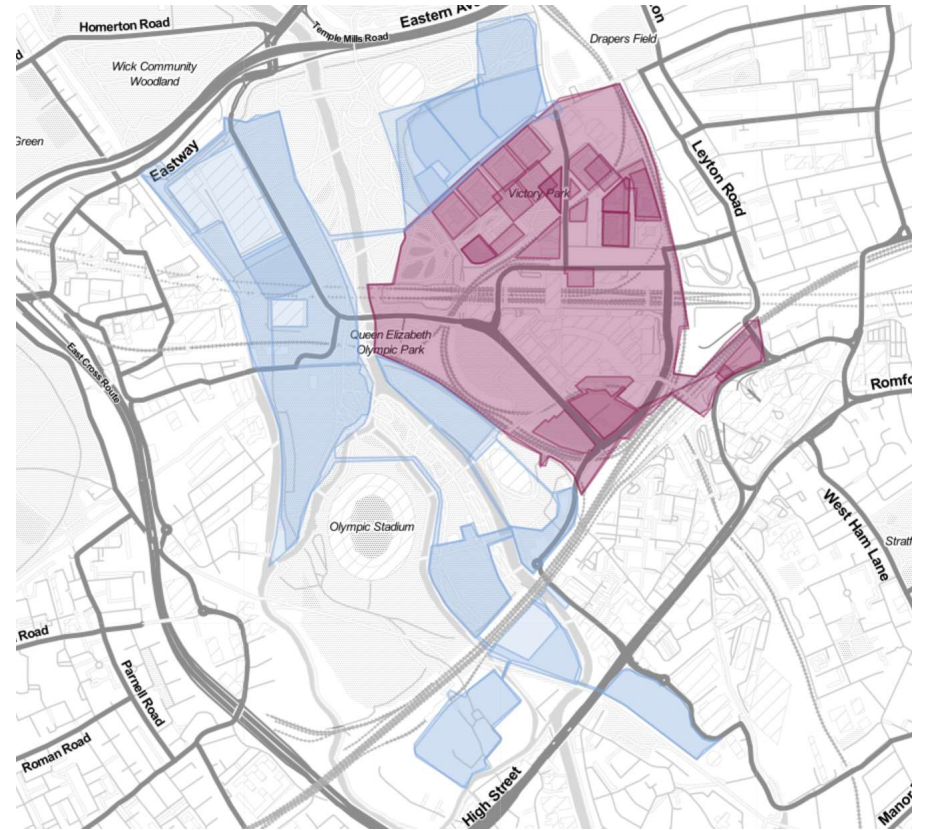


# Local estimates and projections

Ben Corr, Chris Fairless, Marta Lapsley  
Greater London Authority

# Fast-changing London

- New development
- Gentrification/affordability
- Changing migration patterns
- Welfare reforms



Stratford City development (red)

- 3152 units built
- 4172 more proposed

Olympic Park (blue)

- 258 units built
- 3232 more proposed



# Our population projection models

- Trend
  - Gives London level constraint
  - Published on London Datastore



# Our population projection models

- Trend
  - Gives London level constraint
  - Published on London Datastore
- Housing-led, small area
  - Includes local development trajectories
  - Published projections: Strategic Household Land Availability Assessment (SHLAA) for development trajectory
  - 100s of bespoke projections per year
  - Widely used for Planning for local service demand

# Current challenges in small area model

- SHLAA is about land availability, not planned development
  - **Will** it be built?
  - **When** will it be built?
  - **What** will be built? Tenure/bedrooms etc
- Only gives us number of units
- Occupation of units assumed to follow historic pattern for area
- Official data sources less reliable at small area
- How to run, QA, communicate, and document
  - 100s runs
  - 1000s areas



# Talk outline

- Understanding housing stock
- Understanding housing occupation
- Model implementation

# Understanding housing stock

- 2011 census
  - Full coverage
  - Limited categorical detail at small area

**Slide 7**

---

**ML1** Is microsimulation appropriate term for getting existing housing stock details?

Marta Lapsley, 10/09/2019



# Understanding housing stock

- 2011 census
  - Full coverage
  - Limited categorical detail at small area
- Planning Permissions Data (LDD)
  - Up to date
  - Tenure/bedrooms
  - Doesn't include existing stock
  - Don't know how many permissions will actually get built, and when
  - Not linked to statistical geographies

**Slide 8**

---

**ML1** Is microsimulation appropriate term for getting existing housing stock details?

Marta Lapsley, 10/09/2019

# Understanding housing stock

- 2011 census
  - Full coverage
  - Limited categorical detail at small area
- Planning Permissions Data (LDD)
  - Up to date
  - Tenure/bedrooms
  - Doesn't include existing stock
  - Don't know how many permissions will actually get built, and when
  - Not linked to statistical geographies
- Code-Point residential address data

**Slide 9**

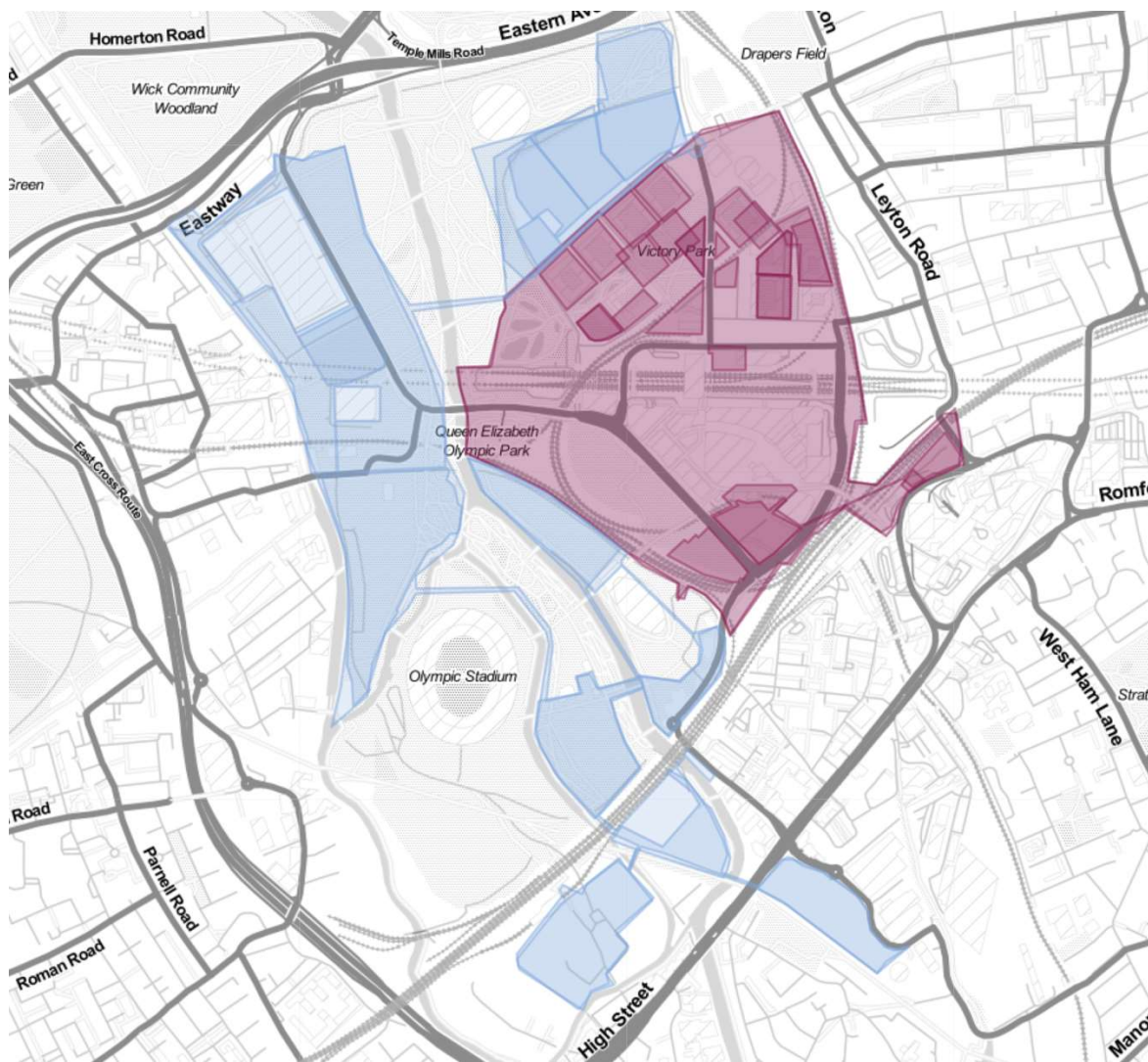
---

**ML1** Is microsimulation appropriate term for getting existing housing stock details?

Marta Lapsley, 10/09/2019

# London Development Database modelling

Libby Rogers (GLA)



