

The role of locational data in 2011 Census and national statistics

Andy Tait,
Office for National Statistics

What I'll cover

- Geography and National Statistics
- The challenges of UK geography
- Geography policy for statistics
- 2011 Census – addresses
- 2011 Census – output geography
- The Open Data opportunity

Statistics are independent



UK Statistics Authority (UKSA) UK Statistics Authority

- independent body at arm's length from government
- established 1 April 2008
- reports directly to Parliament, not a Government Minister

Chair: Sir Michael Scholar

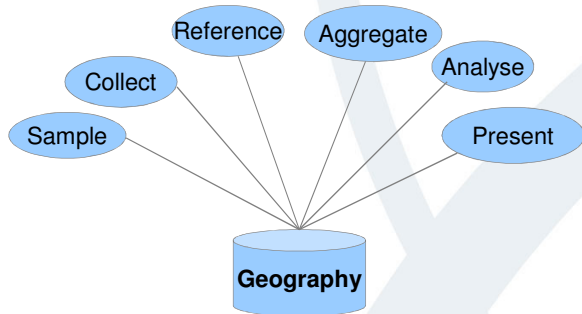


- **ONS** is the Executive arm of the UK Statistics Authority
 - headed by the National Statistician: Jil Matheson
- A National Statistic has to comply with Code of Practice
 - "integrity, honesty, objectivity, impartiality"
 - accessible
 - "proportionate burden"
- Minimum "pre-release" time to Ministers
- Continuous audit/assessment

Geography in UK statistics

“Everything
happens
somewhere”

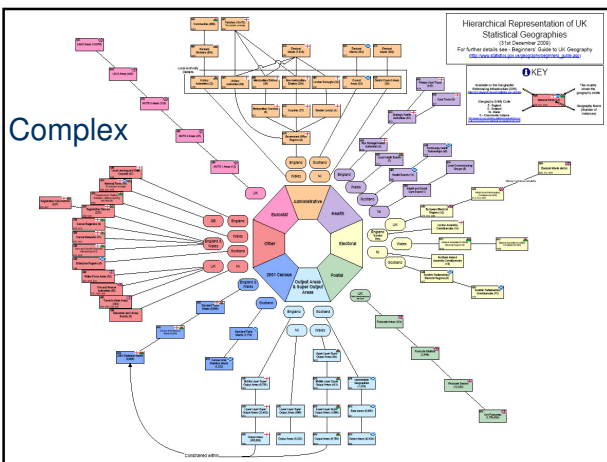
Geography in UK statistics



Geography - why is it a problem?

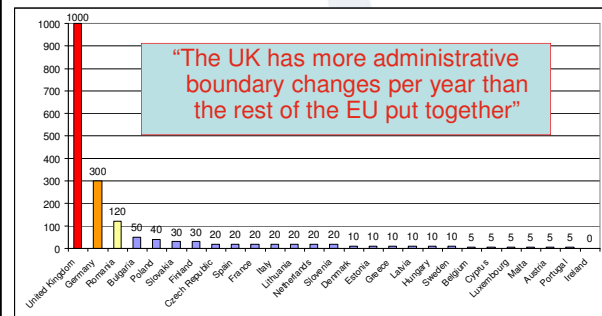
- It's complex
- There are a lot of them
- They change all the time
- They don't align with each other
- Different codes/names used to describe them

...Get it wrong and really mess up your statistics...



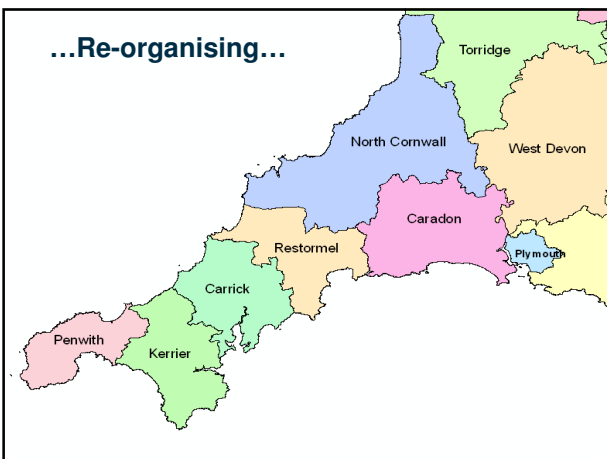
Complex

...Changing....



