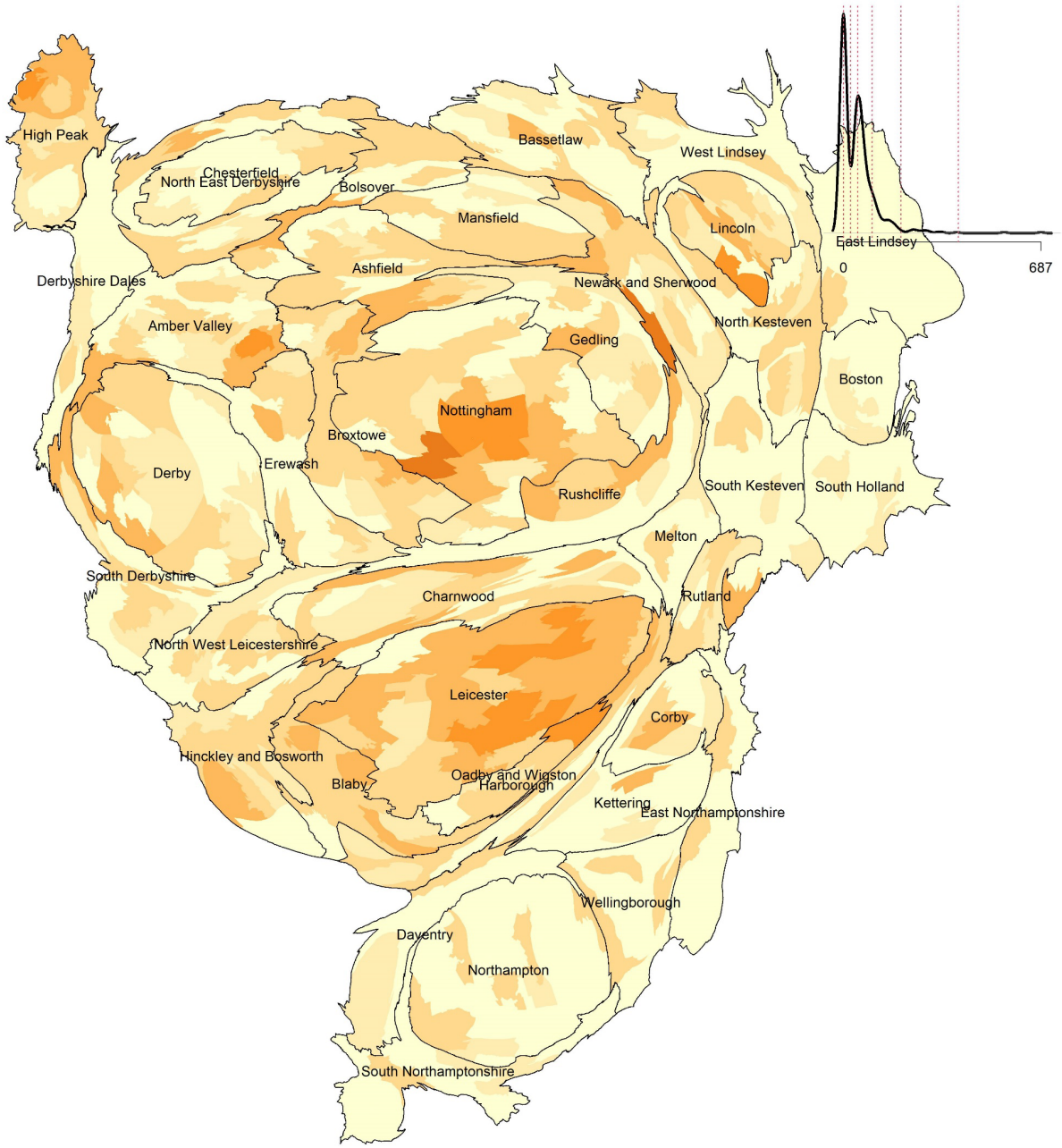
Cartogram of Covid cases in East Midlands 21st September 2020



Rate per 100,000 population

- more than 3,200
- more than 1,600 – 3,200
- more than 800 – 1,600
- more than 400 – 800
- more than 200 – 400
- more than 100 – 200
- more than 50 – 100
- more than 25 – 50
- 0 - 25

