Consumer Data Research Centre

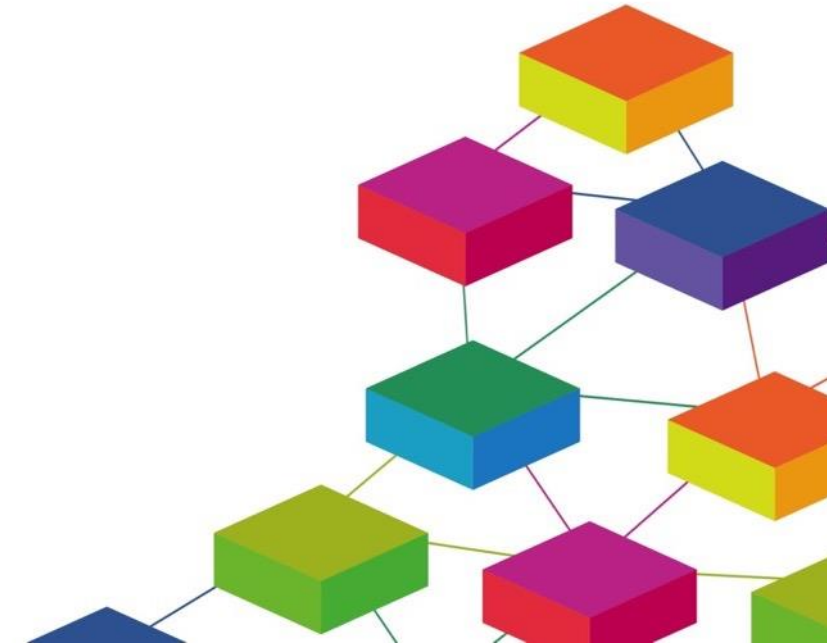
An ESRC Data Investment

UCL

Individual Level Model(l)ing and Simulation

Paul Longley and Tian Lan

Consumer Data Research Centre and Department of Geography
University College London
Department of Cartography and GIScience
Wuhan University



Comprehensive UK Data Agreements

Didobi

LDC The leader in retail location data & insight

Office for National Statistics

bps The British Population Survey

GambleAware

RT Registry Trust Ltd

pdv®

EDGE ANALYTICS

SALAD MONEY

FCA FINANCIAL CONDUCT AUTHORITY

huq

GEOLYTIX

SPEEDCHECKER

FUSION Data Science

OOMap

YouGov® What the world thinks

bike citizens

Domestic Energy Provider
High Street Retailer
Fast Food Restaurant Chain

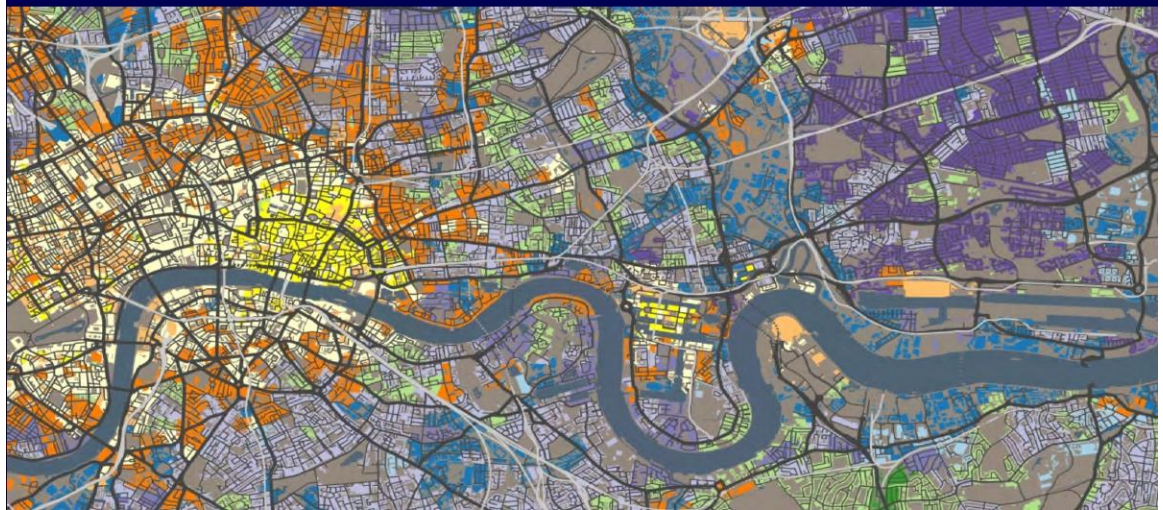
Regional Transport Provider
Youth Banking Card Provider
Family Restaurant Chain

Geospatial Analysis

A Comprehensive Guide to Principles
Techniques and Software Tools

7th edition; Updated June 2024

Dr Michael J de Smith, Prof Michael F Goodchild
Prof Paul A Longley & Associates



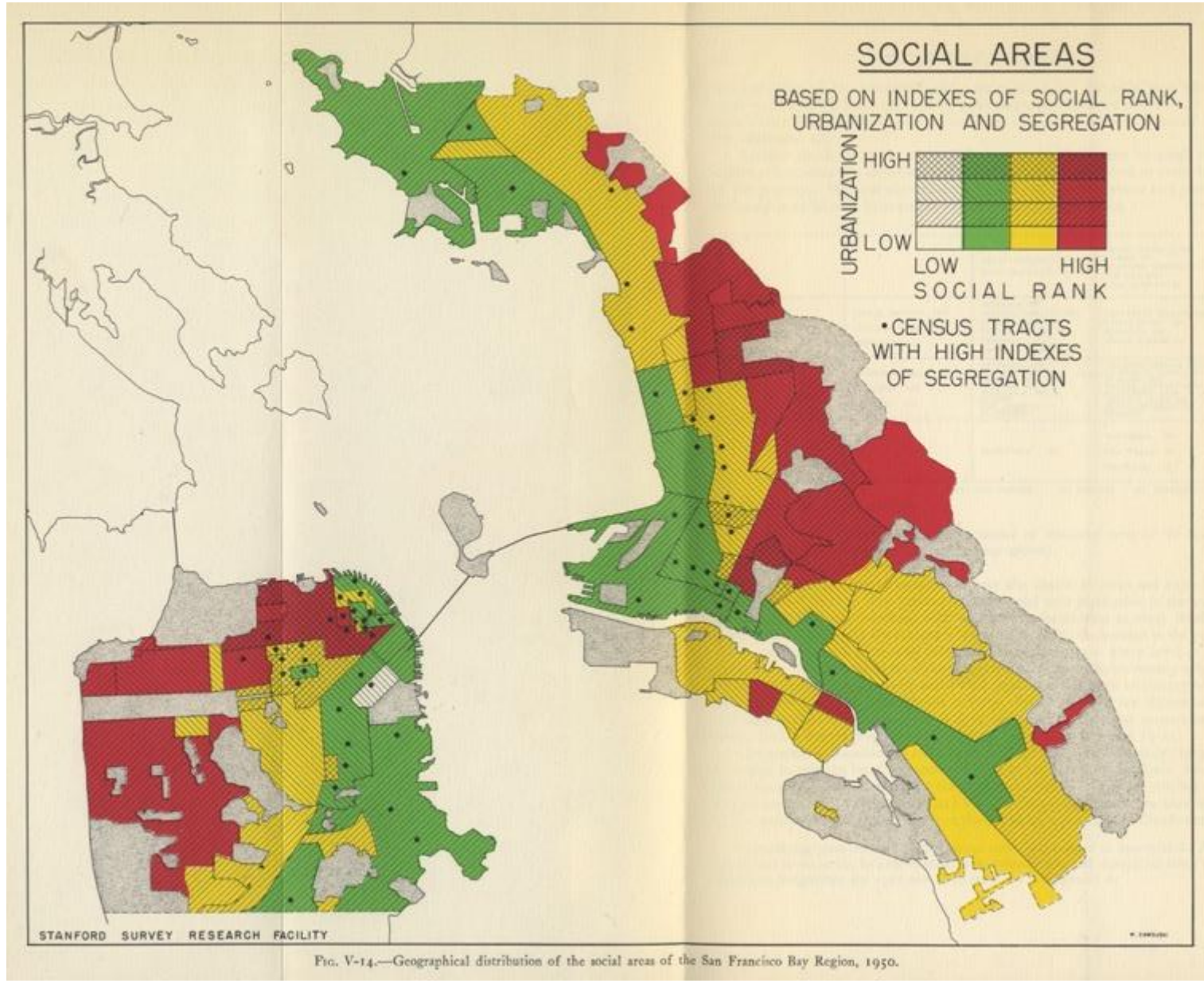


FIG. V-14.—Geographical distribution of the social areas of the San Francisco Bay Region, 1950.