"The UK has more administrative boundary changes per year than the rest of the EU put together"

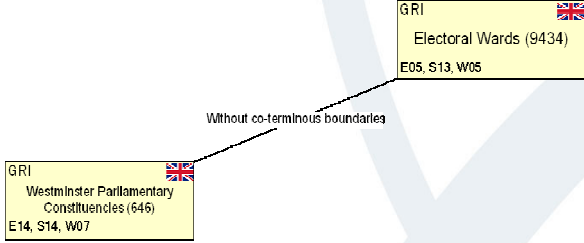
...Re-organising...



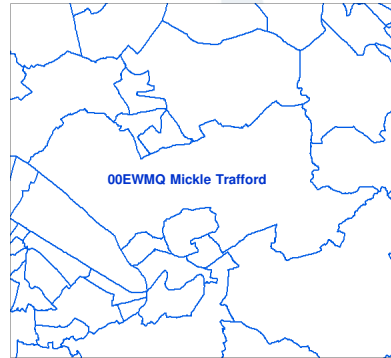
...Changing....

- Boundary Committee for England
- Local Government Boundary Commission for Wales
- Local Government Boundary Commission for Scotland
- Local Government Boundary Commission for Northern Ireland

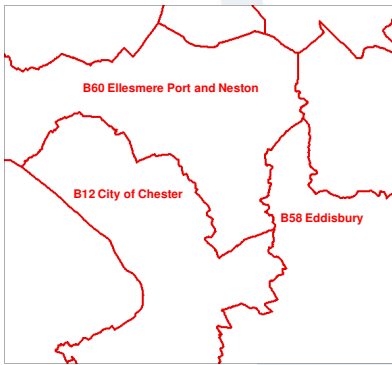
Different Geographies



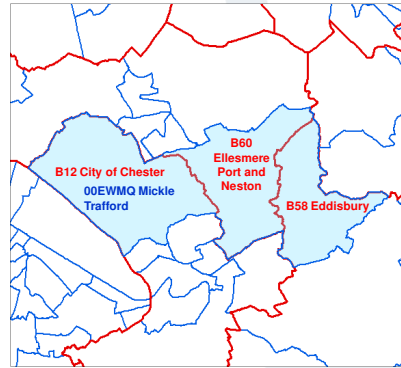
Different Geographies



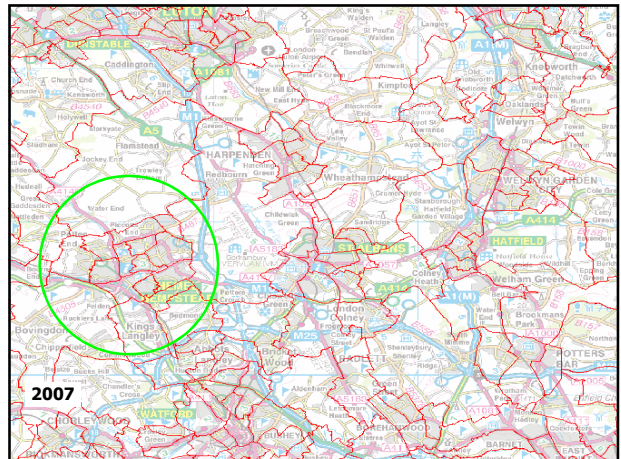
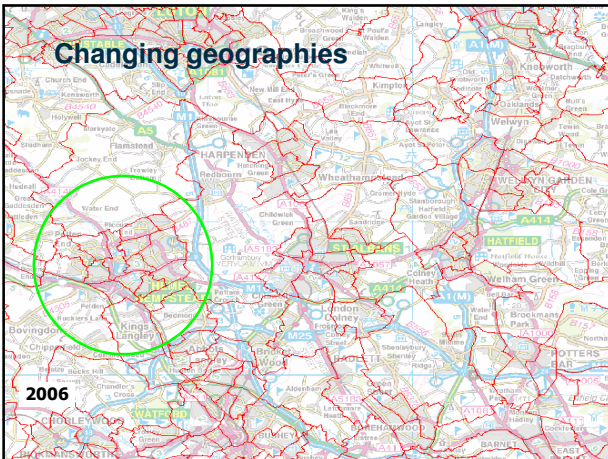
Different geographies

