



Using administrative data to estimate local service needs

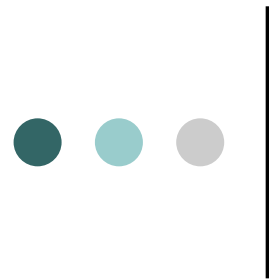
Les Mayhew

Cass Business School

Mayhew Harper Associates Ltd.

Lesmayhew@googlemail.com



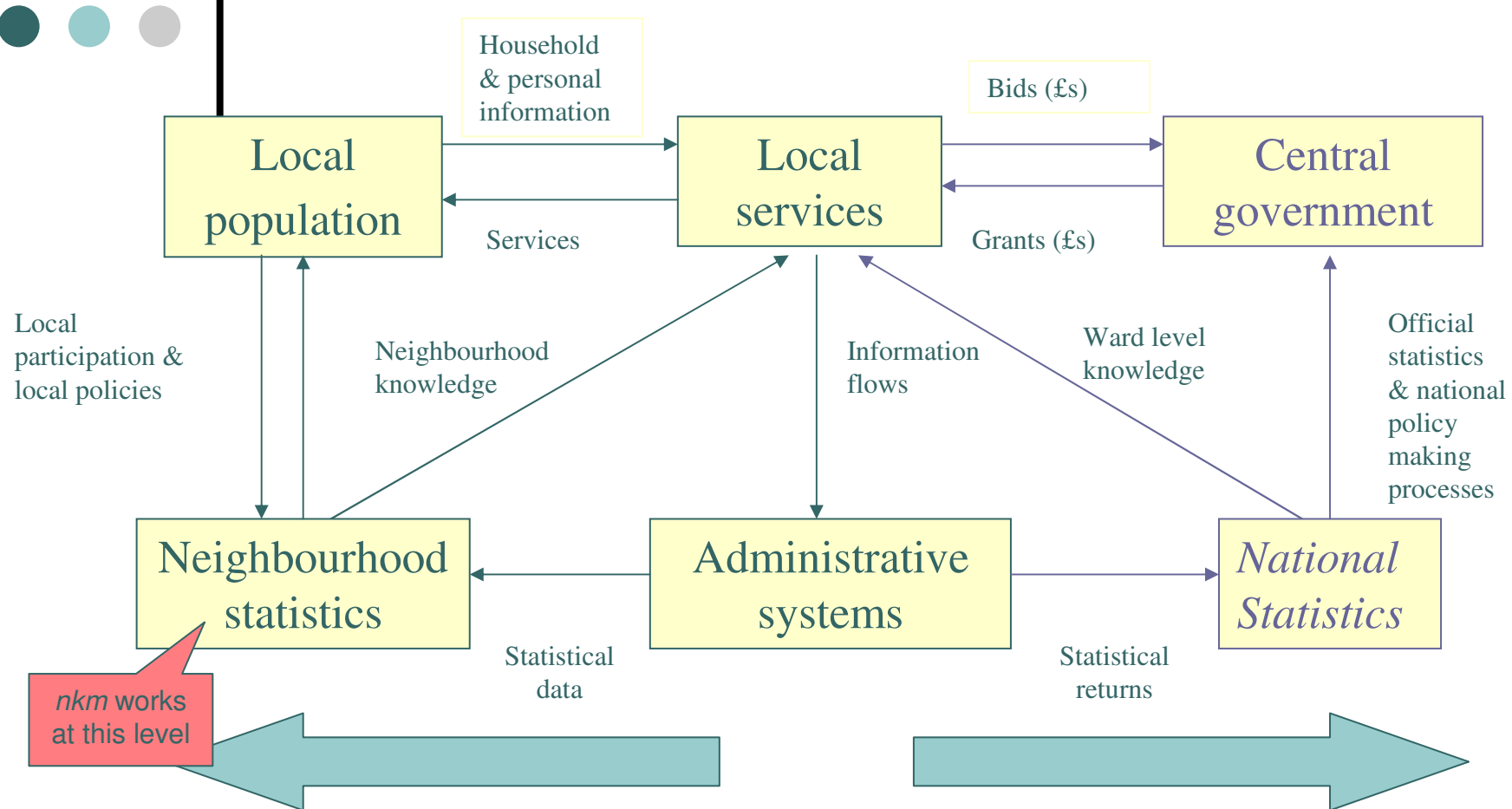


Outline

- Sources and uses of administrative data
- Stages in the use of administrative data
- Households versus individuals
- Ethnicity assignment
- Applications
 - Housing regeneration
 - Health care consumption
 - Religious demography
 - Troubled families



Where the information comes from



- re-generation
- targeting services
- resource allocation
- local policy and planning

- statistical returns
- bids for money
- meeting national targets
- publicity





The benefits of using administrative data

- More timely population intelligence to support council and PCT
- Less reliance on external data sources, product licences, duplication etc.
- More responsive analytical services using better quality data
- Data sharing with local partners and joined up working
- Better data management across council and partner organisations and so cheaper
- More efficient services, better outcomes



FAQs that elude official data sources

- What is the population of my community council area?
- How many single parents live in social housing and are on benefits?
- How many 75+ older households are asset rich and income poor?
- How many nurseries are there within pram pushing distance of households with young children?
- What is the social profile of newcomers and what are their needs for services?
- How do services segment in terms of socio-economic profile and uptake?
- Are services accessible to those that need them and how much unmet demand is out there?
- Who needs to have face to face contact and where should caller centres be located?
- Are there special groups that need more personalised services and how many are there (e.g. older people, single parent households, ethnic groups)?