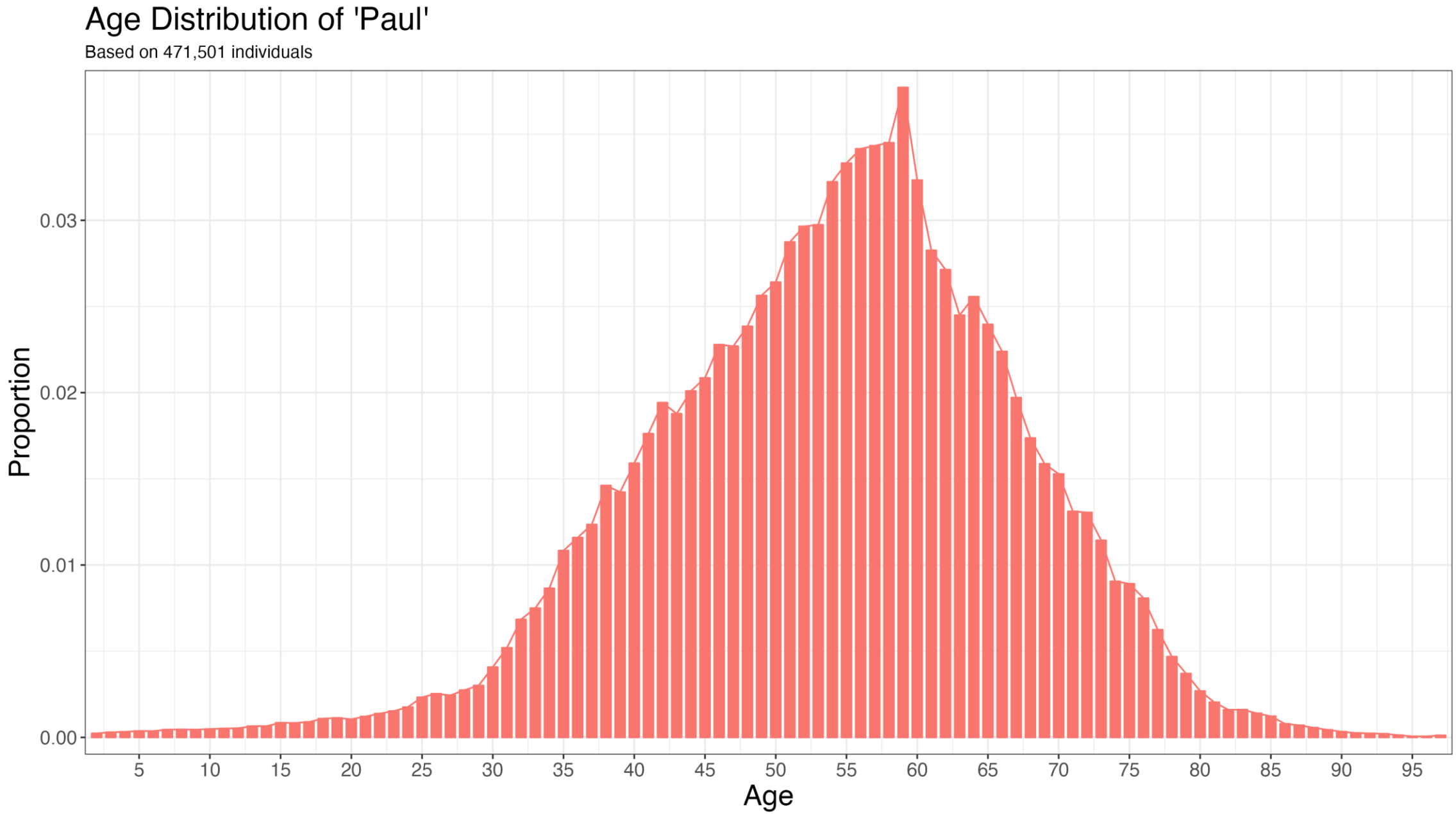
19971998 1999 2000
2001 2002 2003
2004 2005 2006
2007 2008 2009
2010 2011 2012
2013 2014 2015
2016 2017 2018
2010 2020 2021
2022 2023 2024

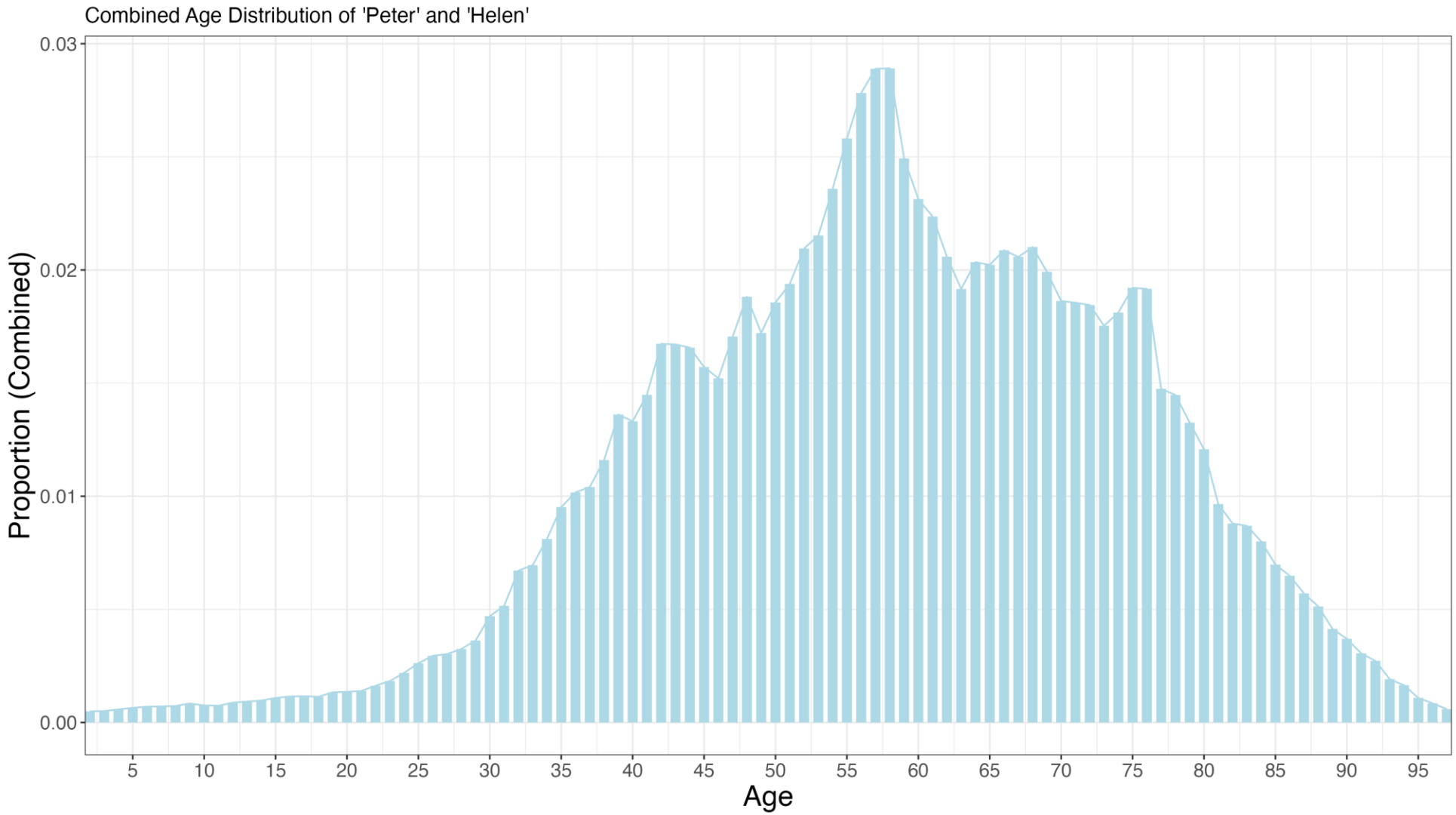
The linked consumer registers contain the ***individual*** names and addresses of adults in the UK in annual snapshots between 1997 and 2024.