### Olympic park (blue)

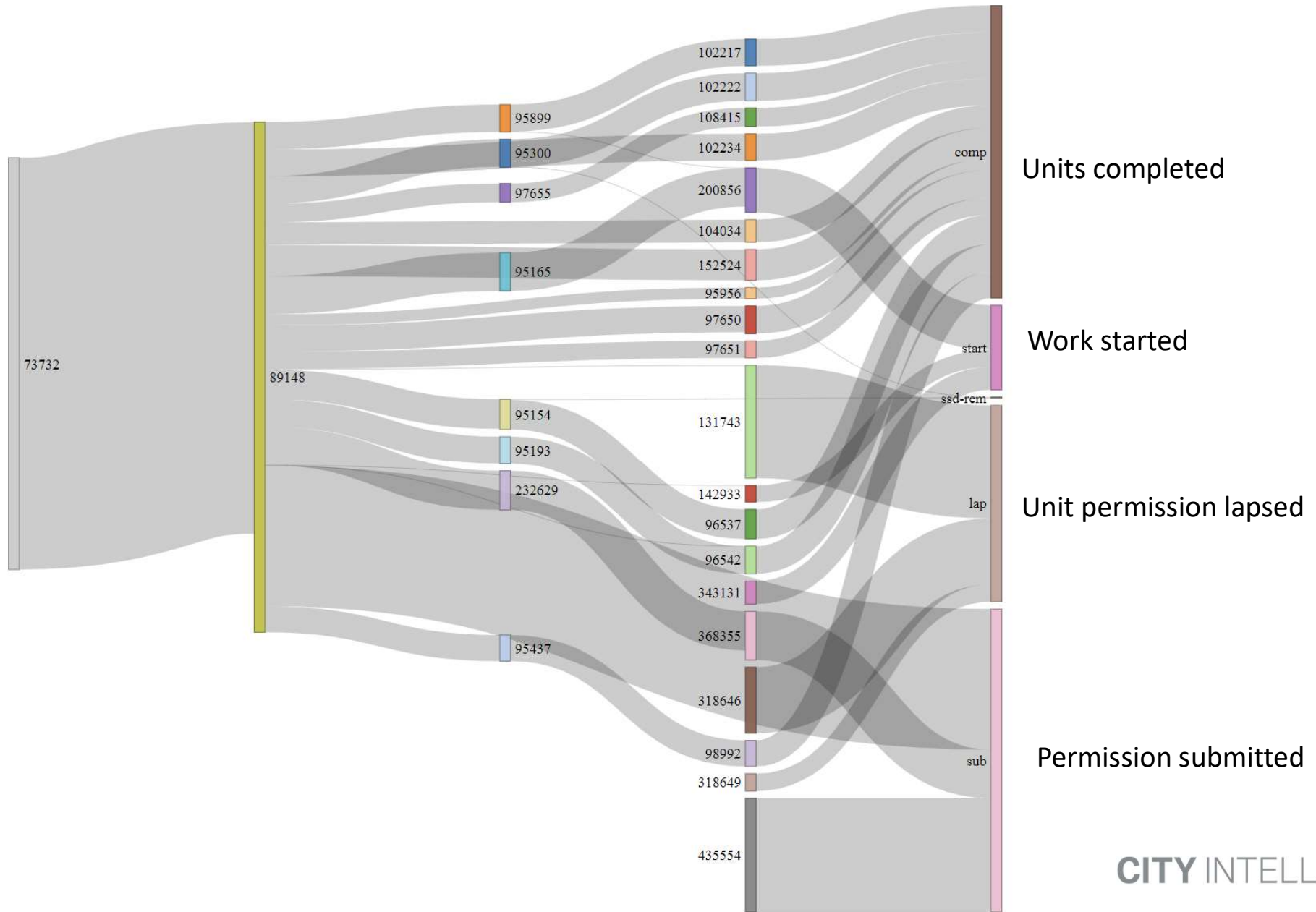
- 258 units built
- 3232 more proposed

### Stratford city development (red)

- 3152 units built
- 4172 more proposed



# Linking the permissions





**Slide 13**

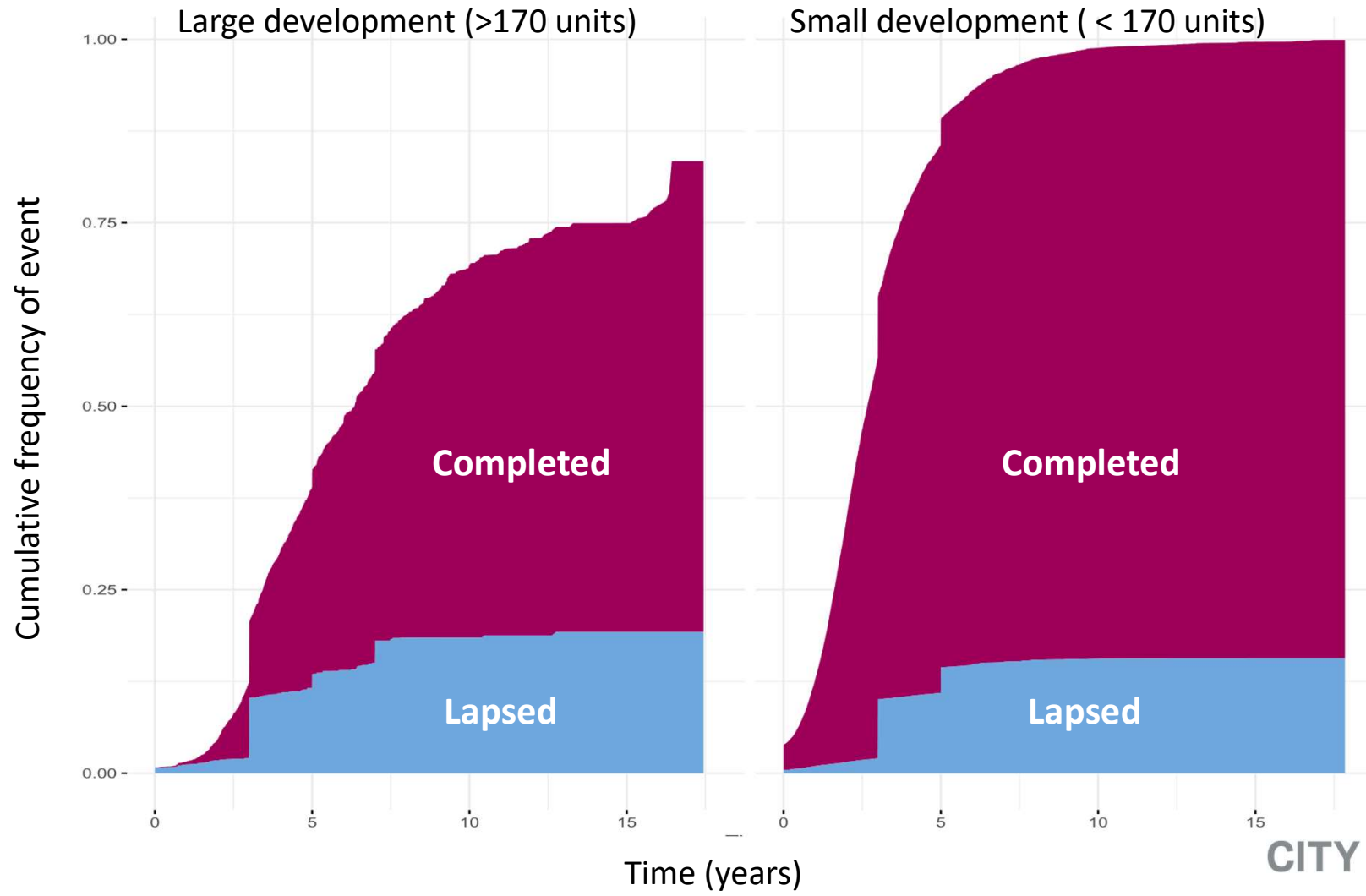
---

**ML2**

What was the methodology Libby used here?