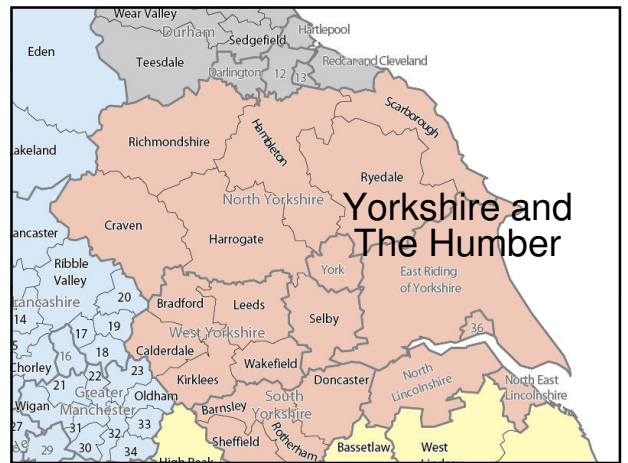
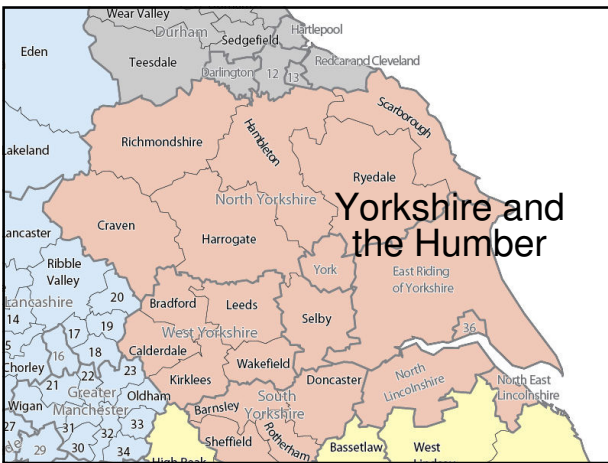
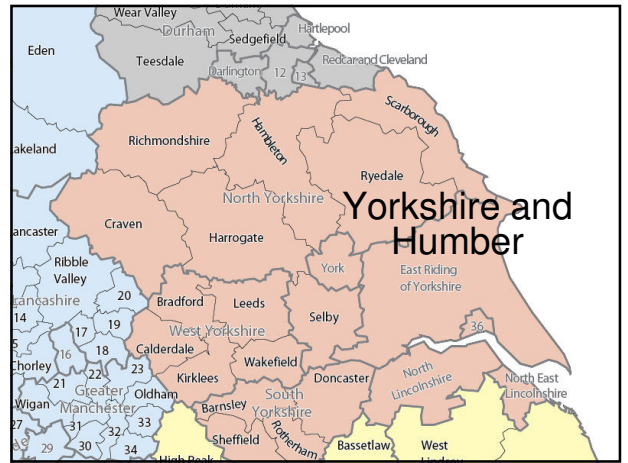
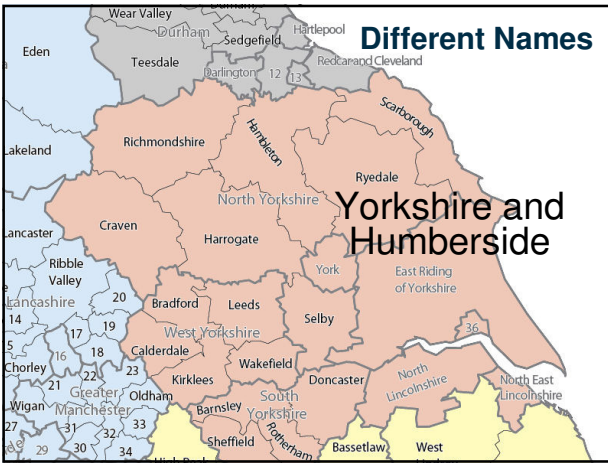
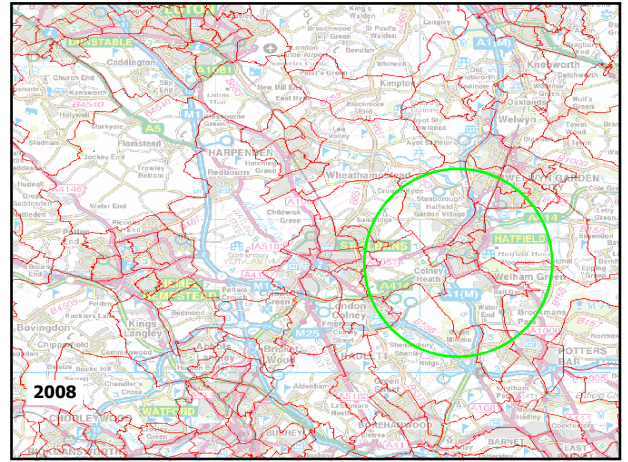
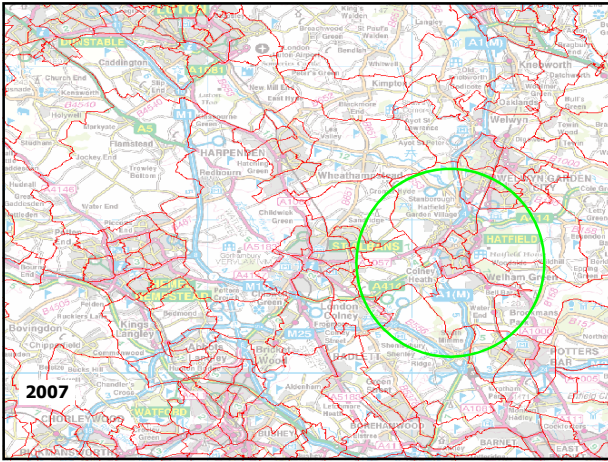