Mayhew Harper Associates Ltd





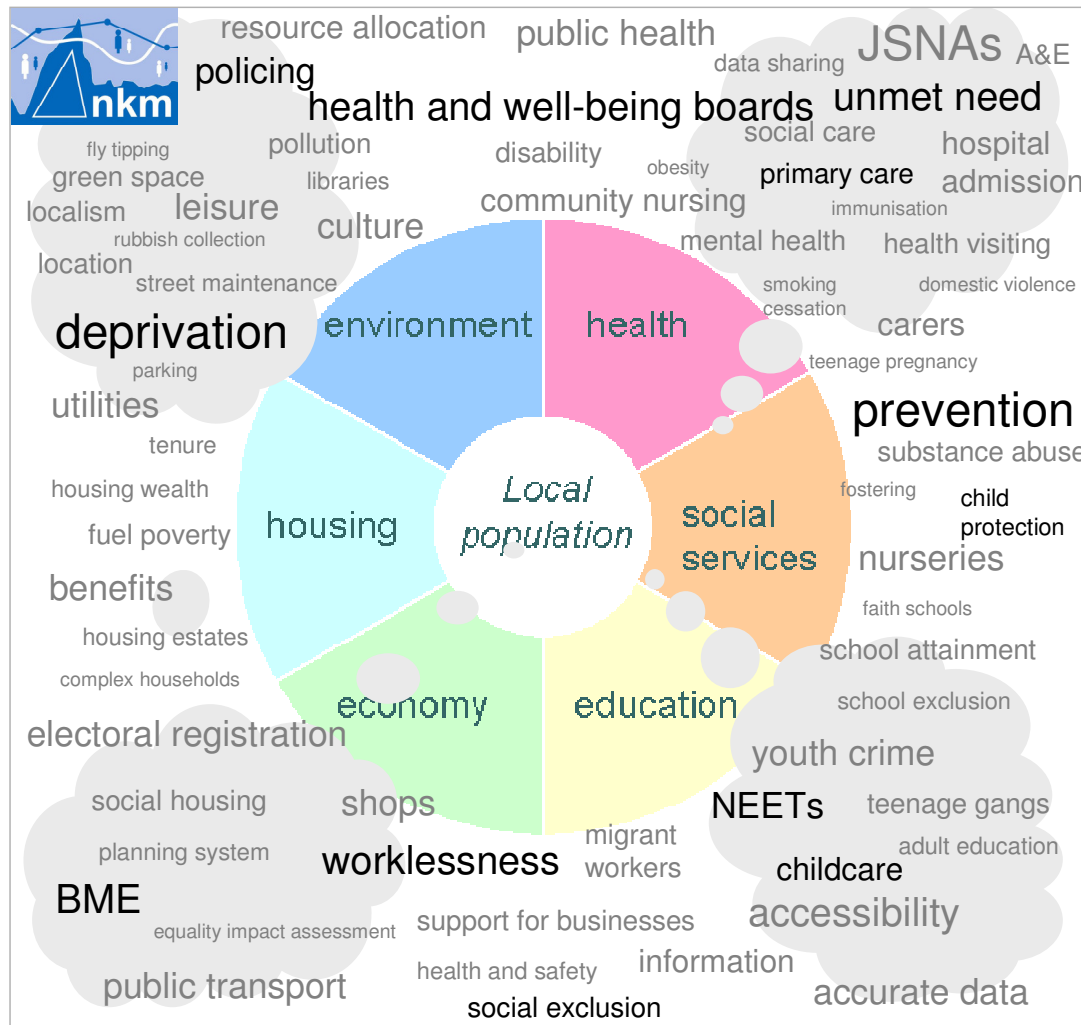
Information silos in health and local government