https://data.cdrc.ac.uk/dataset/linked-consumer-registers-1997-2016

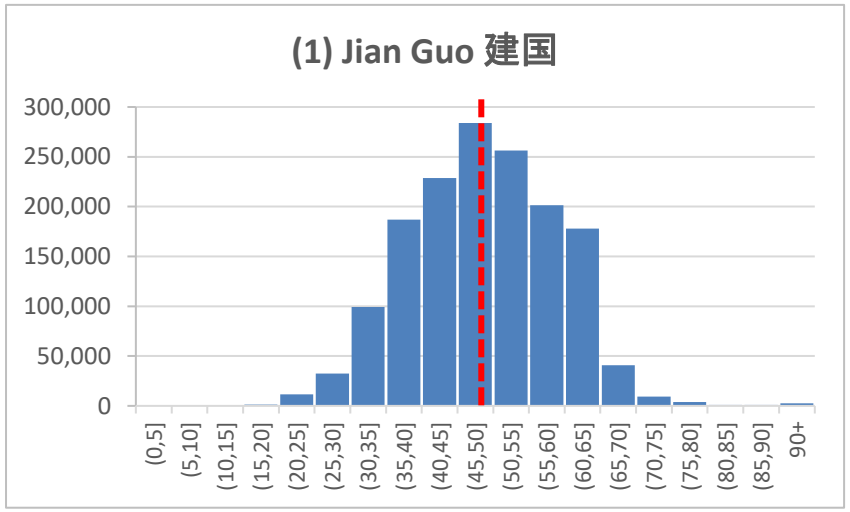
Gender

Data Sources: Birth certificate records (ONS) & Market Data (CACI).
Produced by Guy Lansley, UCL

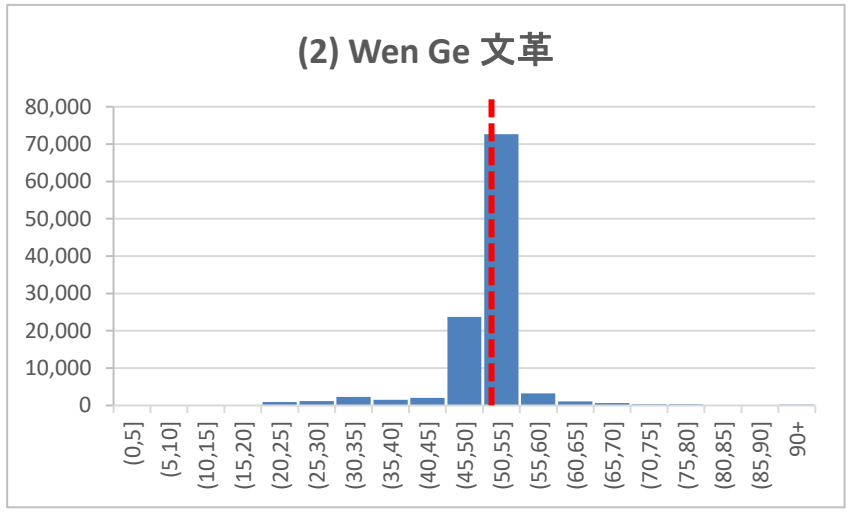

Age Distribution of 'Paul'
Based on 471,501 individuals
Proportion
0.03
0.02
0.01
0.00
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
Age


Combined Age Distribution of 'Peter' and 'Helen'
Proportion (Combined)
0.03
0.02
0.01
0.00
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
Age

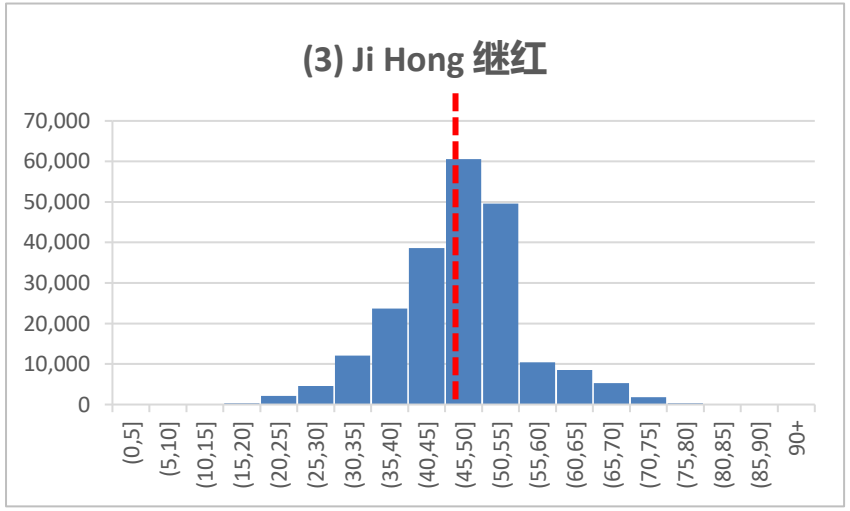
Popular forenames 1950s and 1960s



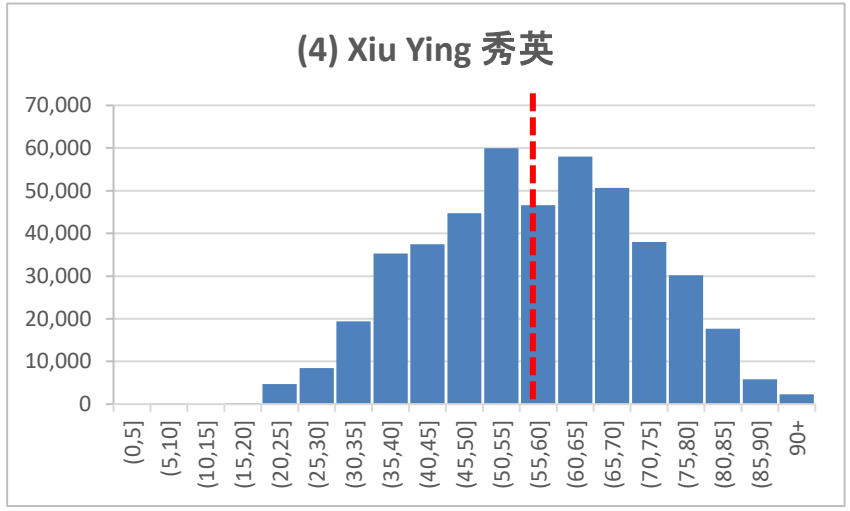
N= 1,538,712; Avg.= 49; Skewness=0.07; Kurtosis=0.33