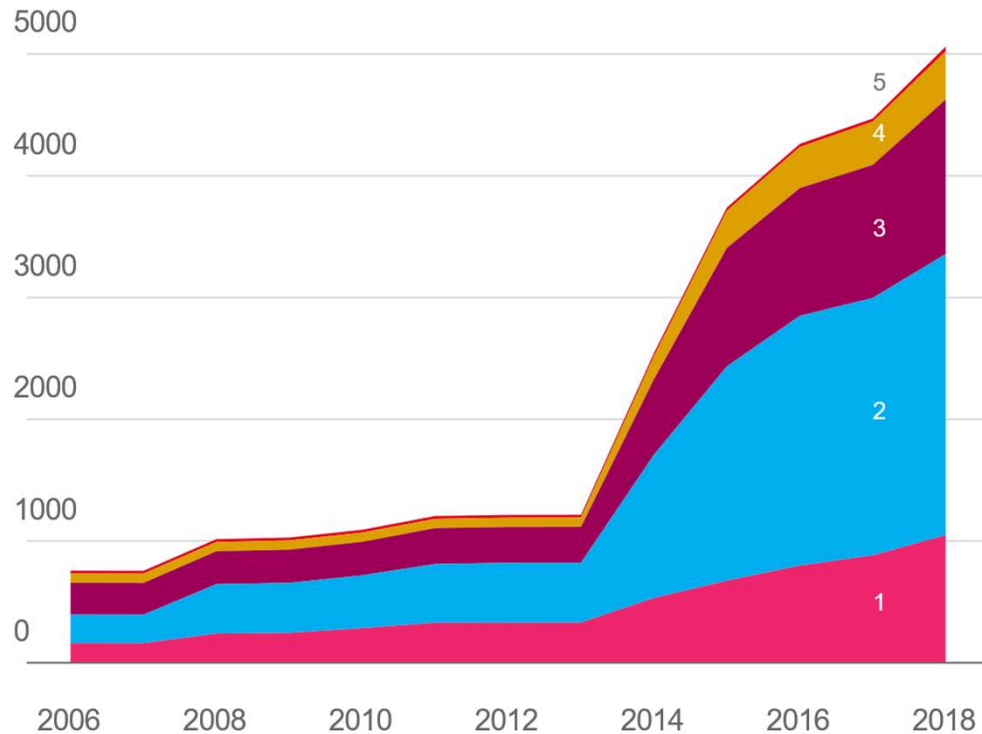
Marta Lapsley, 10/09/2019

# Time to completion



## Validation against address data e.g. Stratford City

New units by bedrooms in Stratford City

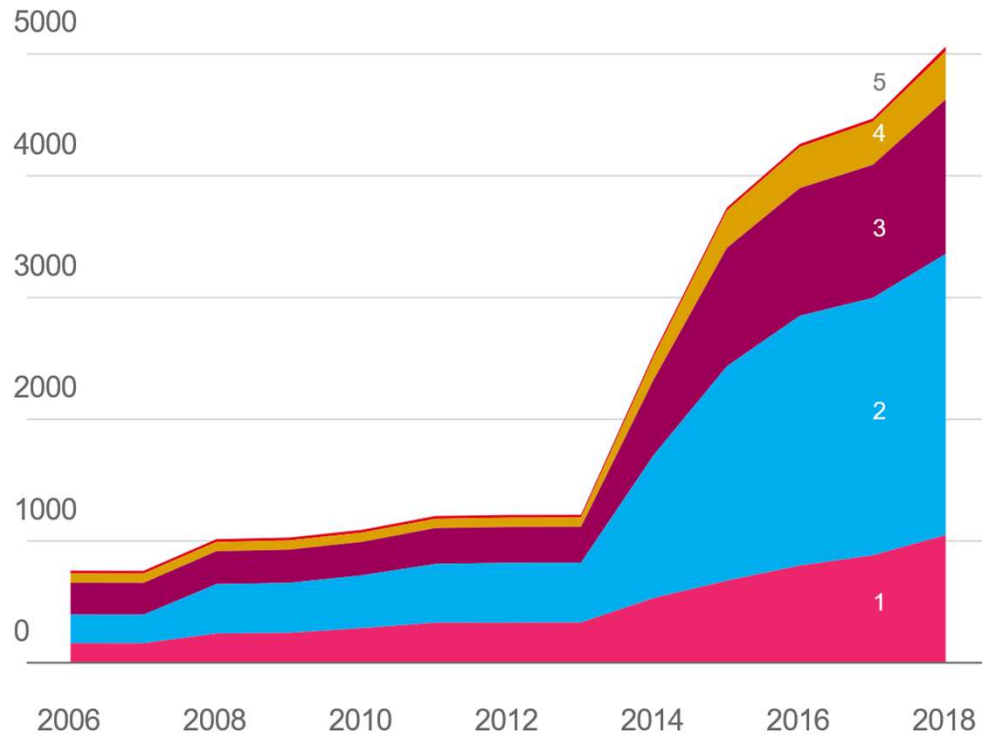


Code-Point domestic addresses in Stratford City

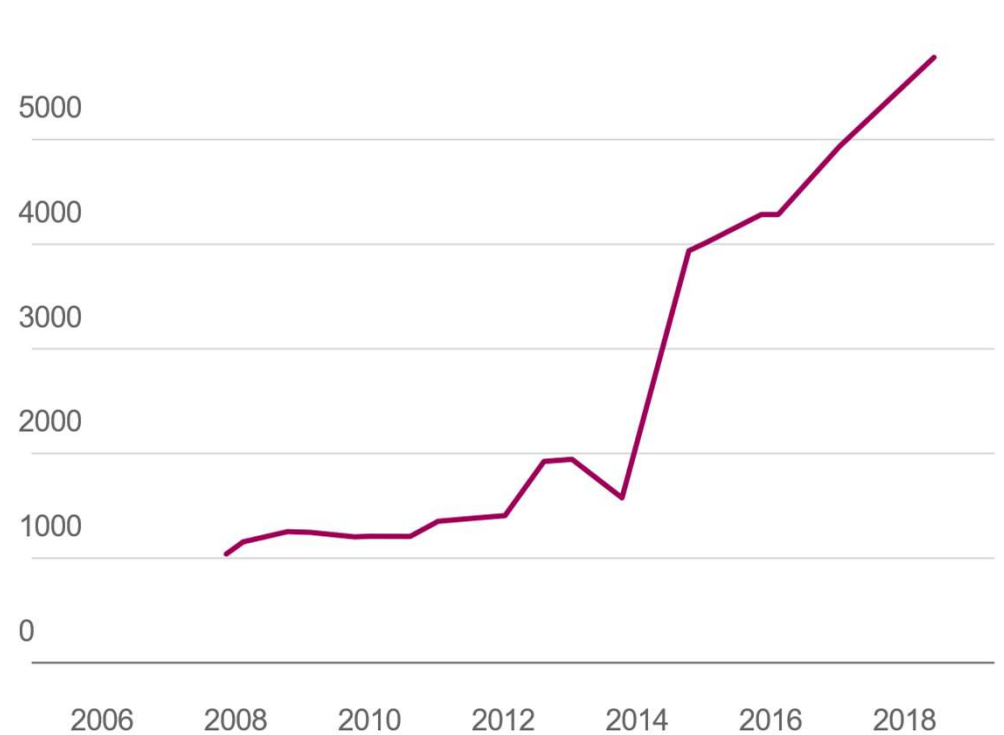
- Number of domestic delivery points per postcode
- Ordnance Survey product
- Quarterly data