Information silos in health and local government



Applications based on joined up data

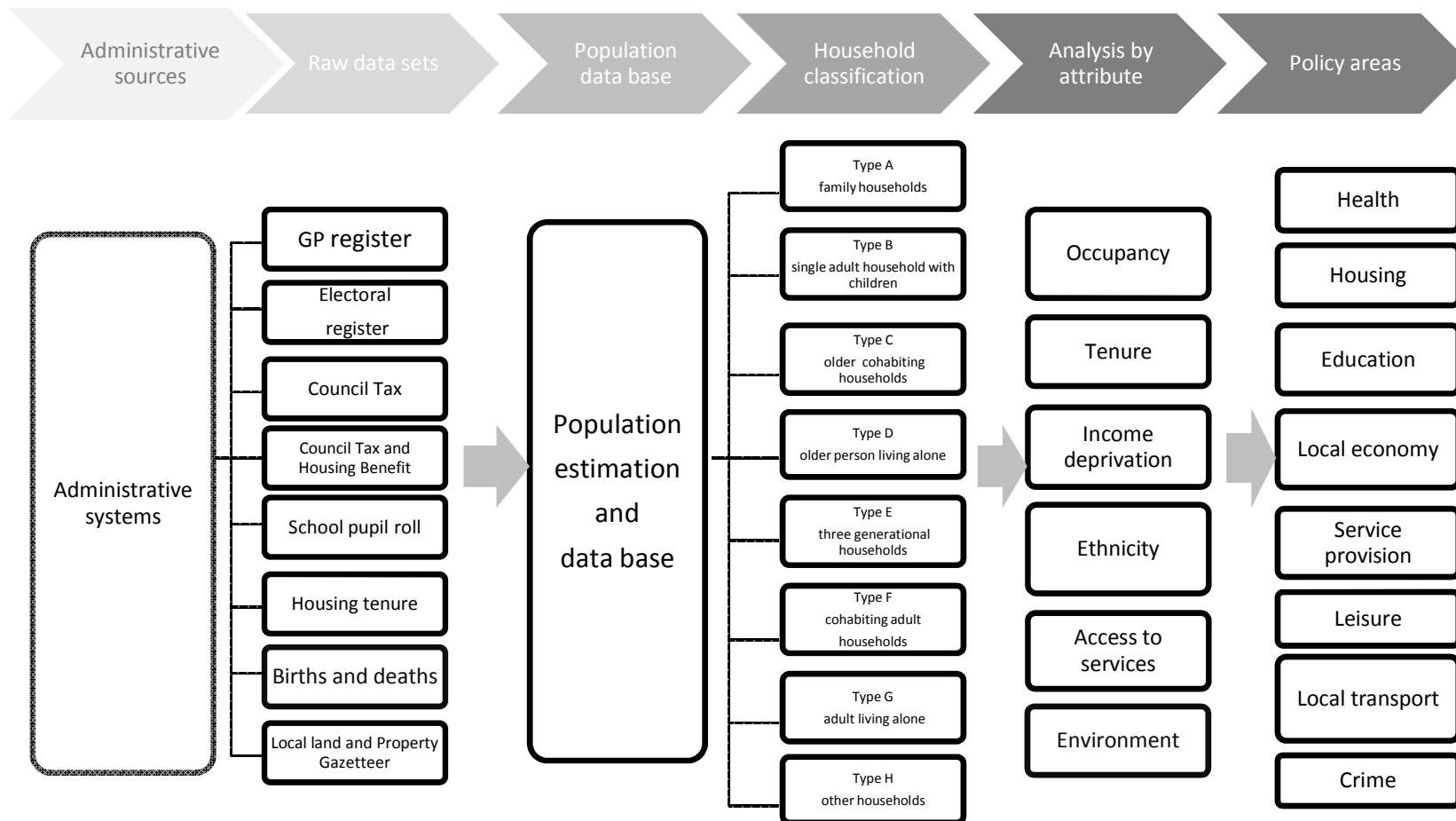


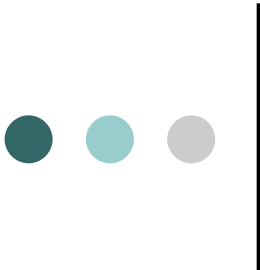
Examples of applications using *nkm* approach over past 10 years





From raw data to applications





Attributes can be assigned to individuals, households or properties

Some examples

- Income poverty
- Tenure
- Occupancy
- Housing wealth
- Age of house
- Ethnicity
- Consumption of services
- Religion
- Crime
- Single sex households
- Married households
- Other sub-classifications





Further considerations (1): Working with households

- *“A household is: one person living alone; or a group of people (not necessarily related) living at the same address, who share cooking facilities **and** share a living room, sitting room, dining room or kitchen.”*

Source: **Helping to shape tomorrow**
The 2011 Census of Population and
Housing in England and Wales





Household classification system

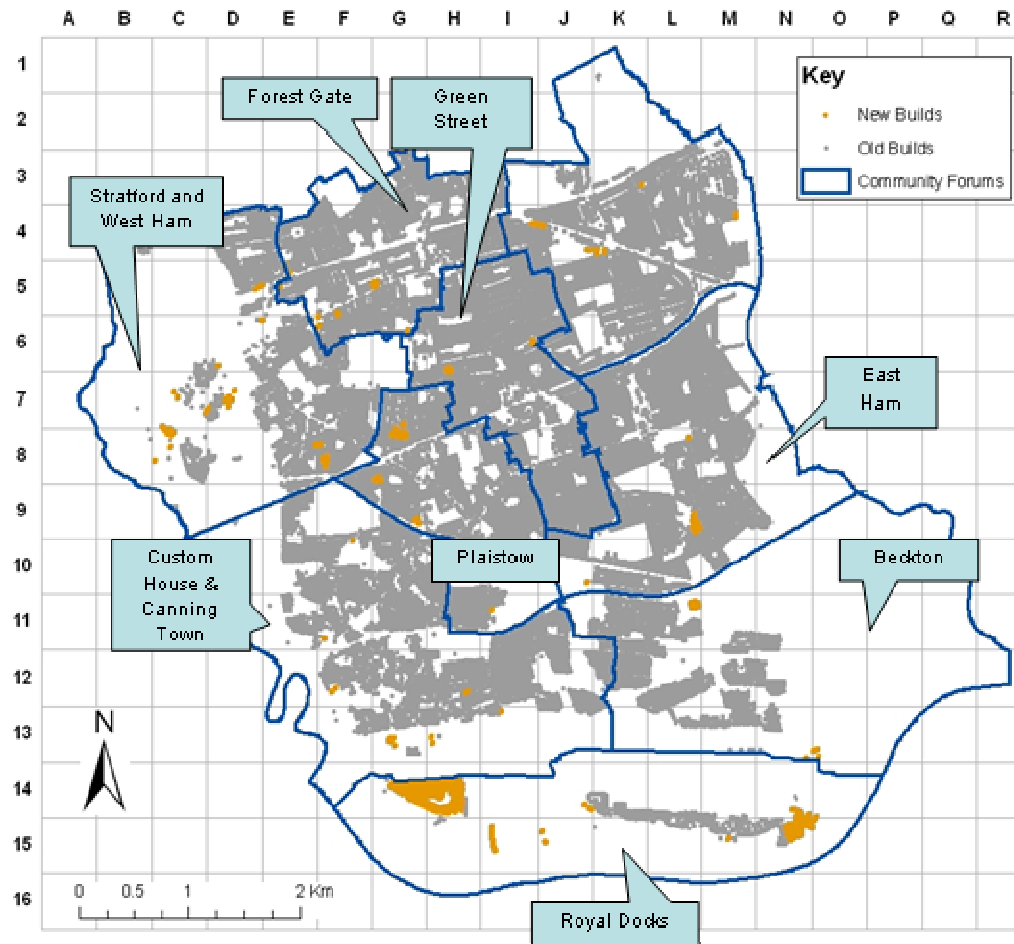
category	description
A	family households with dependent children
B	single adult households with dependent children
C	older cohabiting person households
D	older persons living alone
E	three generational households
F	cohabiting adult households no children
G	single adult households
H	other households