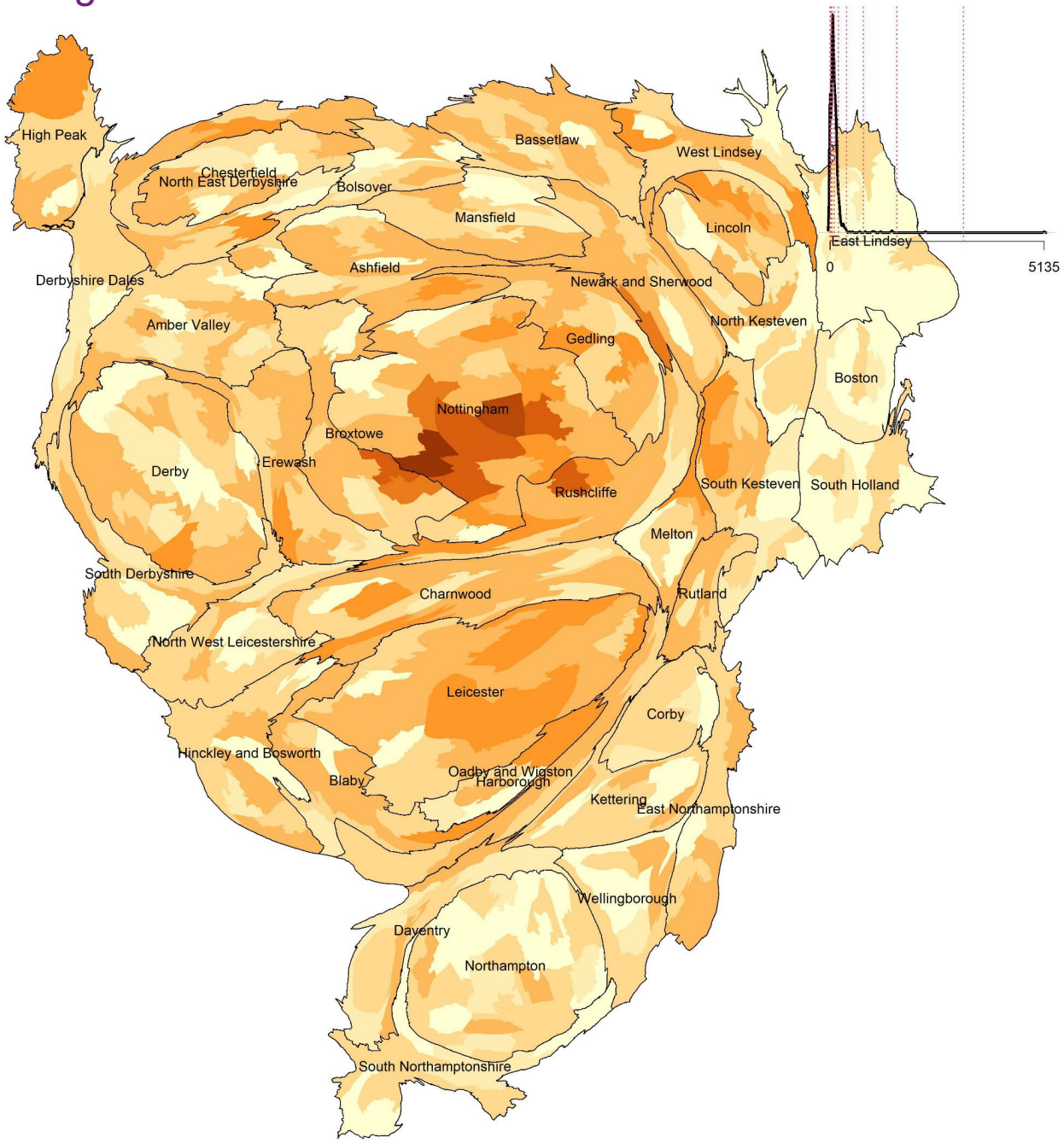
Cartogram of Covid cases in East Midlands 28th September 2020



Rate per 100,000 population

- more than 3,200
- more than 1,600 – 3,200
- more than 800 – 1,600
- more than 400 – 800
- more than 200 – 400
- more than 100 – 200
- more than 50 – 100
- more than 25 – 50
- 0 - 25