## Validation against address data e.g. Stratford City

### New units by bedrooms in Stratford City



### Code-Point domestic addresses in Stratford City





# Understanding housing occupation

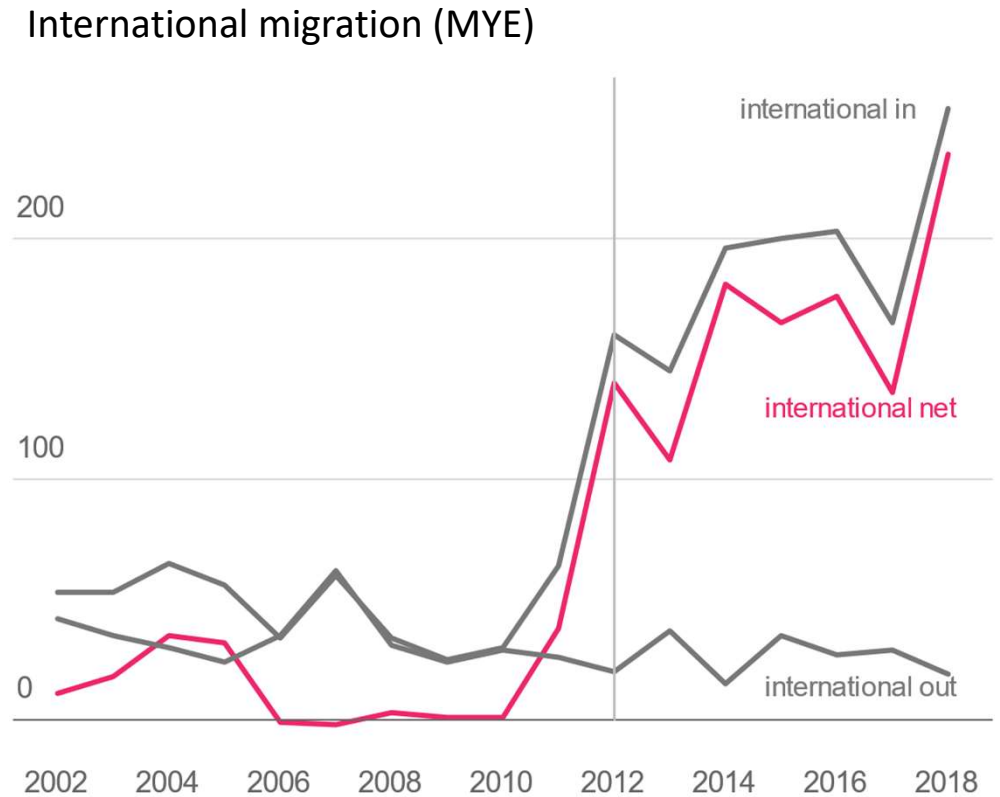
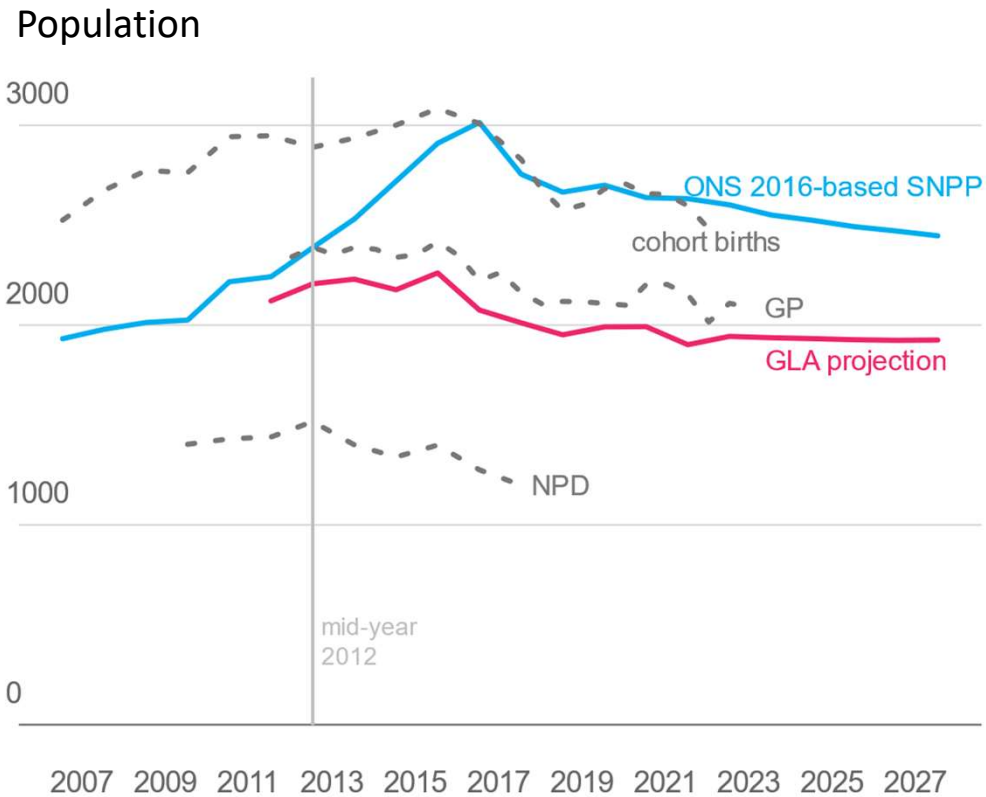
- Data we've used for a while
  - Census (2001, 2011, ...)
  - Mid-year Estimates: Local Authority population, births, deaths and migration
  - Small area population estimates

# Understanding housing occupation

- Data we've used for a while
  - Census (2001, 2011, ...)
  - Mid-year Estimates: Local Authority population, births, deaths and migration
  - Small area population estimates
- New data sources
  - GP register
  - National Pupil Database
  - Consumer Data Research Centre data
  - Small area births data (LSOA)
  - Small area housing occupation data (census, LSOA)
  - Travel time/cost/accessibility data
  - House and rental price

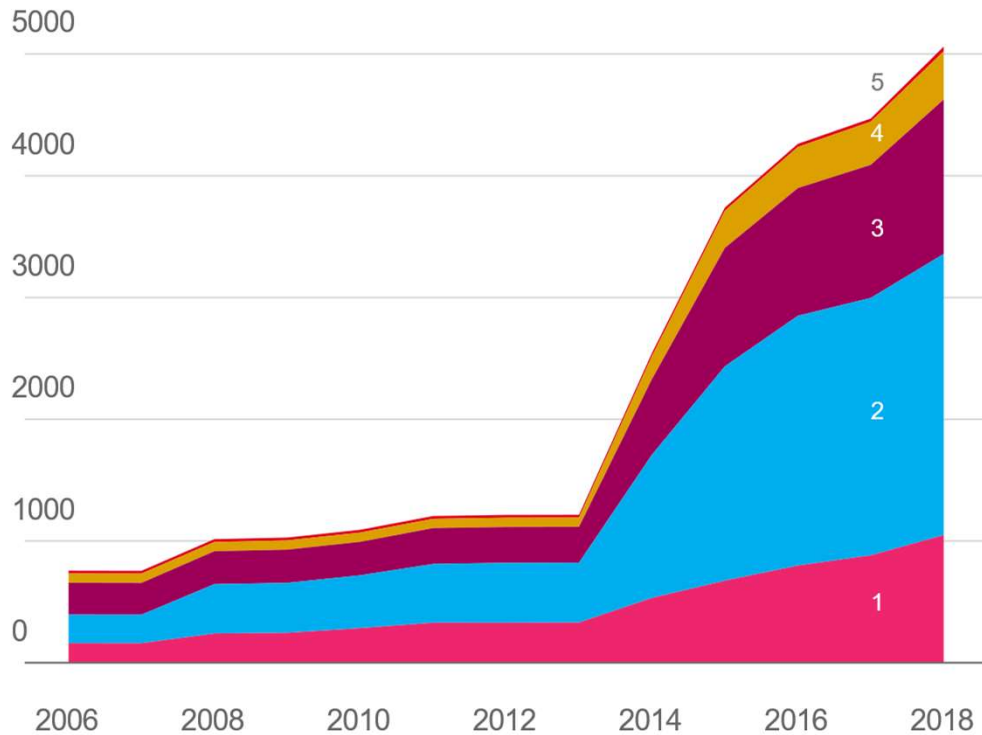
# Overcoming limitations of traditional data sources at small areas

e.g. 4 year olds in Westminster



# Understanding who occupies new development e.g. Stratford City

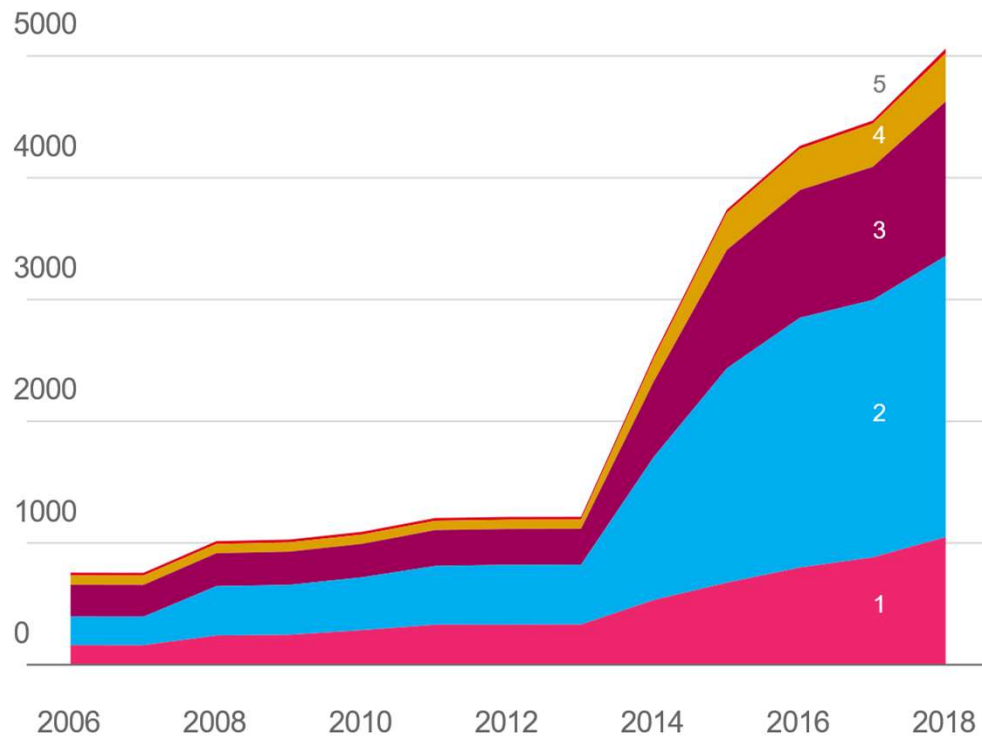
## New units by bedrooms in Stratford City