N= 110,263; Avg.= 51; Skewness=-0.66; Kurtosis=14.6



N= 218,041; Avg.= 47; skewness=0.05; kurtosis=1.12

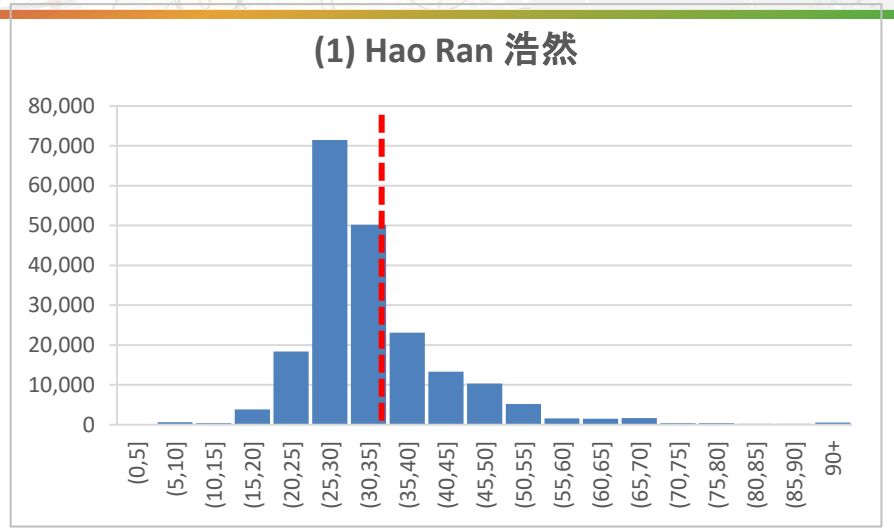


N= 459,654; Avg.= 57; skewness=-0.07; kurtosis=-0.59

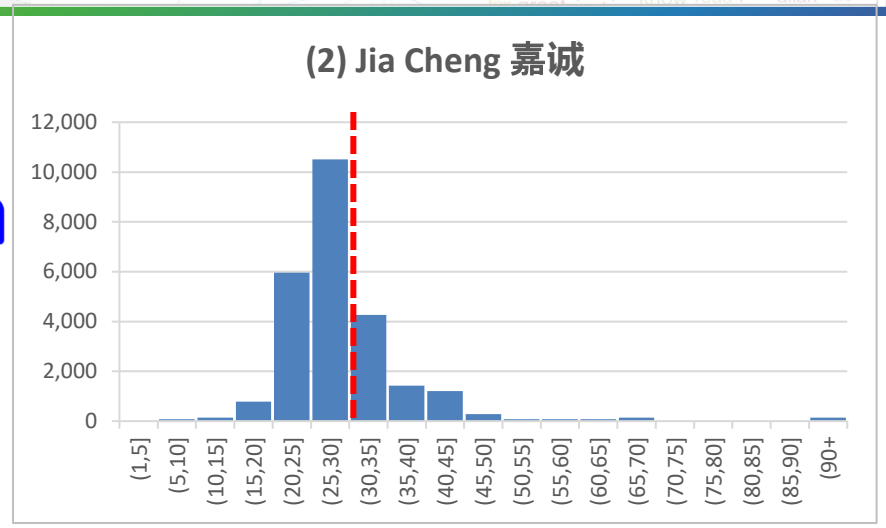
Consumer Data Research Centre

An ESRC Data Investment

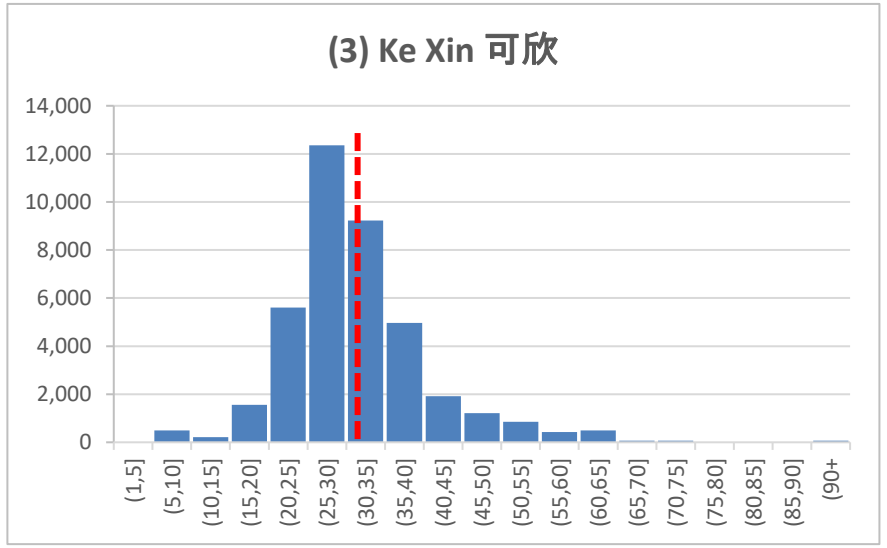
Popular forenames in 1980s and 1990s



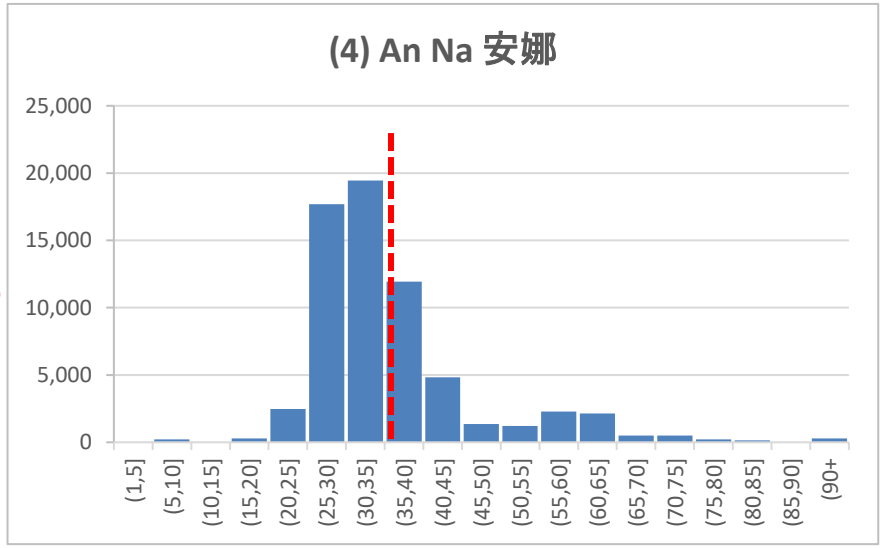
N= 203,486; Avg.= 34; skewness=2.06; kurtosis=7.38



N= 25,134; Avg.= 30; skewness=3.6; kurtosis=21.7



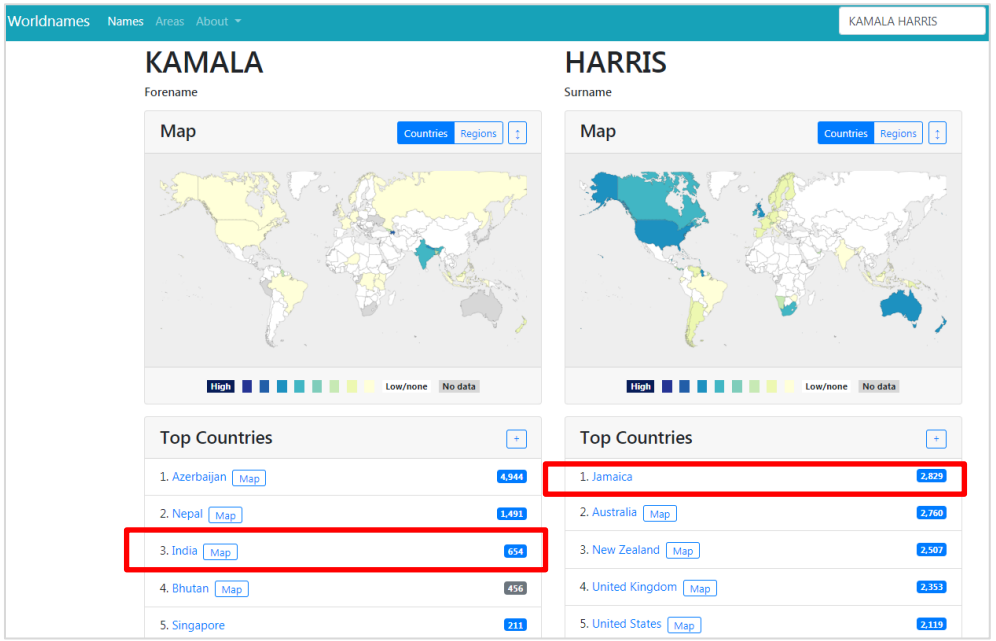
N= 39,547; Avg.= 32; skewness=1.5; kurtosis=5.5