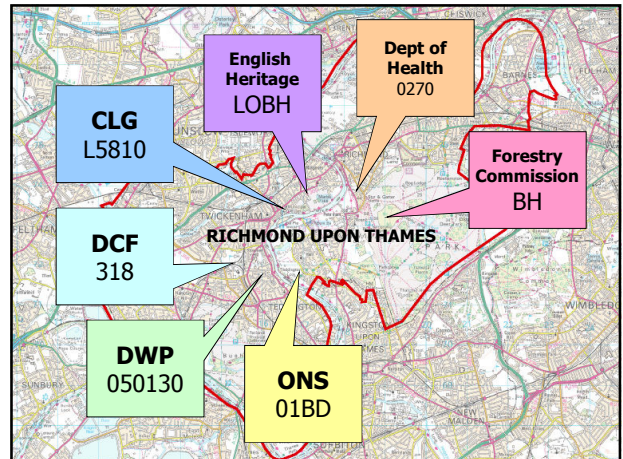
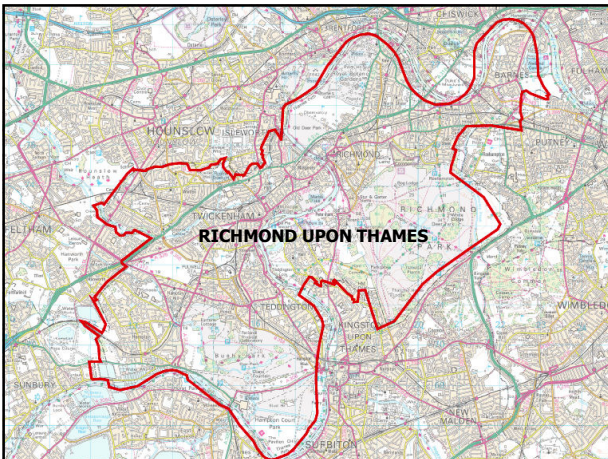
Different geographies



Changing geographies







Supporting statistics with geography

National Statistics Geography Policy

What?