First tier household classification based on detailed household demography and sub-types

Type	age group 1	age group 2	age group 3	size	Description
A	00--	00	----	4	Couple household with two children
B	-0--	0--	----	2	Single adult household with one child
C	----	0--	0--	2	Older couple household with one person aged 65+
D	----	----	0--	1	Older person living alone
E	--0-	00--	0--	4	3-generational with one child ,couple and an older person
F	----	000-	----	3	Cohabiting adult household
G	----	0--	----	1	Adult living alone
H	00--		00--	4	Split generation household
H'	0000	----	----	4	Young household (e.g. students, teenage parent)



Households by age of build



Aimed to find out if new properties being occupied by higher income people as key to regeneration

Based on Newham 2007

Are new builds occupied by higher income households?

Row	number of households	new build	Council Tax Band > A	adults no children	private tenure	% of h'holds not on benefits
1	733	Y	Y		Y	99.7
2	3,286	Y	Y	Y	Y	90.7
3	154	Y		Y	Y	90.3
4	34,508		Y	Y	Y	74.6
5	37,759		Y		Y	67.1
6	601				Y	66.6
7	44	Y			Y	63.6
8	1,942			Y	Y	58.2
9	316					48.4
10	1,301			Y		37.3
11	7,795		Y			34.1
12	7,560		Y	Y		33.7
13	347	Y	Y			0.3
14	1	Y		Y		0.0
total	96,347	4,565	91,988	48,752	79,027	11.4

Based on further analysis, new builds are:

- 2.7 times less likely to be on benefits than old builds
- 1.3 times more likely to be tax band B and above
- 1.3 times less likely to have children living in them
- 4.8 times more likely to be privately owned





Further considerations (2): Ethnicity assignment system

- We use actual data to assign ethnicity when available (e.g. School census data, HES)
- Data based on self reported ethnicity
- We infer the ethnicity of others living in children's households based on children's known ethnicity
- For other households we use name recognition based on extensive database
- About 80%-90% of population can be assigned ethnicity in this way
- Results presented in form of 3-tier classification system
- Household ethnicity assigned based on dominance principle



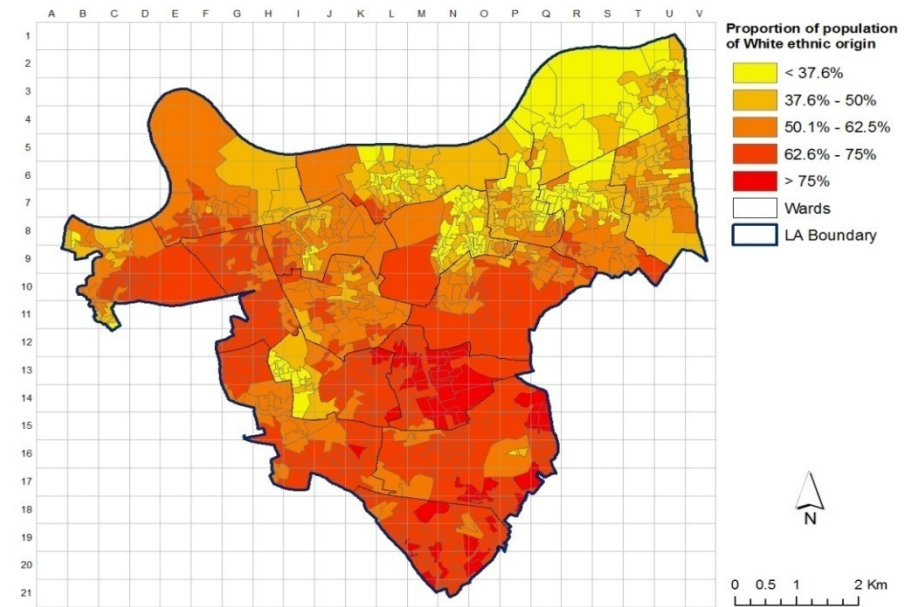
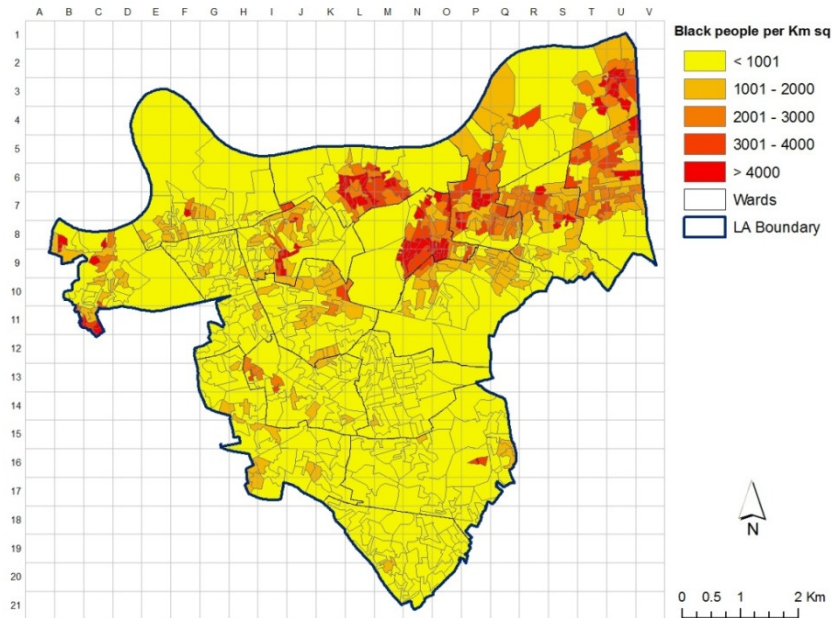
Simplified example of ethnicity assignment