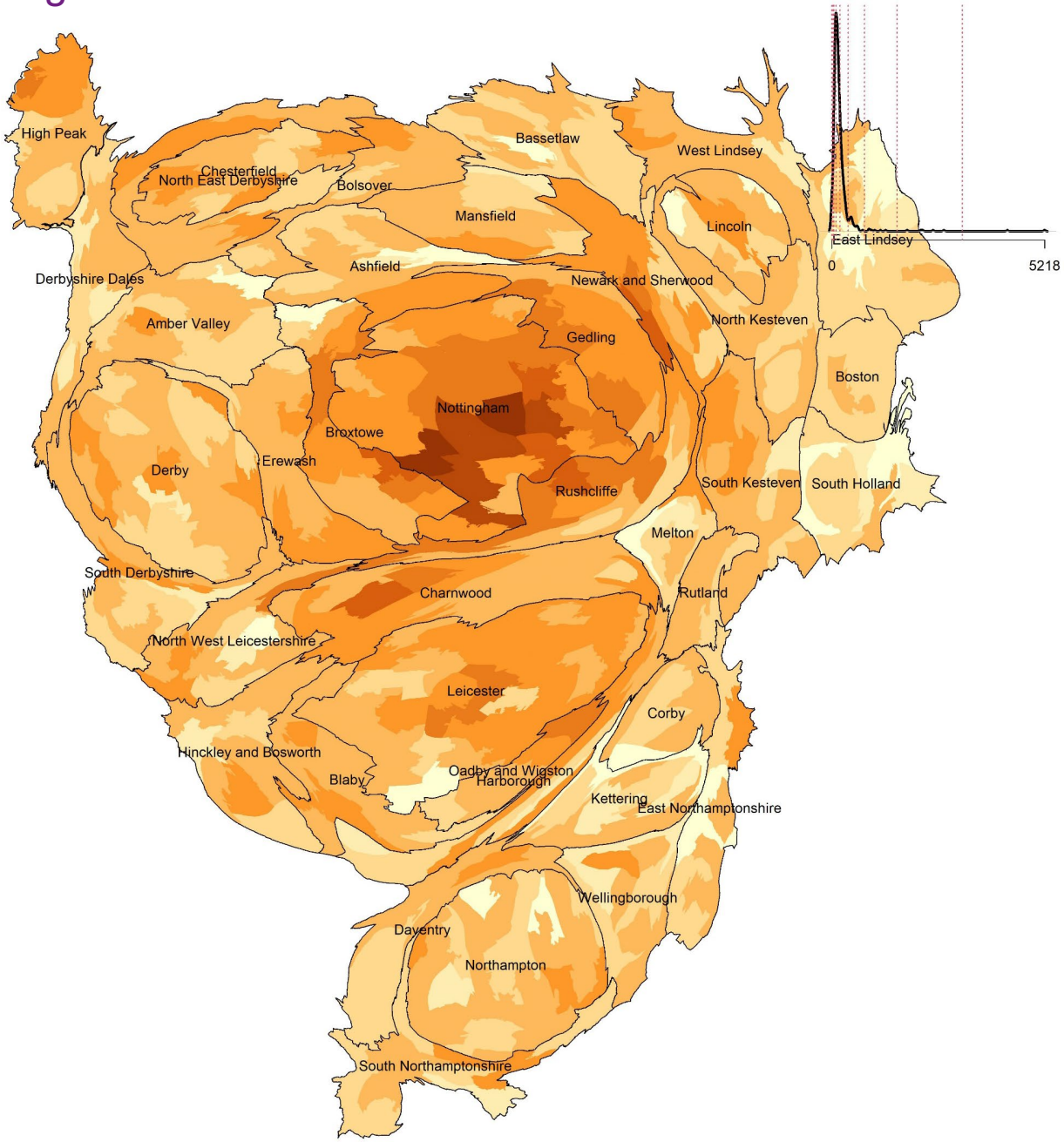
Cartogram of Covid cases in East Midlands 4th October 2020



Rate per 100,000 population

- more than 3,200
- more than 1,600 – 3,200
- more than 800 – 1,600
- more than 400 – 800
- more than 200 – 400
- more than 100 – 200
- more than 50 – 100
- more than 25 – 50
- 0 - 25