Standards and guidance for using geography to produce National Statistics
 "standard geographic referencing standards"
 (Code of Practice for Official Statistics, Principle 4, Practice 7)

Why?

for the geographic element, the statistics are **accurate, consistent and comparable**

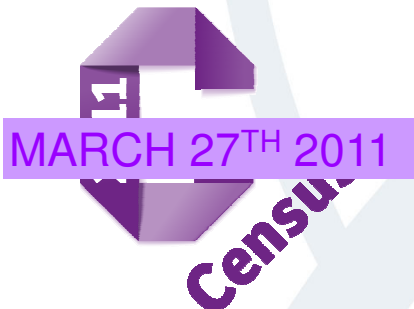
How?

ONS Geography delivers products and services for NS producers to implement that policy

NS Geography Policy - summary

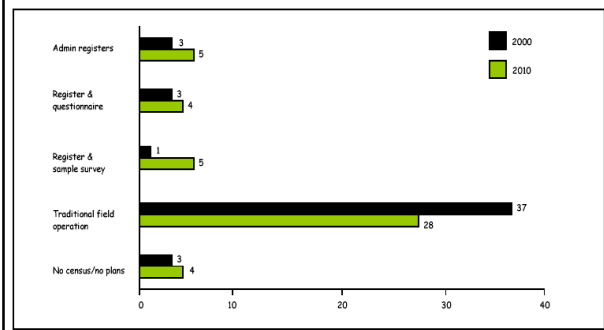
- Referencing your data
- Stable small area building blocks to build National Statistics
- Standard Coding and Naming of statistical geographies
- Defining the currency/versions of geography
- Standard Area Measurement
- Presentation Order

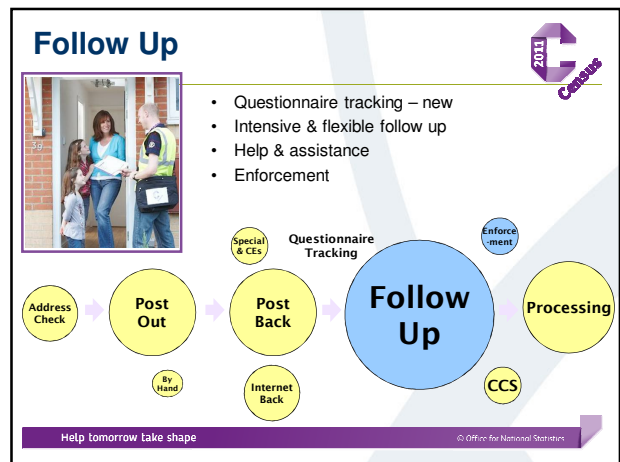
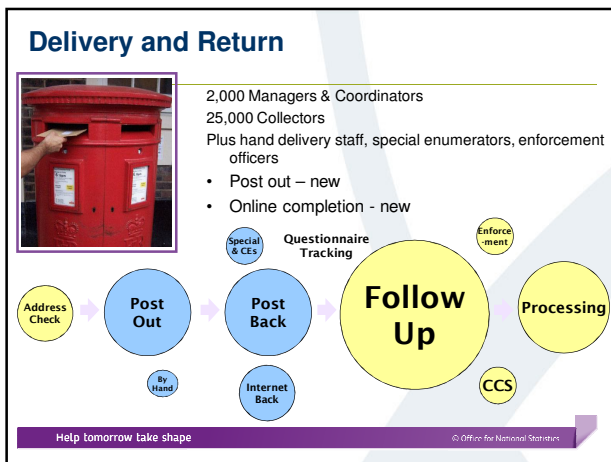
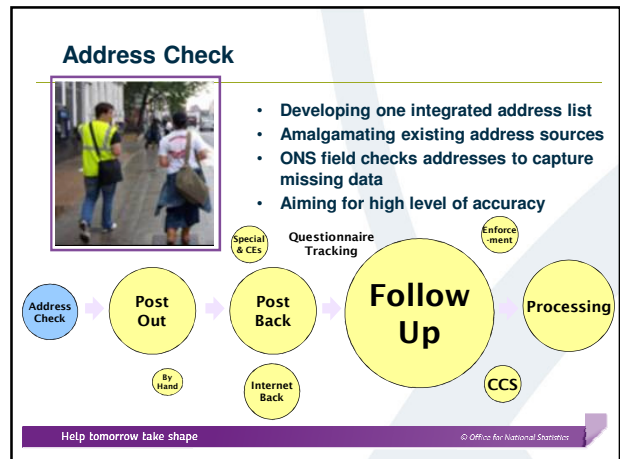
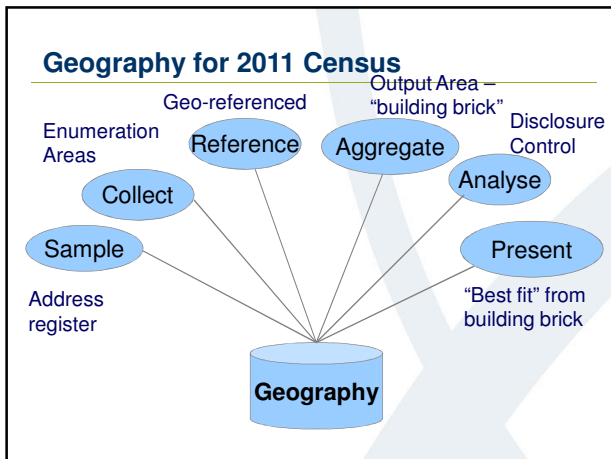
Census 2011