Person ID	Household ID	Black	White	Asian	Other	sum
1	a	1	0	0	0	1
2	b	0	1	0	0	1
3	b	0	1	0	0	1
4	a	0.25	0	0.7	0.05	1
5	a	0.2	0.3	0.1	0.4	1
6	a	0.9	0.1	0	0	1
7	c	0.2	0	0	0.8	1
8	c	0	0.1	0.8	0.1	1
9	c	0.4	0.3	0.2	0.1	1
10	d	0	0	0	1	1
total		2.95	2.8	1.8	2.45	10

Each individual is assigned a probability of being in a certain ethnic group. We actually use 100 ethnic categories.



Mapping ethnicity



*Based on London Borough of Greenwich,
2011*





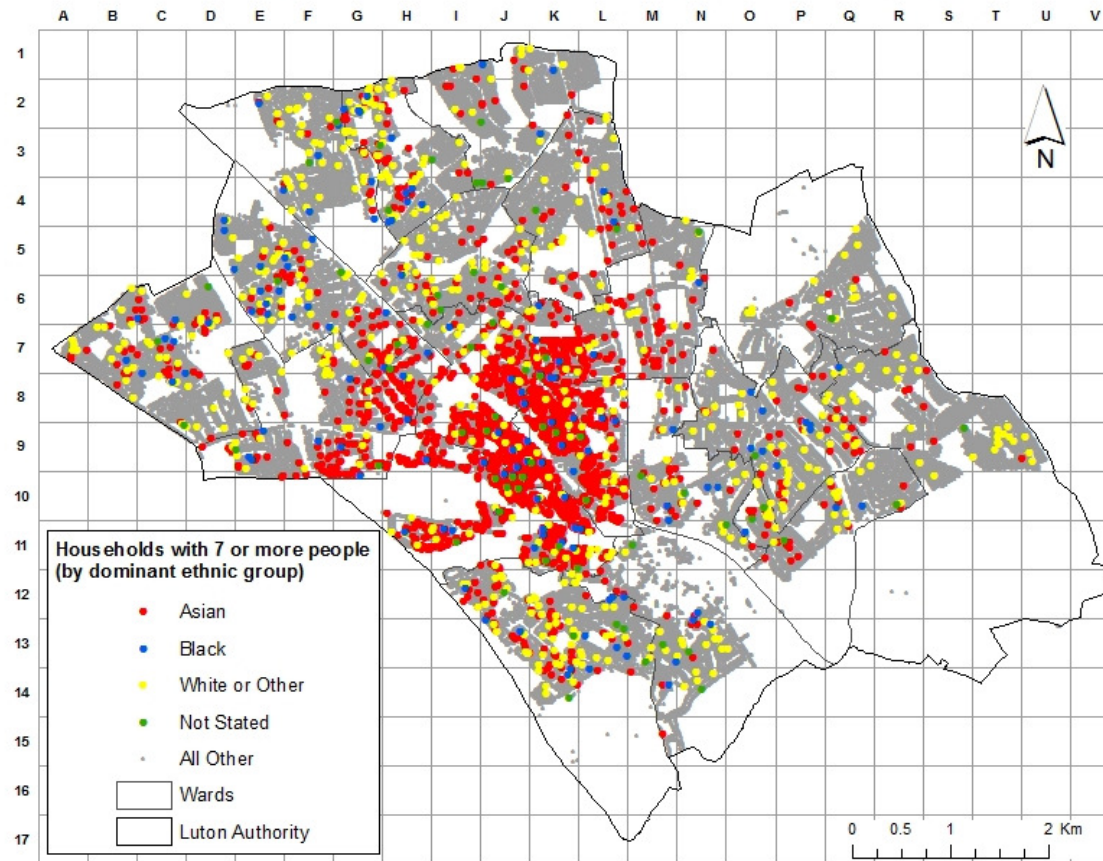
Assigning ethnicity to households

household ID	Black	White	Asian	Other	Household category
a	2.35	0.4	0.8	0.45	Black
b	0	2	0	0	White
c	0.6	0.4	1	1	Asian
d	0	0	0	1	Other
total	2.95	2.8	1.8	2.45	

Number of
people by
ethnic
category

Rule: Ethnicity of a household is determined
by the highest cumulated probability based on
all occupants