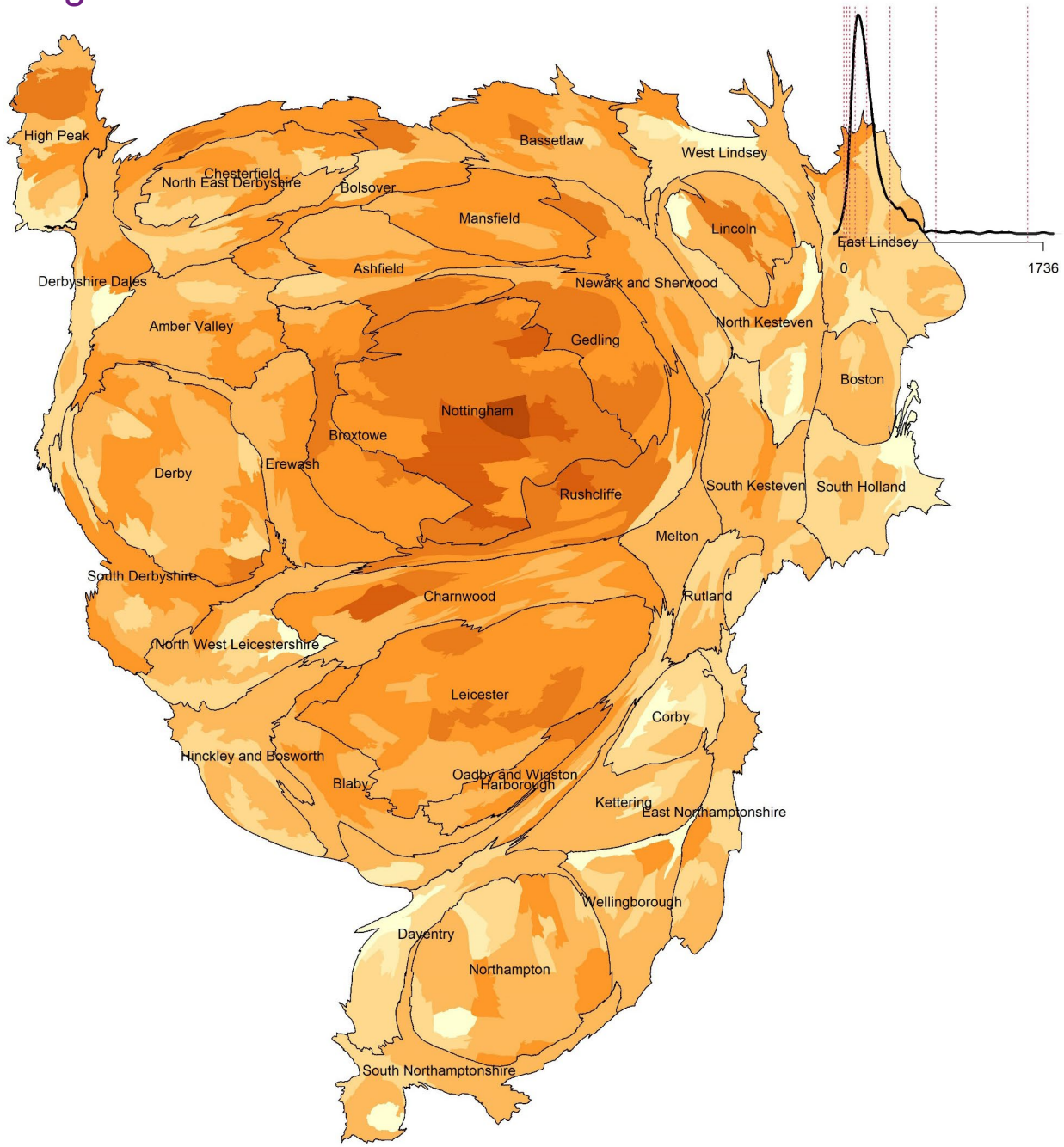
Cartogram of Covid cases in East Midlands 11th October 2020



Rate per 100,000 population

- more than 3,200
- more than 1,600 – 3,200
- more than 800 – 1,600
- more than 400 – 800
- more than 200 – 400
- more than 100 – 200
- more than 50 – 100
- more than 25 – 50
- 0 - 25