## Understanding who occupies new development e.g. Stratford City

### New units by bedrooms in Stratford City

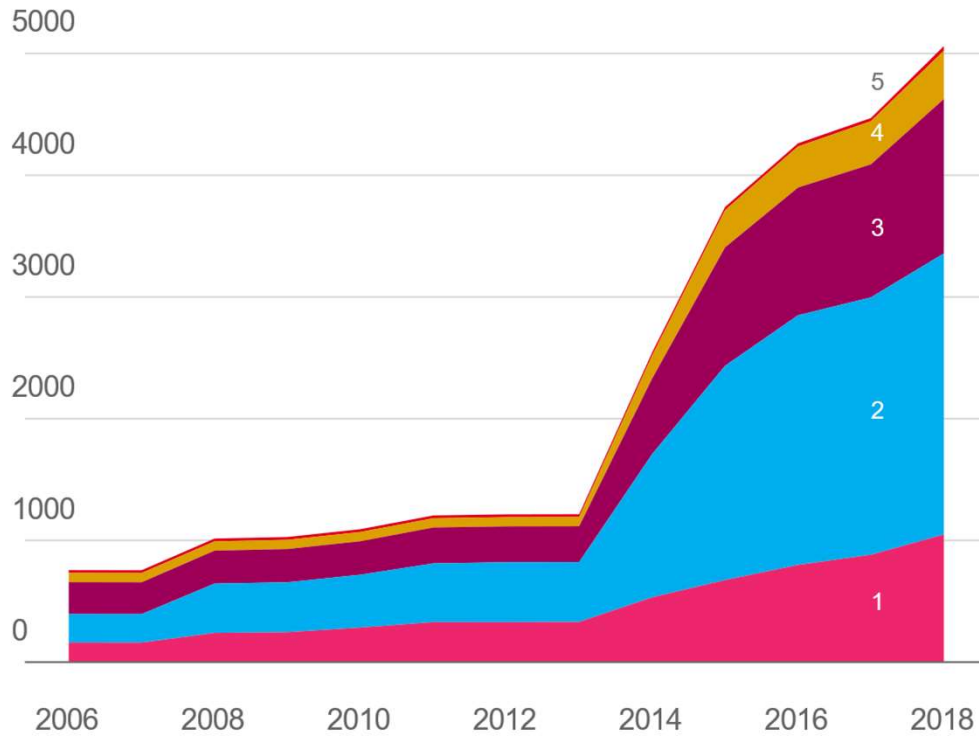


### GP registrations in Stratford City by age

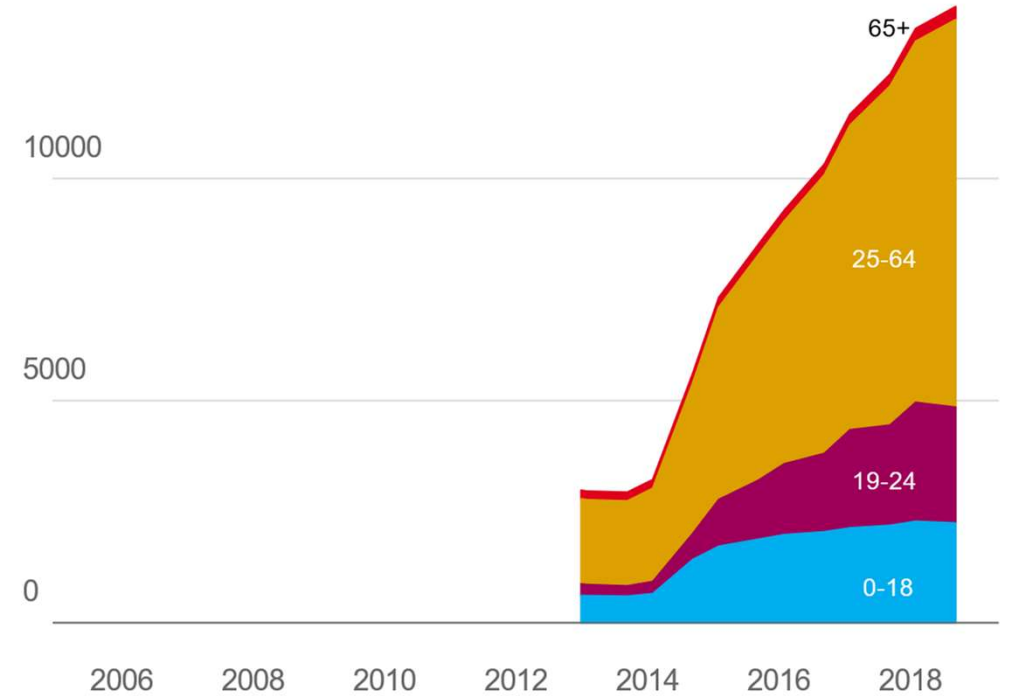
- Down to LSOA
- By sex and single year of age
- Only a few weeks lag
- 4 times per year
- Good for children, women & older people

# Understanding who occupies new development e.g. Stratford City

## New units by bedrooms in Stratford City



## GP registrations in Stratford City by age



# Understanding drivers for occupation

## Current methodology:

- Fill housing using characteristics of historical migrants

## In development methodology:

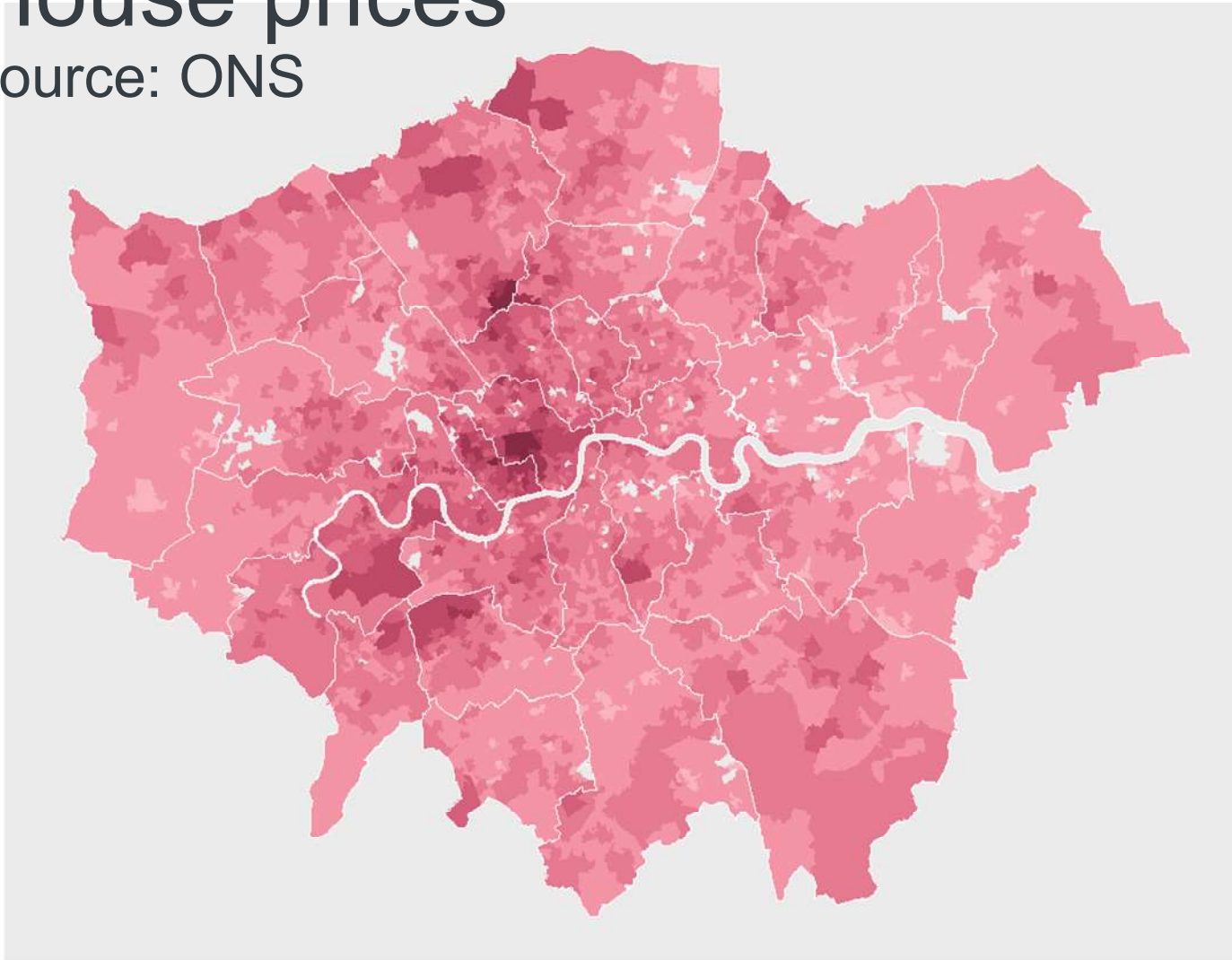
- Migrate adults as above, then migrate children according to observed numbers of children for each housing type and number of bedrooms