Assigning ethnicity to households

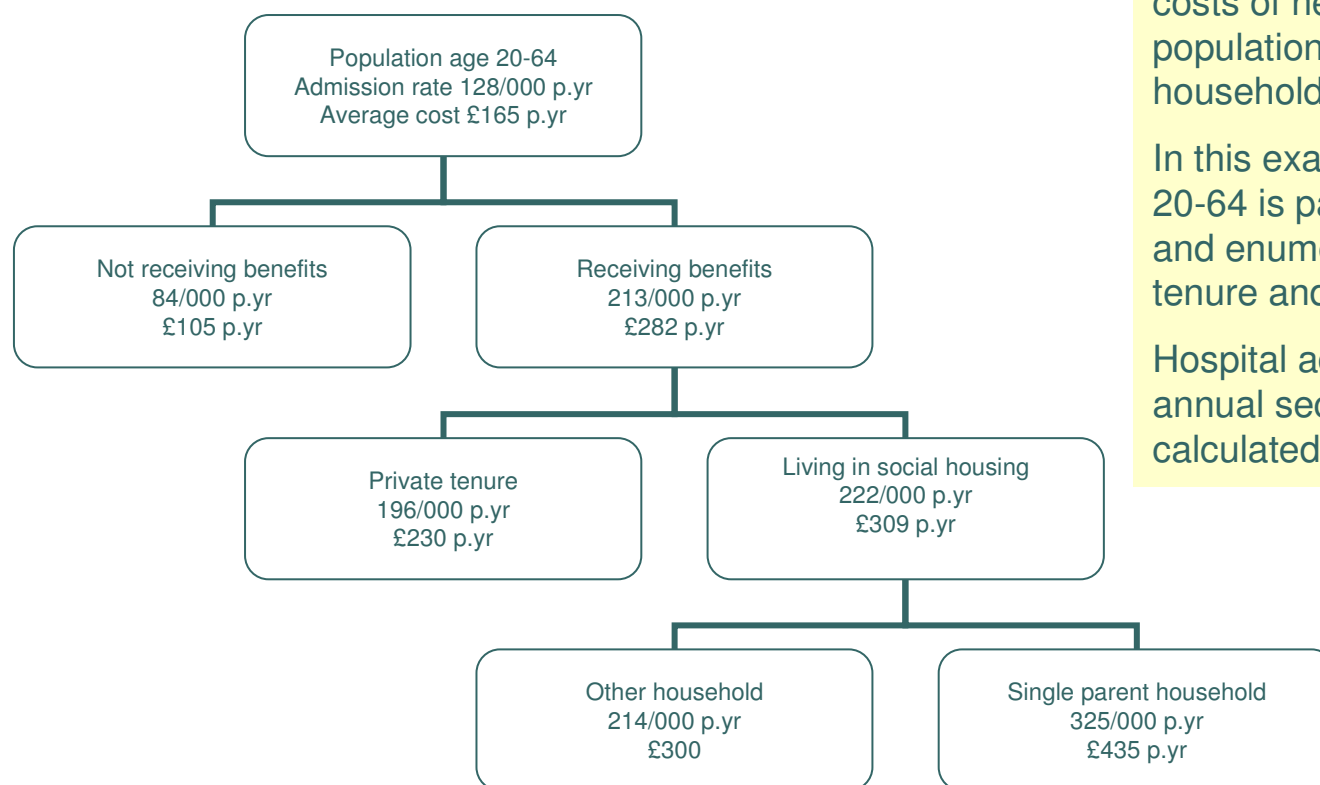


This analysis was used during the census to identify households with 7+ occupants. Colour coding is used to identify ethnic status.

Based on Luton BC 2010



Application 1: Consumption of health care based on hospital admissions



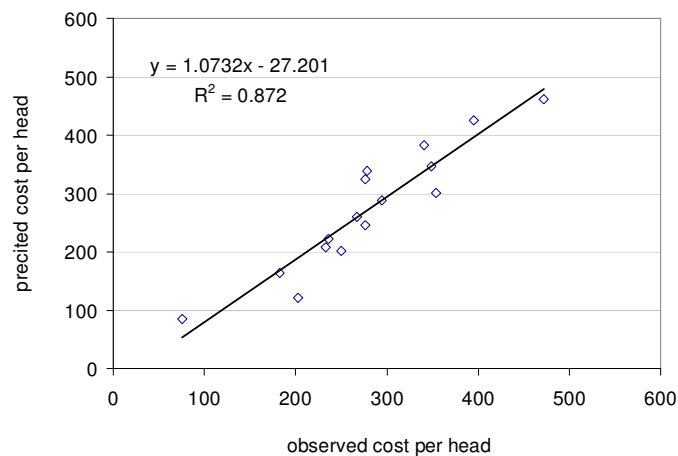
Aim is to understand and quantify the costs of health care among the population according to different household characteristics.

In this example, the population aged 20-64 is partitioned into sub-groups and enumerated by benefit status, tenure, and single parenthood.

Hospital admission rates and annual secondary care costs are calculated for each sub-group.