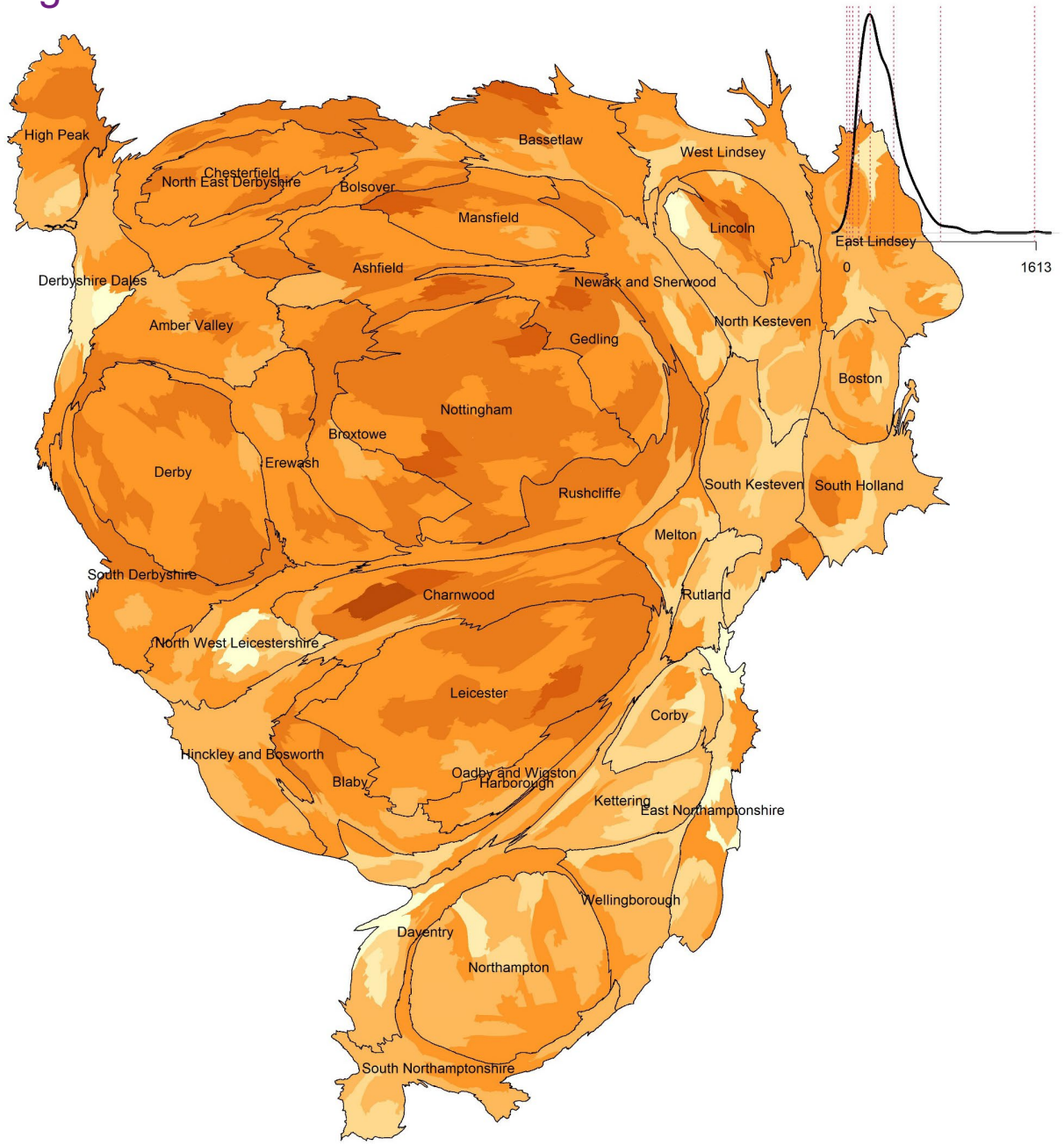
Cartogram of Covid cases in East Midlands 18th October 2020