## Future work:

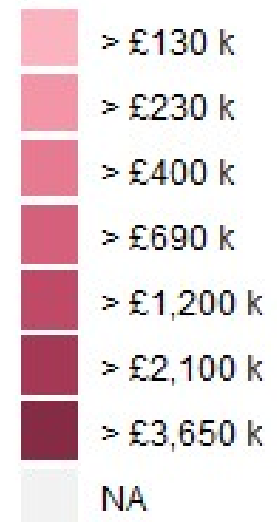
- Better understand patterns and drivers of occupation
  - CDRC (collaboration to understand potential use of consumer data)
  - TfL travel time data
  - Housing value data

# House prices

source: ONS

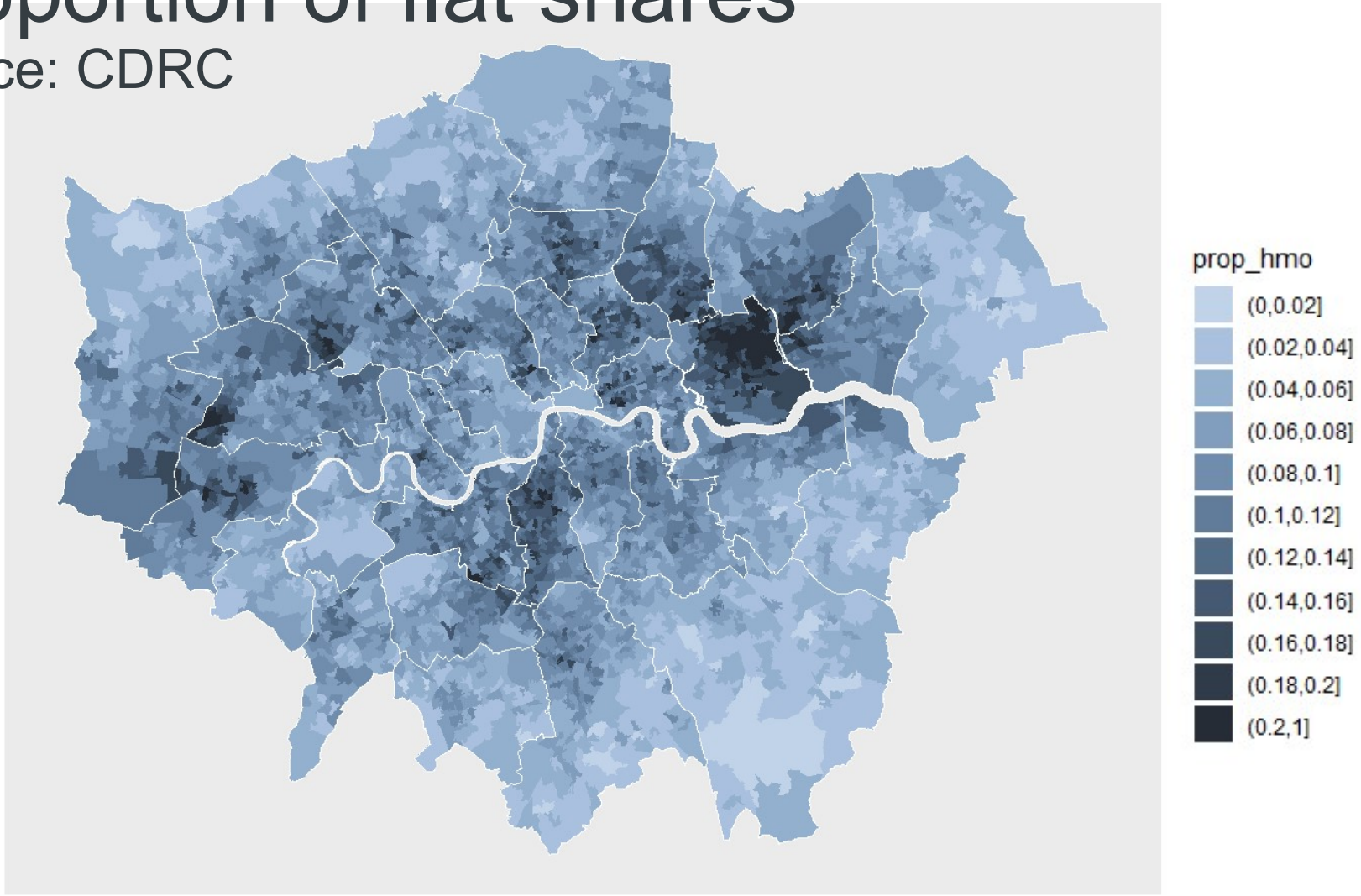


priceband



# Proportion of flat shares

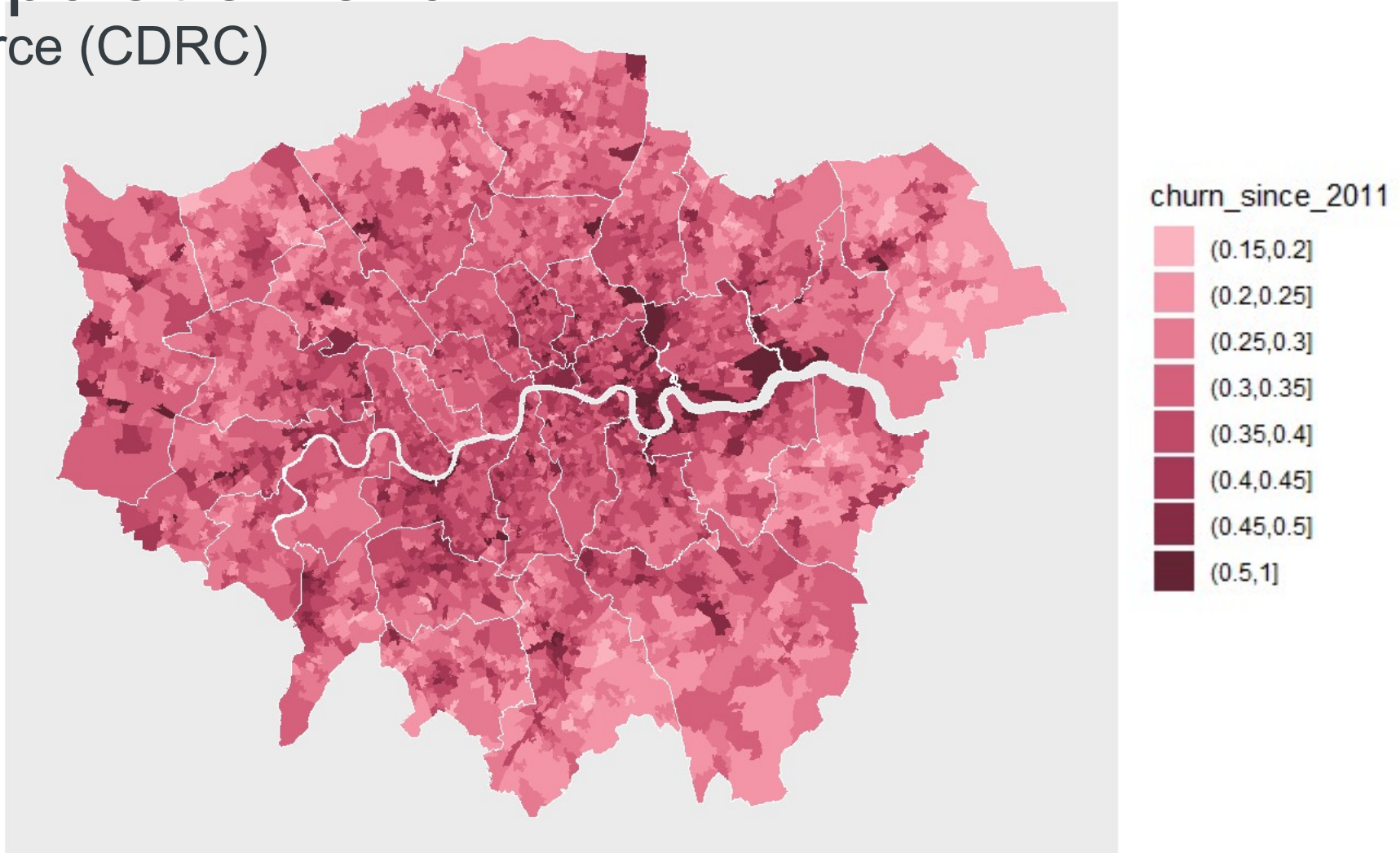
source: CDRC





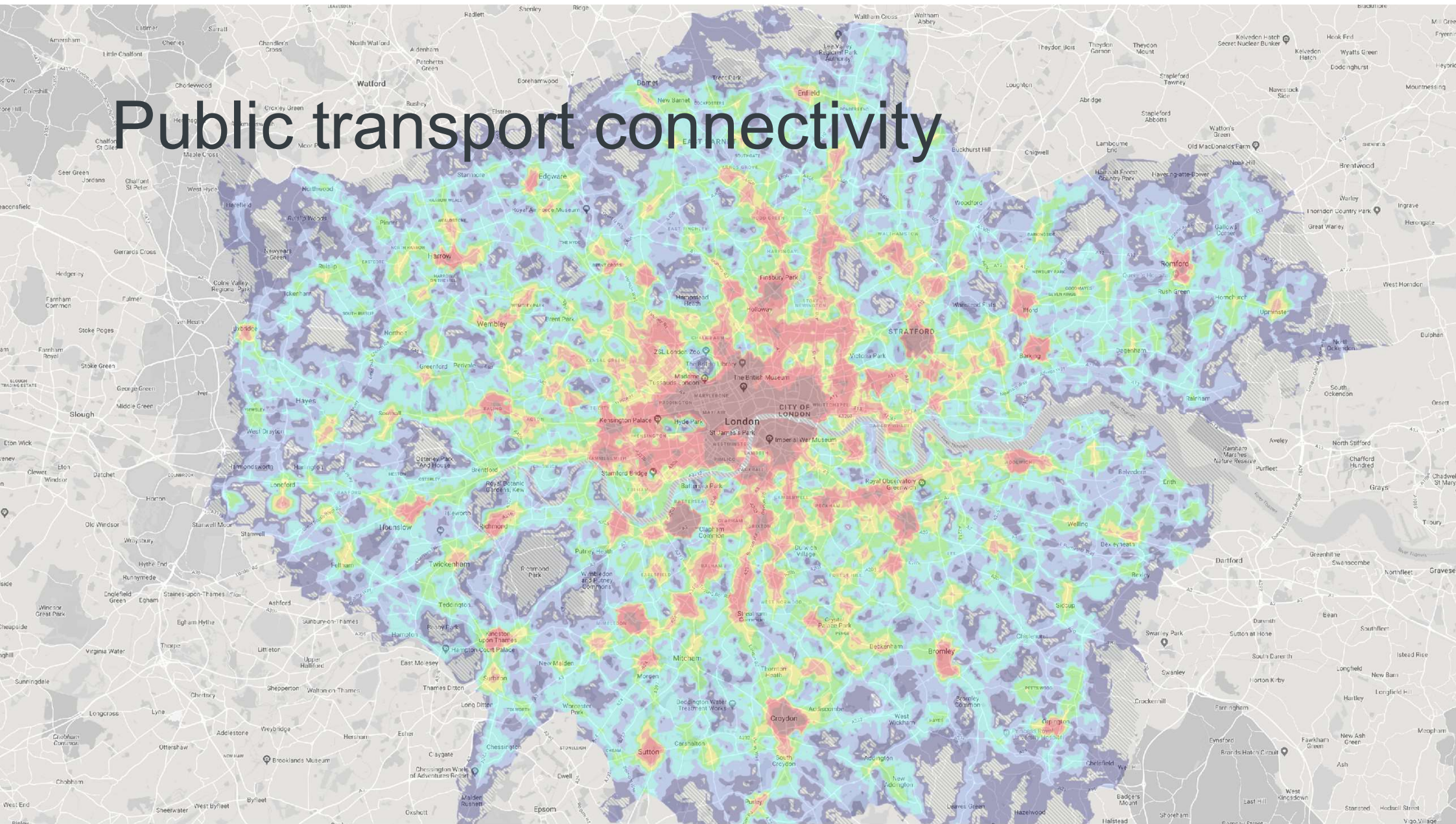