Contribution of risk factors to hospital admission rates and secondary care costs

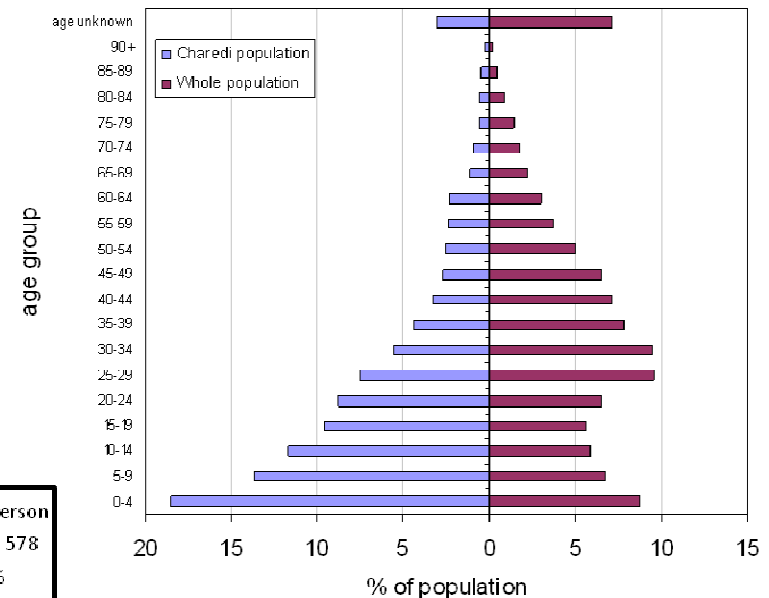
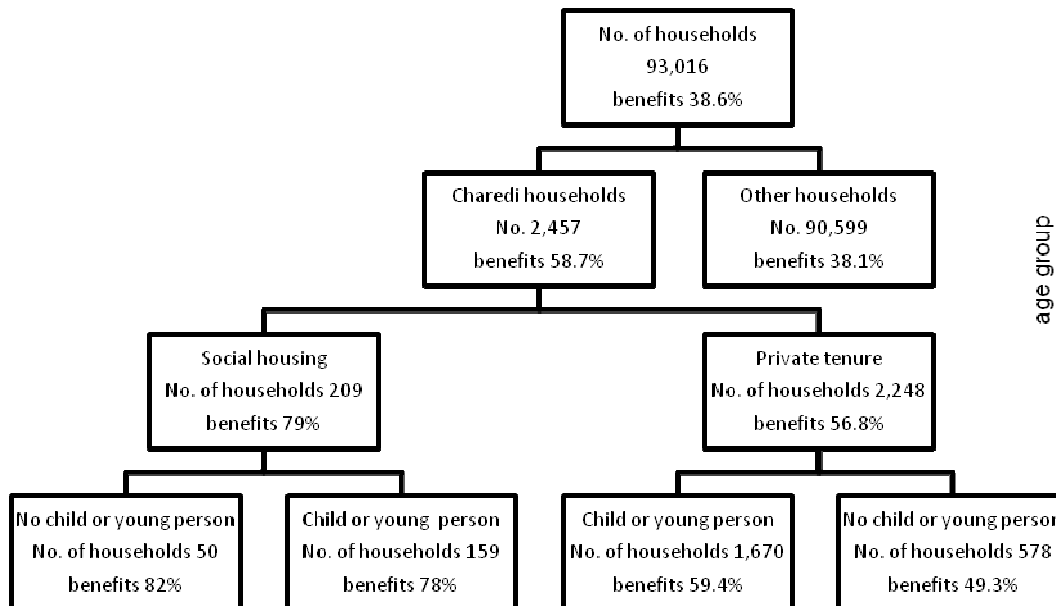
Category	risk factor	incremental admission rate 000s p yr	incremental annual average cost £s p. head
O	none	69	85
A	single parent household	117	137
B	social housing	60	79
C	receiving benefits	86	123
D	3+ children in household	34	37
	total (p.000), £s p. head	366	462



Dummy variable
multiple regression
is used to estimate
factor costs



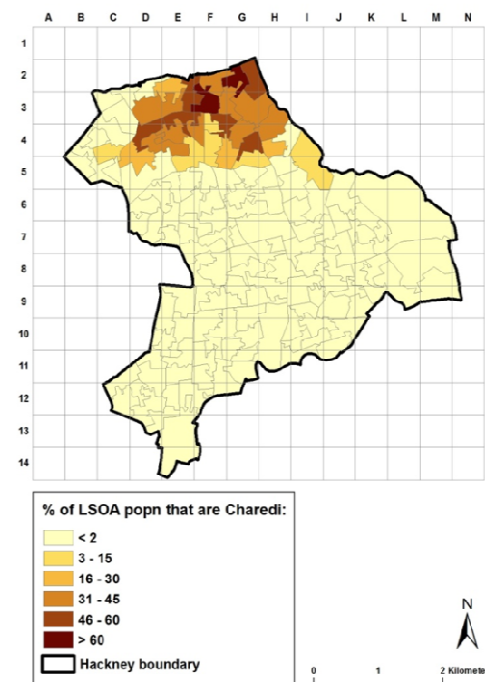
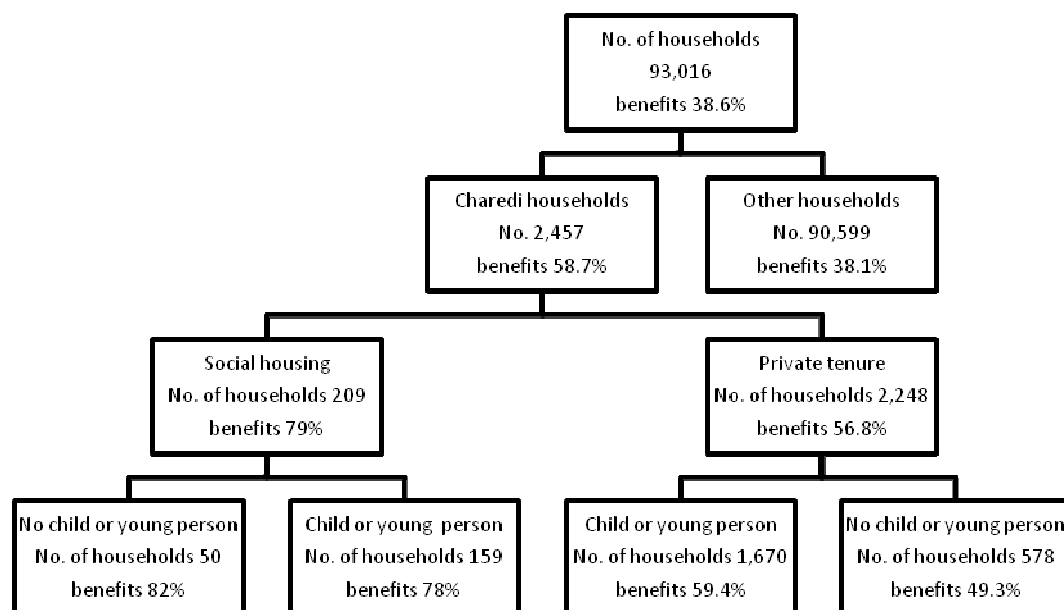
Application 2: Assigning religious attributes (example of Charedi households in Hackney)