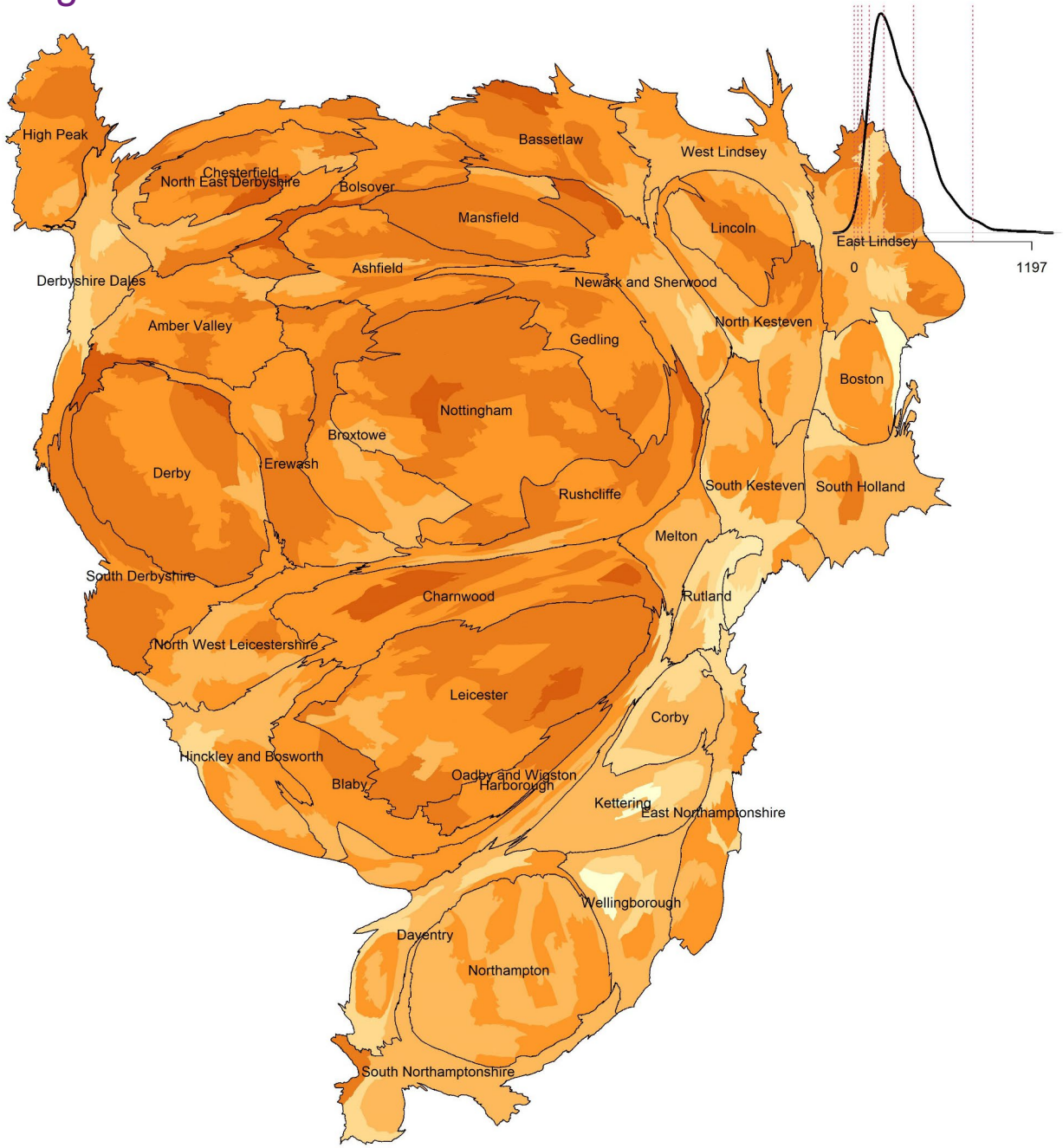
Rate per 100,000 population

- more than 3,200
- more than 1,600 – 3,200
- more than 800 – 1,600
- more than 400 – 800
- more than 200 – 400
- more than 100 – 200
- more than 50 – 100
- more than 25 – 50
- 0 - 25