N= 65,462; Avg.= 36; skewness=1.9; kurtosis=5.1

Names-based ethnicity classifications

- **Onomap3**
  - Worldnames2 apps.cdrc.ac.uk/worldnames/ (O'Brien & Longley, 2018)
  - Census groups & country origins (Lan & Longley, 2022)



Forename and surname Frequency per Million (FPM) by countries
(Source: https://apps.cdrc.ac.uk/worldnames/)

Forename component Surname component

$$S_E = w_f * p_{E,f} + w_s * p_{E,s}$$

Where

S_E : a name pair's composite score for country E;

$p_{E,f}$: FPM of a forename for country E;

$p_{E,s}$: FPM of a surname for country E;

w_f : forename weighting factor;

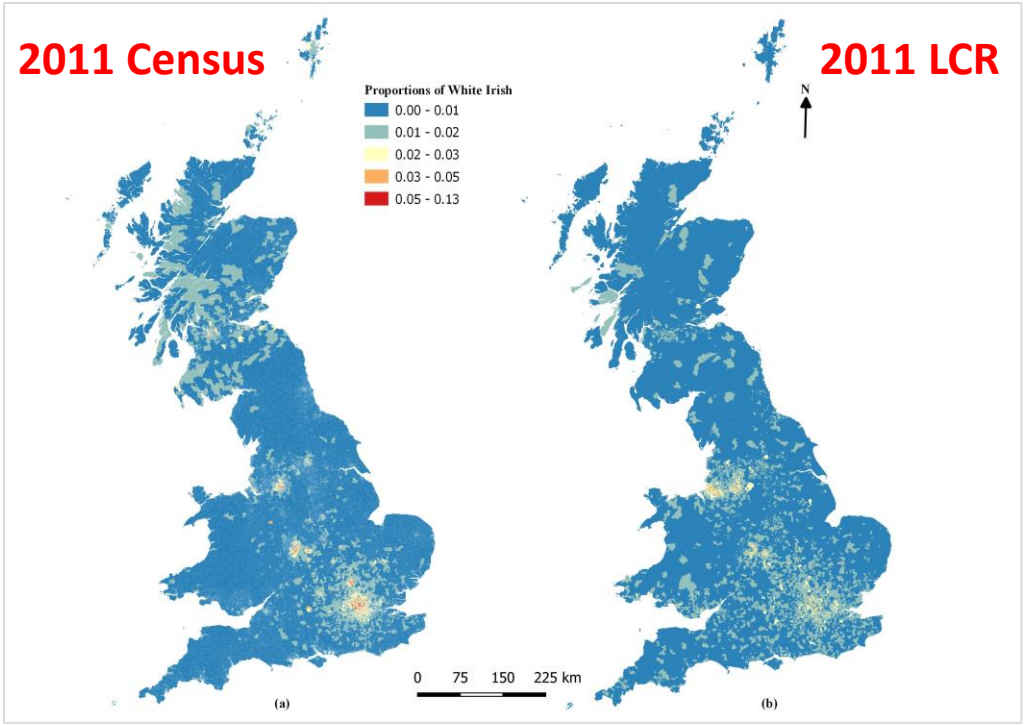
w_s : surname weighting factor.

Individual basis to ‘digital twins’

- **Digital twin validation**
  - 2011 LCR vs. aggregate 2011 Census ethnic data

Group	2011 Census	2011 LCR	Ratio
AAO	663,124	604,428	**91%**
ABD	294,505	285,953	**97%**
ACN	371,521	230,604	**62%**
AIN	1,167,436	1,354,202	**116%**
APK	788,849	779,968	**99%**
BAF	713,257	606,049	**85%**
BCA	496,195	377,245	**76%**
OXX	1,298,097	608,318	**47%**
WAO	2,286,231	2,775,057	**121%**
WBR	41,245,227	39,581,857	**96%**
WIR	551,410	546,743	**99%**
Total	49,875,852	47,750,424	**96%**

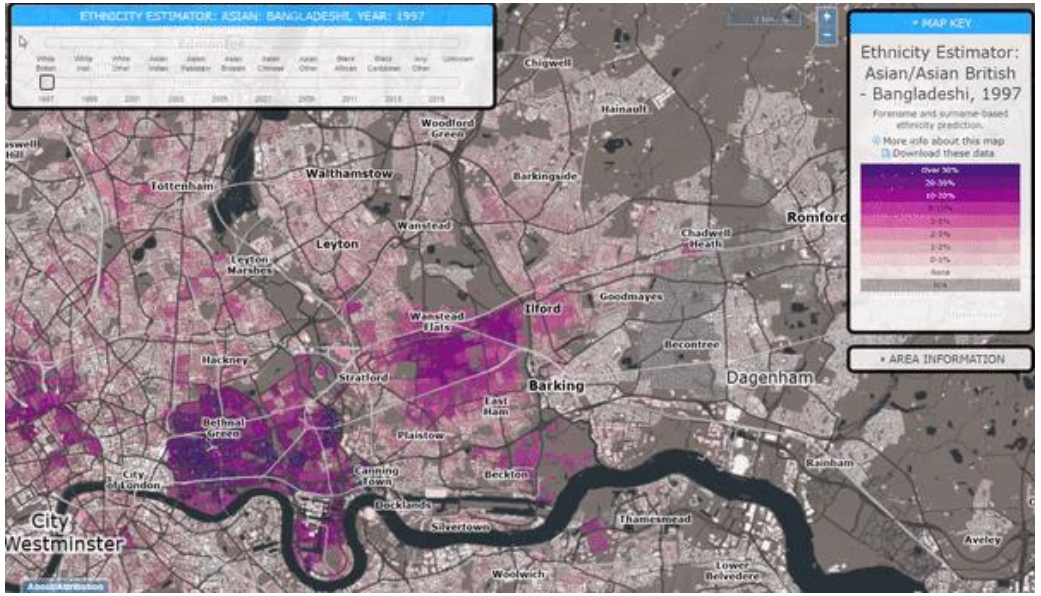
Comparison of populations by ethnic groups between the 2011 Census and 2011 LCR



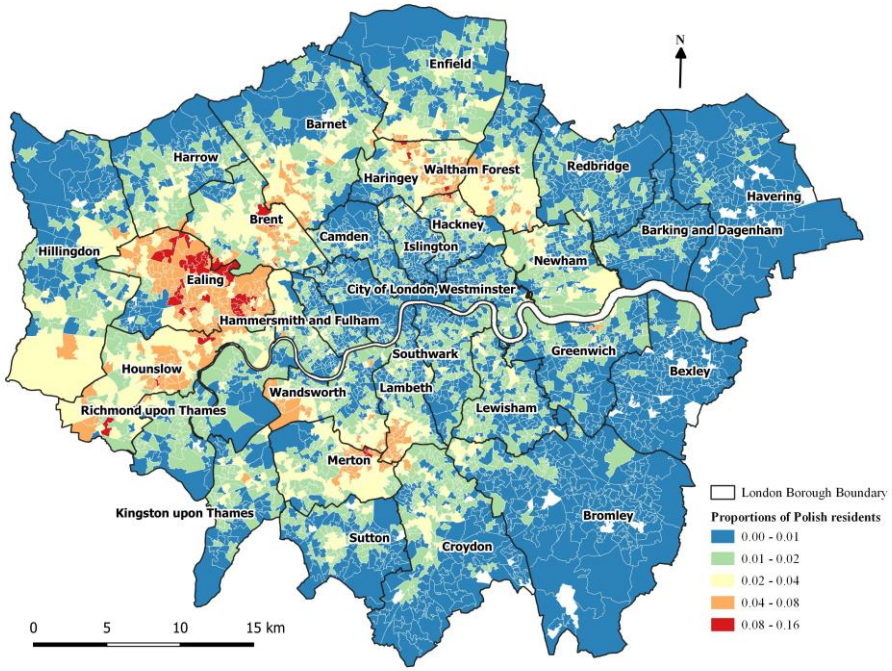
Proportions of White Irish population by LSOA in the 2011 Census and 2011 LCR