2-stage process based on linking Shomar Shabbas names to other data sets and then using probabilistic assignment based on name recognition



Assigning religious attributes (example of Charedi households)



2-stage process based on linking Shomer Shabbos to other data sets and then using probabilistic assignment based on name recognition

Application 3: Troubled families

The 'Troubled Families' programme aims to turn around the lives of 120,000 families by providing targeted interventions to vulnerable adults and children. There is a payment-by-results element of this programme requiring substantial amount of information to correctly target and assess progress of families enrolled in the programme.

category	(1) Youth offender/A SBO/ABC	(2) Persistent absence and/or exclusion, and/or PRU	(3) workless- ness	(4) DADV	total households with given criteria
1	Y	Y	Y	Y	72
2		Y	Y	Y	179
3	Y		Y	Y	134
4	Y	Y		Y	49
5	Y	Y	Y		6
6			Y	Y	139
7		Y		Y	82
8		Y	Y		288
9	Y			Y	87
10	Y		Y		9
11	Y	Y			1
12				Y	124
13			Y		5,187
14		Y			255
15	Y				1
16					12,685
Total	342	900	5995	477	19,298

BASE TABLE ENUMERATING FAMILIES AT RISK

Families must meet any two of criteria. (1) to (3)

DADV is a local filter (drugs, alcohol and domestic violence)

Total households with children

Factors predictive of troubled families and their odds

Odds					
3 plus children in household	a	1.4	1.9	2.7	2.2
Lone parent household	b	1.2	1.4	1.7	1.2
Social housing	c	2.6	2.2	9.8	2.5
SEN(PS)	d	1.8	3.9	1.8	1.9
Children in need flag	e	1.7	2.5	1.5	3.0
Household known to adult social care	f	6.3	6.1	2.8	6.4
Drug/Alcohol/ Violence/DV households	g	38.5	4.6	2.1	n.a





unit cost of each intervention	2500
Key	change value

E
Wider
benefits –
households,
Council,
Government

expected number of households meeting given combination of outcomes based on DCLG definition of success

total expected
successes

wider
benefits
(including
reduction in
DADV)

K	L	M
None of the criteria	Any one of criteria (1),(2) or (3)	Any two or more criteria (1),(2) or (3)
21	32	19
75	82	21
66	56	12
21	23	6
2	3	2
0	0	0
0	0	0
121	132	35
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
305	329	94

[illegible][illegible]

C
Households turned round
meeting DCLG criteria

Household meeting DCLG criteria sum col(S)	94	(Q) reimbursement per successful household	4,000
DCLG success rate sum col(S)/sum col(H) x 100	12.9	reimbursement to LBI sum Col(S) x DCLG reimbursement rate	376.416

Success rate based on all interventions supplied	$\frac{\text{sumcol}(U)}{\text{sumcol}(V)} \times 100$
	30.81

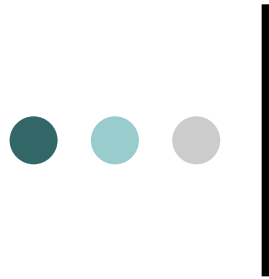
D
Success rate
based on total
number of
interventions
supplied



What we do

- Work with partners to identify issues and identify data sources
- Arrange data sharing protocols that will enable data acquisition and matching
- Clean and link the data before it is anonymized
- Undertake aggregate statistical analyses using risk or related methodologies and geographical information systems (GIS)
- Hand over database and train analysts
- Marginal cost of projects based already completed population base is low





END

lesmayhew@googlemail.com

References:

Harper, G. and L. Mayhew (2011) Using administrative data to count local populations. Journal of Applied Spatial Analysis and Policy DOI: 10.1007/s12061-011-9063-y

Harper, G. and L. Mayhew (2011) Applications of population counts based on administrative data at local level. . Journal of Applied Spatial Analysis and Policy DOI 10.1007/s12061-011-9062-z

Harper, G. and L. Mayhew (2012) Re-thinking households - Using administrative data to count and classify households with some applications. Research paper, Cass Business School, Faculty of Actuarial Science and Insurance

Web: www.nkm.org