# Population churn

source (CDRC)





# Public transport connectivity



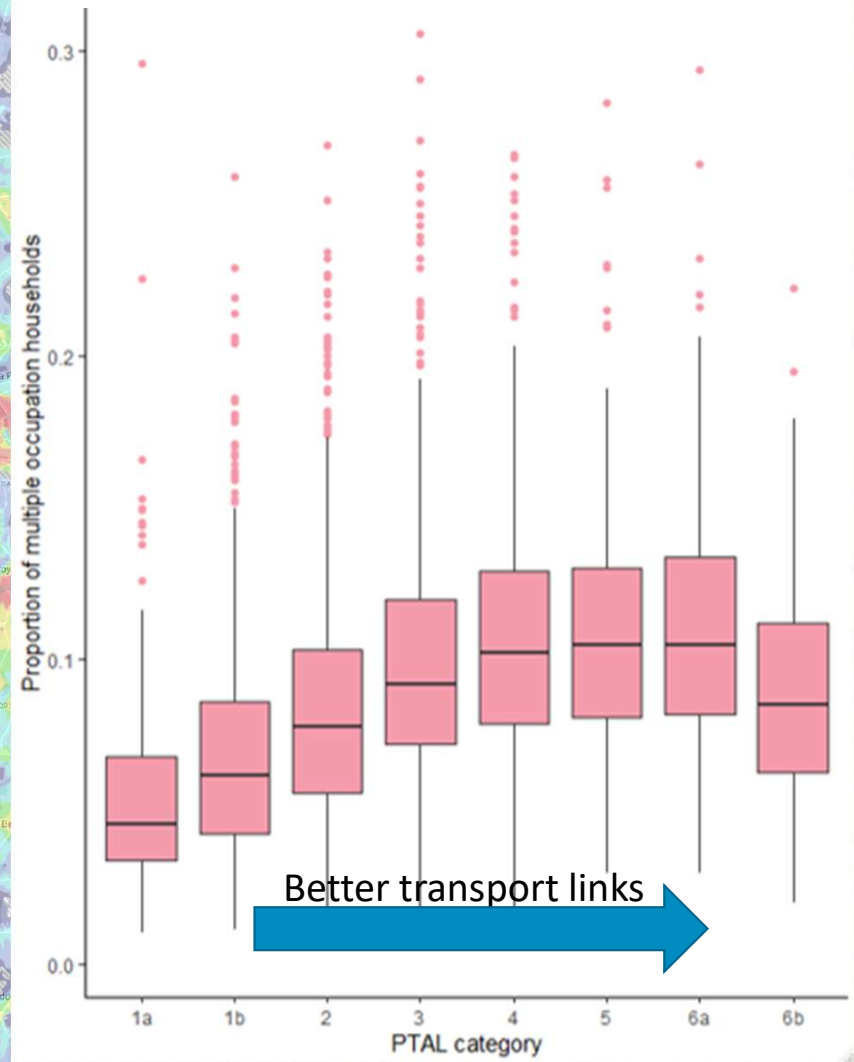


# Public transport connectivity

- London geography is more complex than distance from the city centre
- We've used the *traveltimes* API and TfL Public Transport Accessibility Index to look at the whole South-East of England
- House shares are more frequent in better-connected locations, and households have fewer children