Prediction using individual level estimates

- **Annual estimates and more granular categories**

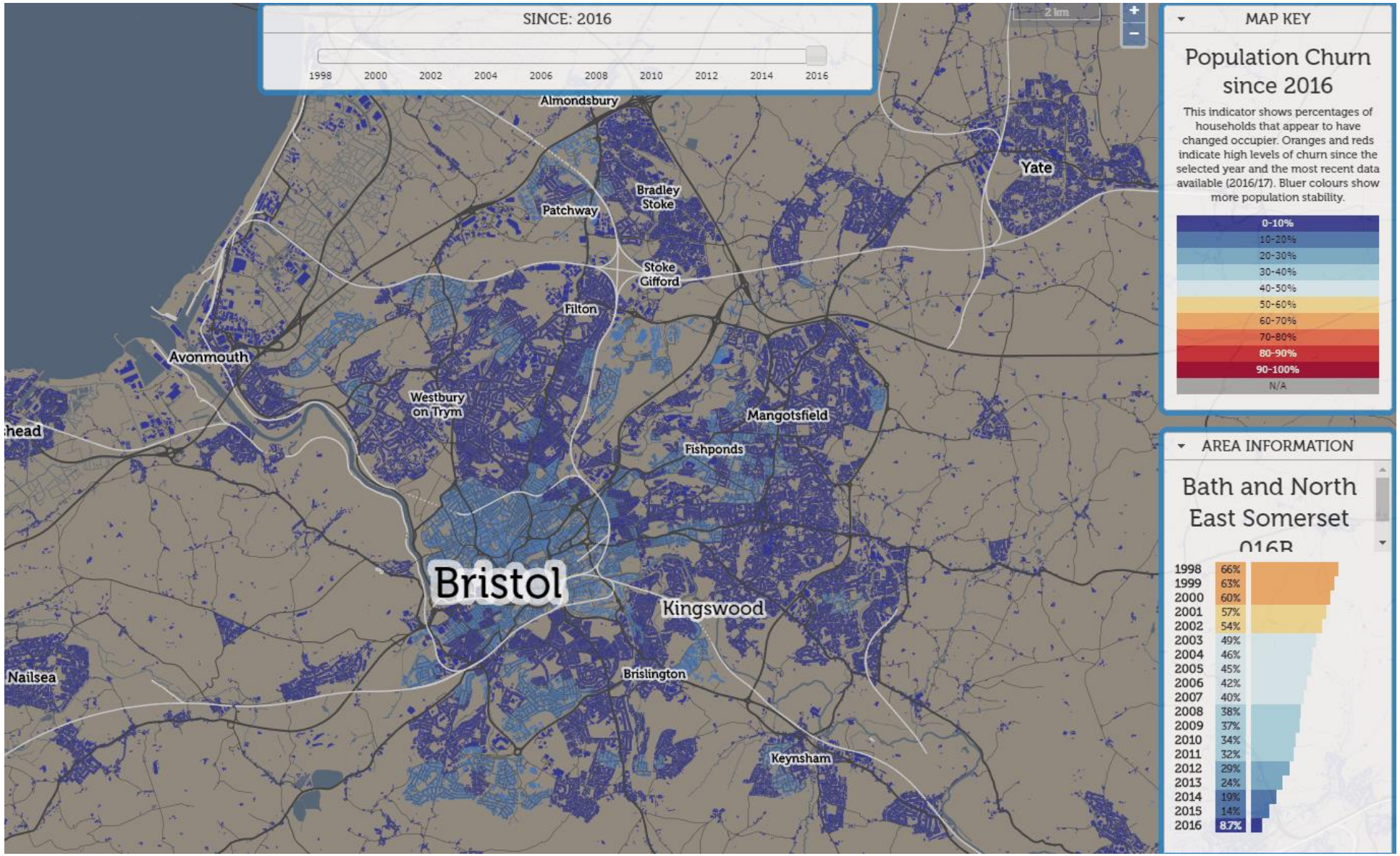


Bangladeshi community in East London moving further eastbound during 1997 - 2016



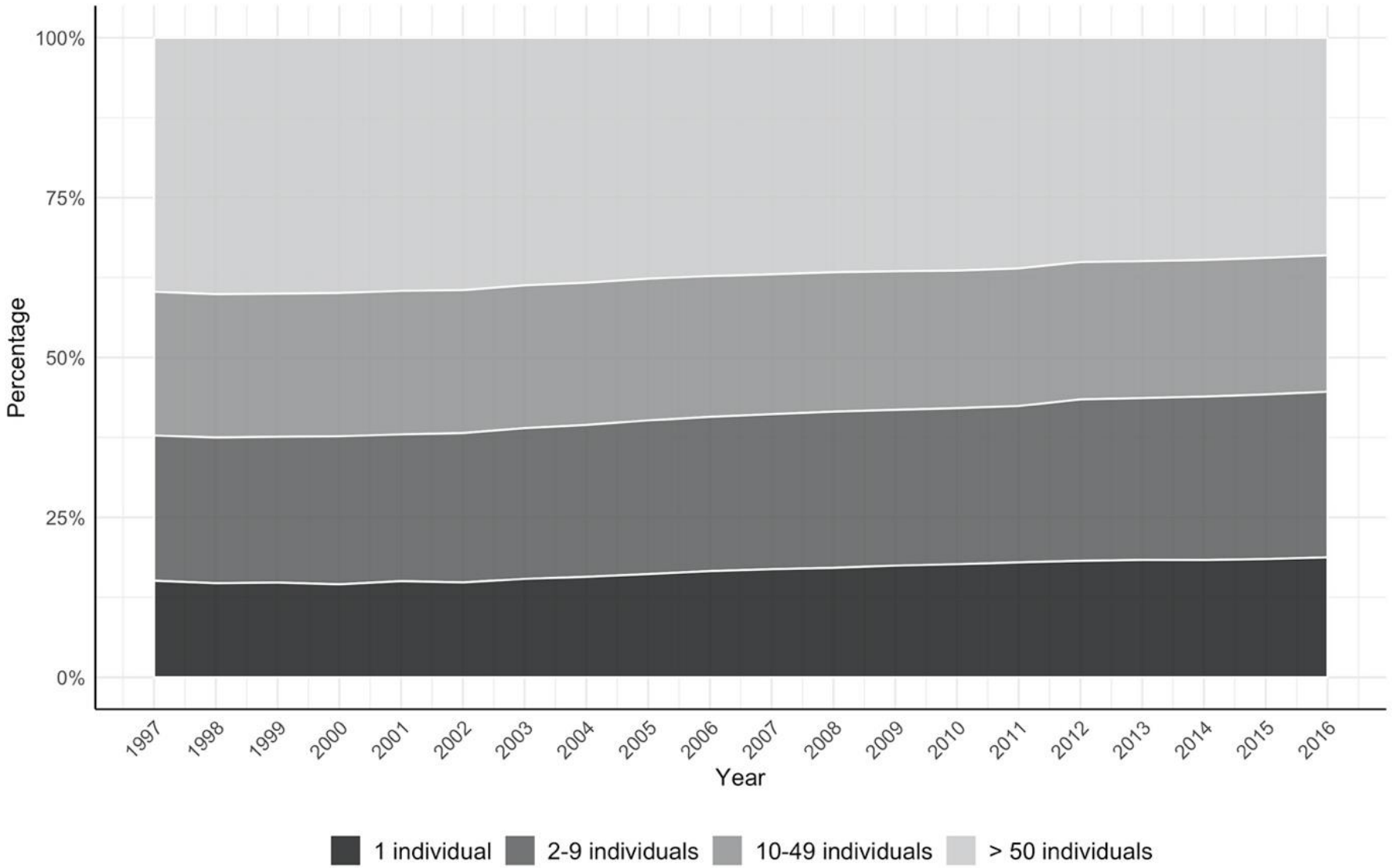
Distribution of Polish residents in London (estimated from the 2011 LCR)