Cartogram of Covid cases in East Midlands 25th October 2020



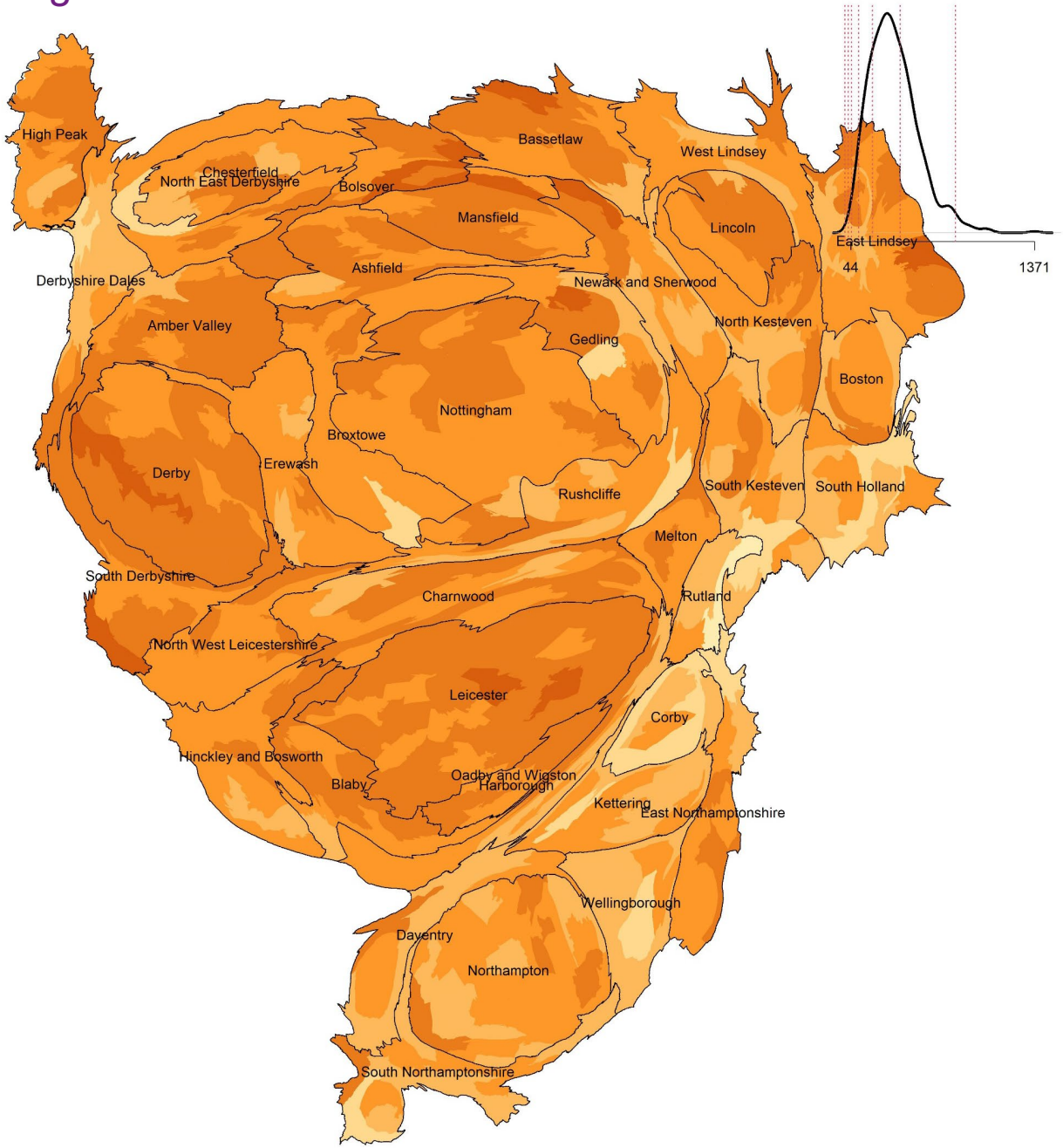
Cartogram of Covid cases in East Midlands 1st November 2020



Rate per 100,000 population

- more than 3,200
- more than 1,600 – 3,200
- more than 800 – 1,600
- more than 400 – 800
- more than 200 – 400
- more than 100 – 200
- more than 50 – 100
- more than 25 – 50
- 0 - 25

Cartogram of Covid cases in East Midlands 8th November 2020



UK Data Service

Potential issues to consider

Issues to consider

- Impact of Covid-19 on e.g. employment, place of residence, work and education: – ONS report will include some consideration in first release with more detailed report to follow
- Learning from the quality assurance process with local authorities – some information will be published
- Coverage and statistical imputation
- Changing geography
- General trust and unwillingness to disclose some characteristics

Census products

UKDS versions

(alternatives provided by ONS, NOMIS and potentially others)

Data

Likely to also be available from other sources

- Aggregate data – interface to select variables (allowed combinations)
- Geographical data – boundaries used in census outputs

Likely to only be available from UK Data Service

- Flow data
- Microdata

Training

- Awareness raising
- How to … sessions
  - Aggregate data
  - Boundary data
  - Flow data
  - Microdata
  - Techniques and substantive areas of interest

Contextual background

- Statistical disclosure control
- Modifiable areal unit problem
- Coverage and imputation
- Implications of Covid-19

Support

- Documents providing updates on release schedules
- Video materials supporting training
- Explainer for key census variables
- Help Desk

UK Data Service

Thank you.

Nigel de Noronha

nigel.denoronha@manchester.ac.uk