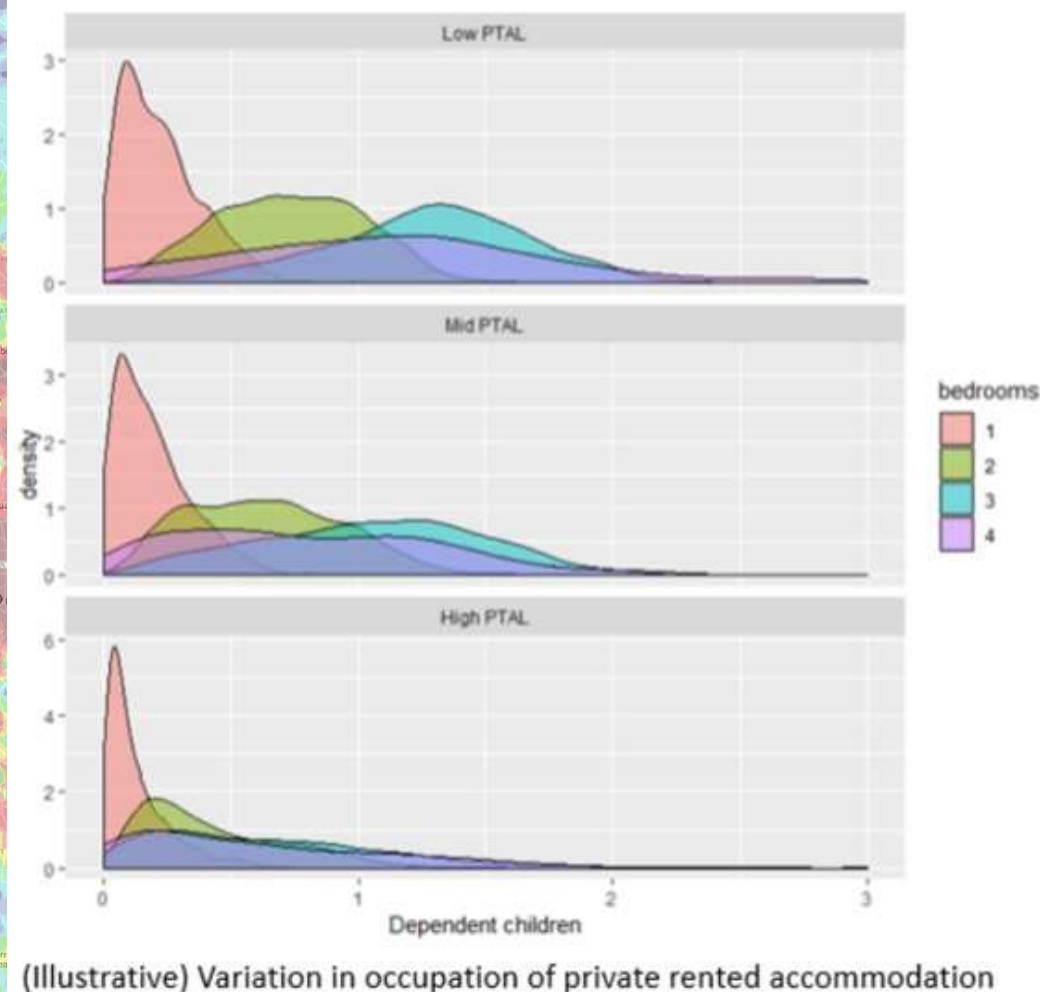
Proportion of adult house shares by transport connectedness





# Public transport connectivity

- London geography is more complex than distance from the city centre
- We've used the *traveltimes* API and TfL Public Transport Accessibility Index to look at the whole South-East of England
- House shares are more frequent in better-connected locations, and households have fewer children



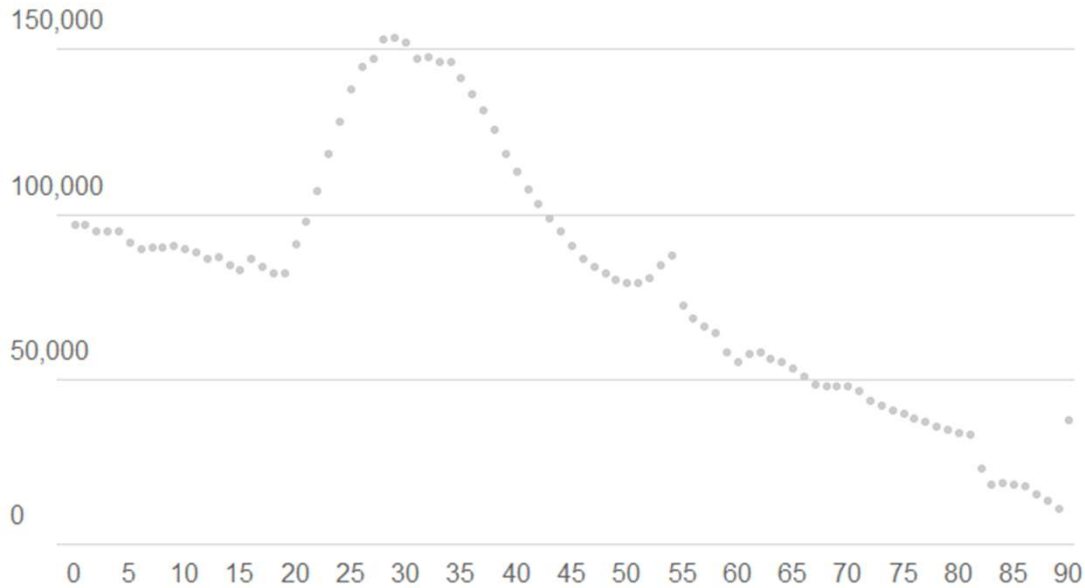
# Model implementation

- Reproducible
  - Version controlled code (coded in R)
  - Configuration & log files for each run
- Modular approach
  - E.g. births/deaths/migration methods can be swapped in and out
  - Allows for better experimentation, calibration and reusability
- Reports, outputs and visualisations
  - Automated quality assurance (for us and customers)
  - For onward use in reports/presentations
- Integrated testing
- Inputs/outputs via datastore API

# Automated outputs for QA

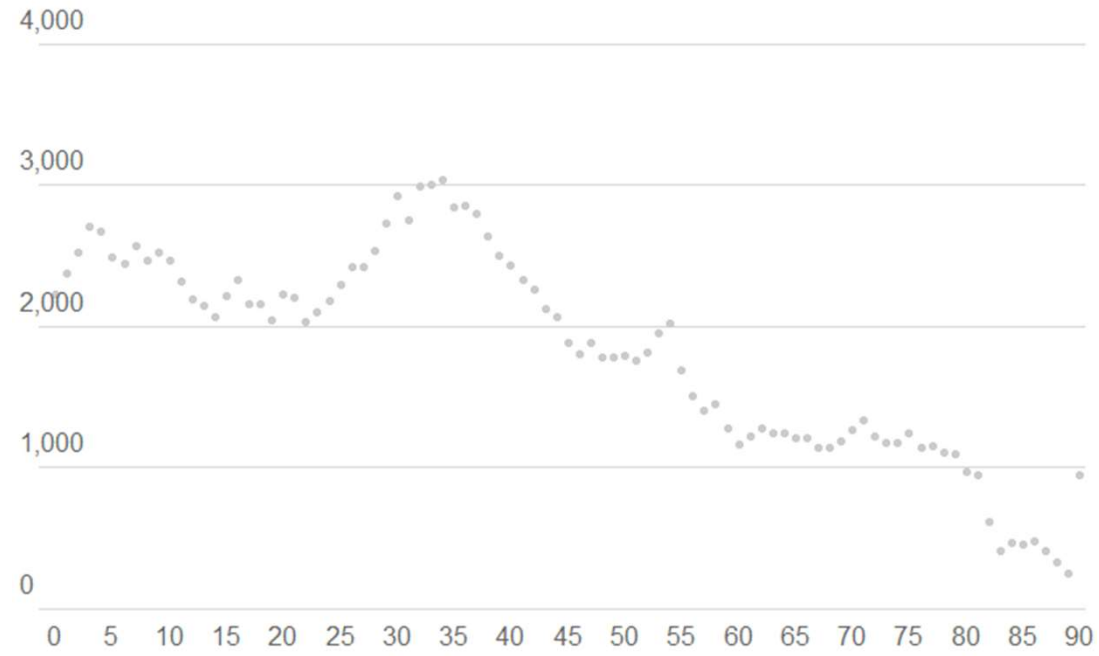
### London age structure over projection years

2001



### E09000002 age structure over projection years

2001



# Reproducible Analytical Pipelines

- Goals: reusability, accountability, collaboration, quality control



[https://ukgovdatascience.github.io/rap\\_companion/](https://ukgovdatascience.github.io/rap_companion/)

# Summary

- Getting London right at small areas is key to school and neighbourhood planning
- Not all data sources are reliable at small areas, but they can be combined to form a more reliable picture
- Planning data can be used to project changes to the housing stock
- But moving people into these new properties is a more complex problem
  - London is a rapidly changing, highly heterogeneous city
  - We are looking at new data sources to help us model internal migration
- Reproducible, adaptable models are key to creating reliable output



Thank you!

Questions?