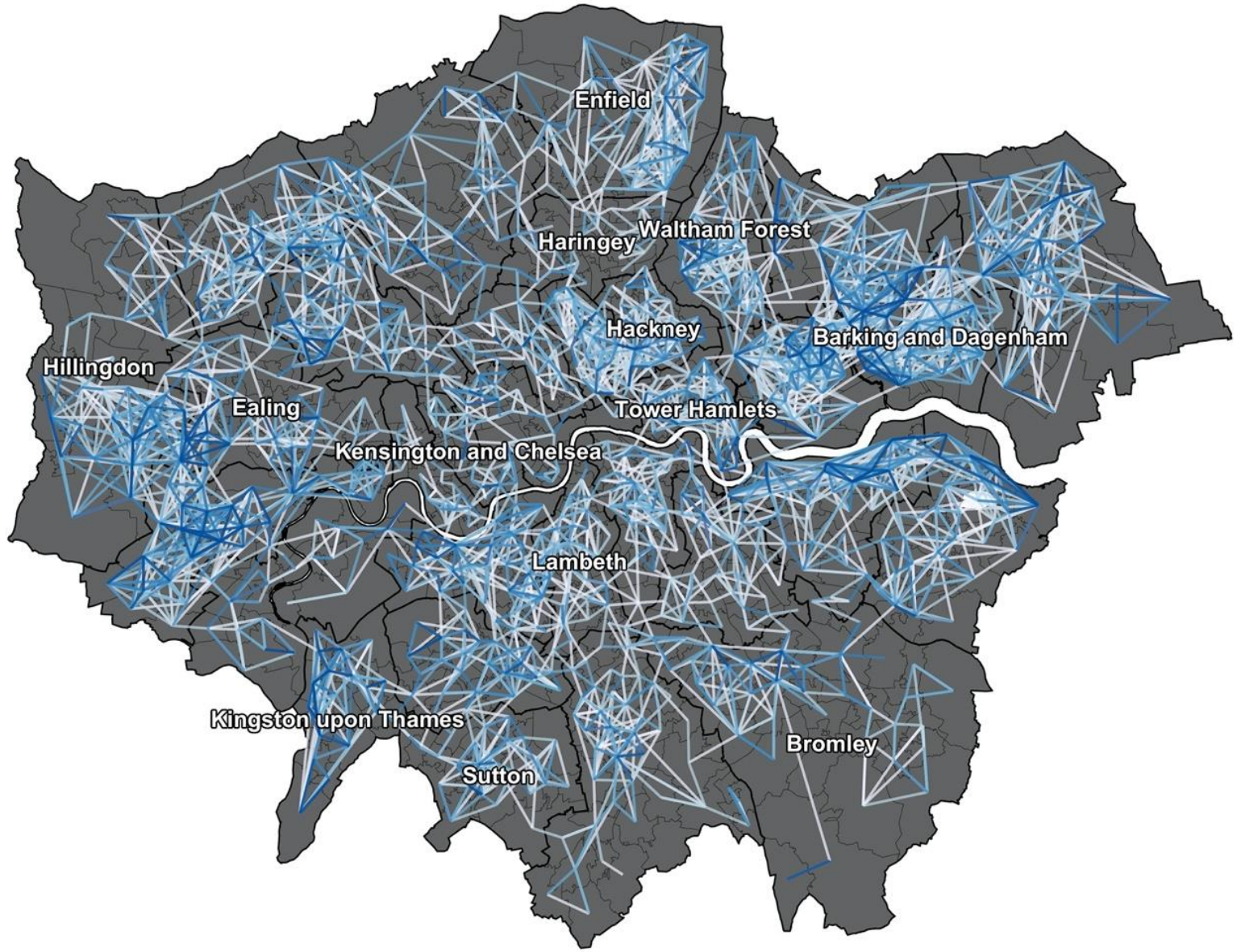
Annual estimates of population churn



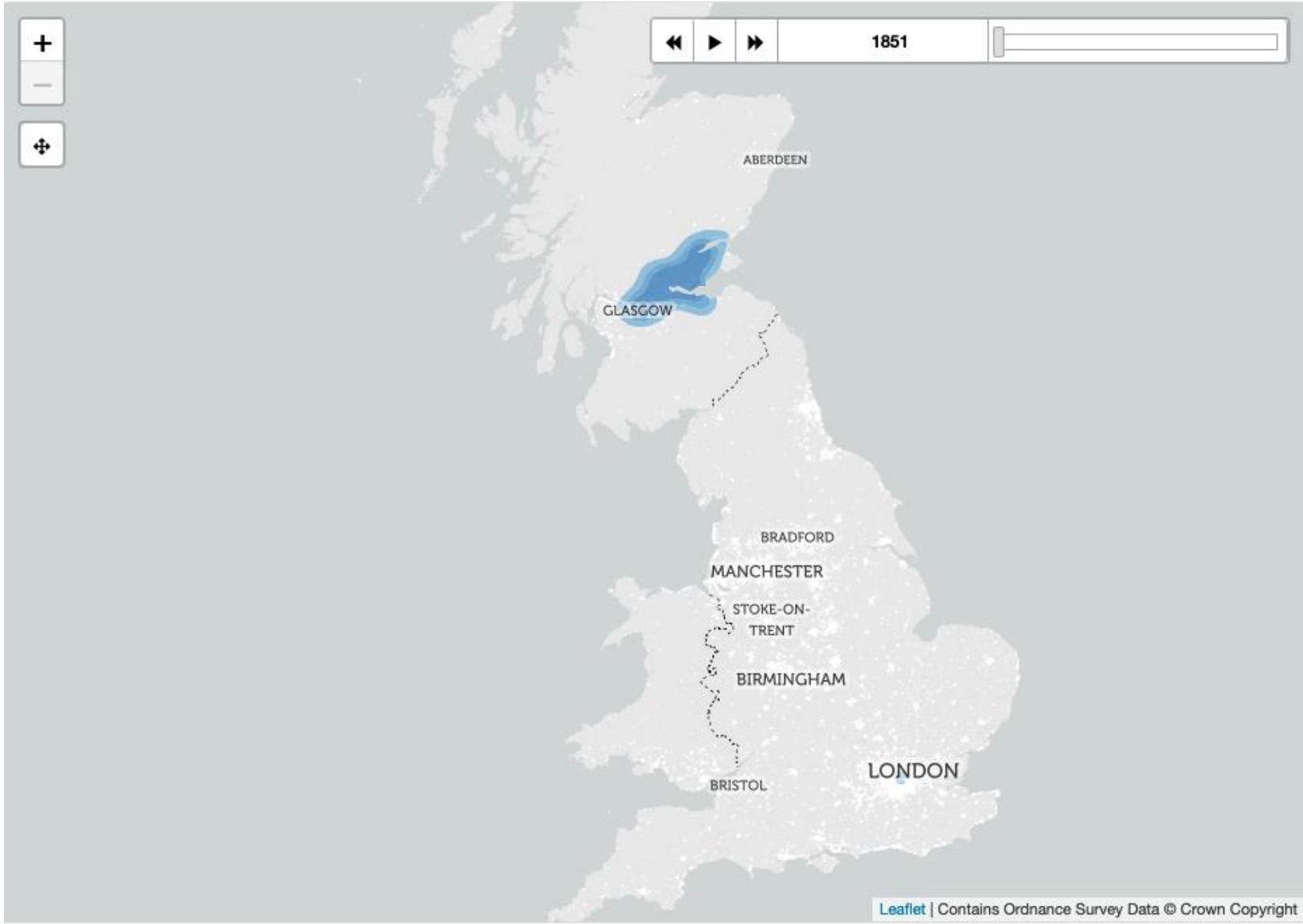
https://maps.cdrc.ac.uk/#/indicators/churn/

Percentages of the adult population captured in the LCRs bearing forename – surname pairings that occur 1, 2-9, 10-49 and 50 or more times over the period 1997-2016



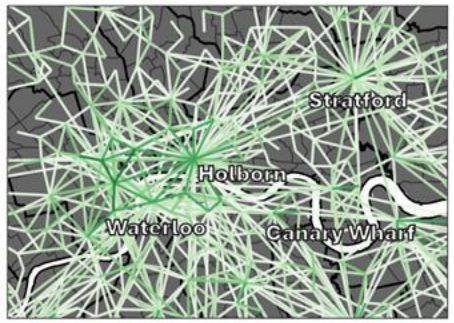

Enfield
Haringey
Waltham Forest
Hackney
Barking and Dagenham
Hillingdon
Ealing
Tower Hamlets
Kensington and Chelsea
Lambeth
Kingston upon Thames
Bromley
Sutton

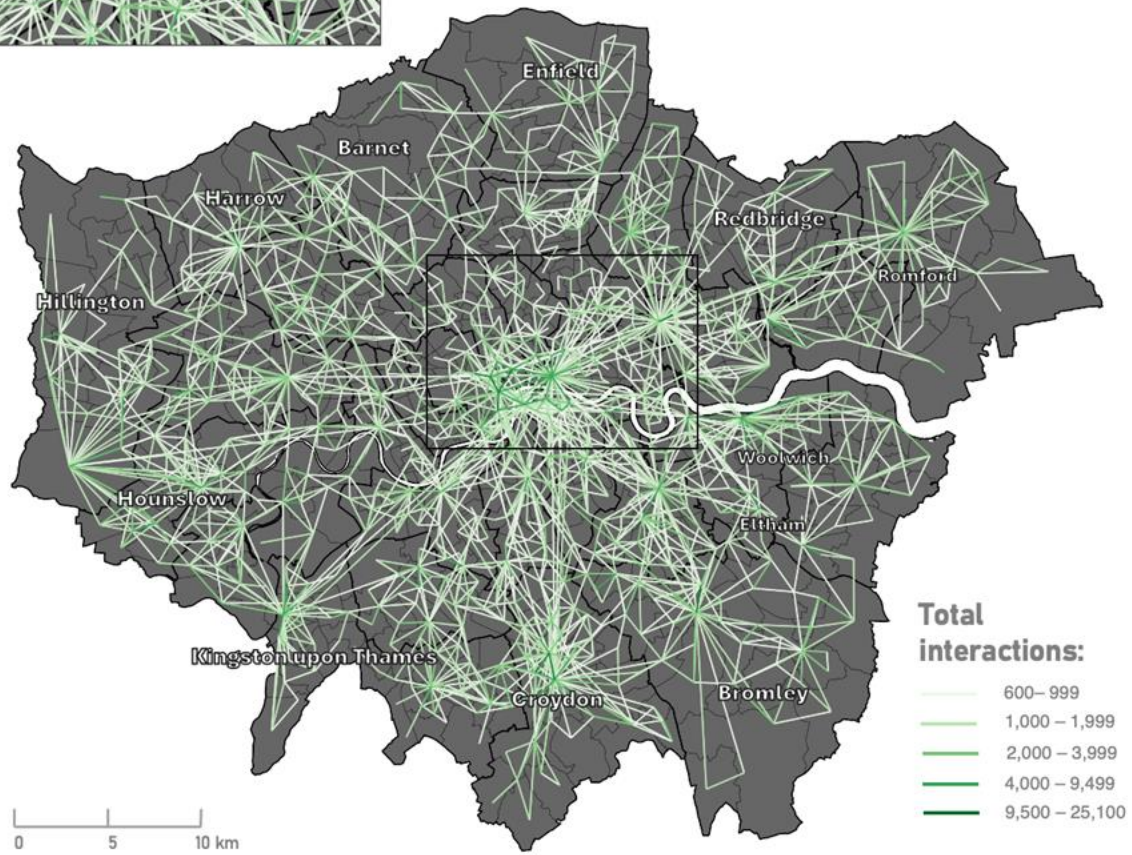
FOTHERINGHAM



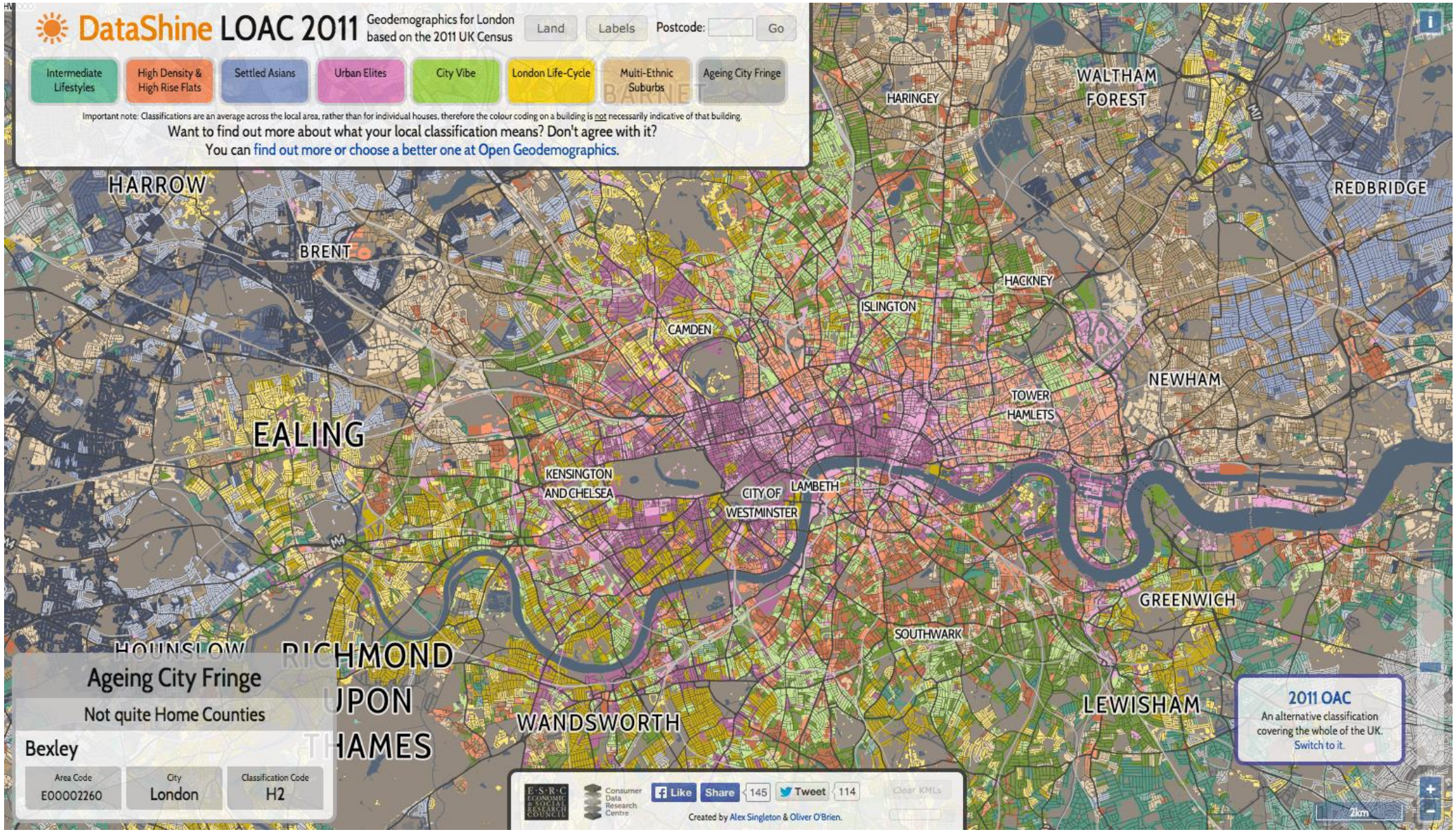
The blue contours enclose areas where bearers of the name were most concentrated, measured as where the population-weighted density of the surname is highest. Use the slider bar to see where the name was found over historical (1851, 1861, 1881, 1901, and 1911) and more recent (1997, 2006 and 2016) time periods.

Non-directional OD Flows between each MSOA in Greater London, 2019 (minimum threshold = 600 interactions).





Courtesy: Mikaella Mavrogeni



London Output Area Classification