Census: traditional or register?

- Enumeration methods: households, ECE Member States





2011 Census – Why are addresses so important to Census?

- They drive Census coverage: addresses = household(s)
- They are used for:
 - Sampling - for Test, Rehearsal and Address Checks
 - Enumeration
 - addresses on pre-printed questionnaires
 - address lists for hand delivery
 - Data Capture
 - questionnaire tracking/management information/follow-up
 - Creating Census Output Geography – Output Areas
 - Allocating outputs for other geographies

Addressing - “challenges”

- Constantly changing
- Descriptions vary between products – structure, content, format
- Method of data capture
- Multiple occupancy
- Address classifications (residential/commercial)
- And even if you find all addresses....
 -it's households the Census is interested in....and..
 -what about those “beds in sheds”?.....
- No or wrong address = no response
- Wrong address location = wrong allocation

How many addresses at this house?





2011 Address Sources

A number of sources used to provide the most complete register of addresses:

1. Postcode Address File (PAF) – Royal Mail
2. MasterMap Address Layer 2 – Ordnance Survey, based on PAF, but with each address located
3. The National Land and Property Gazetteer – from local authorities
4. Council Tax (residential) lists – Valuation Office Agency (VOA)
5. Non-domestic (business) rates - VOA

Recap - why UK Geography a problem

- There are a lot of them
- They change all the time
 - Postcodes/addresses
 - Boundary change
 - New geographies
- They do not align with each other
- Different codes/names used to describe them

Why do we need a statistical output geography for the UK ?

What	Why
Stable – no constant boundary change	to compare the same areas over time
Similar population size	consistency, comparability
Similar social groupings	comparability, support deprivation measurement reflect neighbourhood
Small enough	No individual/hhold revealed
Non-disclosive	More and more consistent data can be collected
Widely used and recognised	Build consistent statistics for higher level geographies
Use as a statistical building block	

Output Areas (OAs)

- new output geography derived from 2001 Census data (England and Wales)
- For first time, output separated from collection
- built from Census data
- Aligned to wards and parishes as at 2003
- socially homogeneous
- Non-disclosive

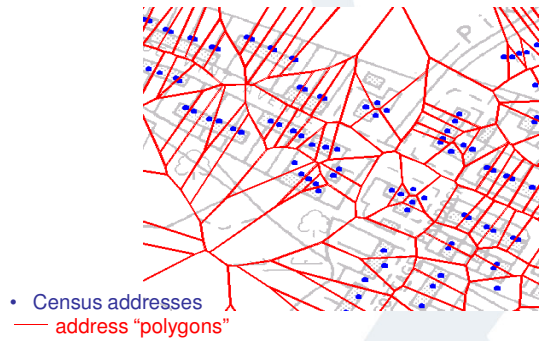
Output Areas – key features

- approximately 175,000 OAs (England and Wales)
- target of 125 households in each OA
 - minimum 40 households or 100 people
- size allows for a fine resolution of data analysis
- boundaries made freely available for non-commercial purposes

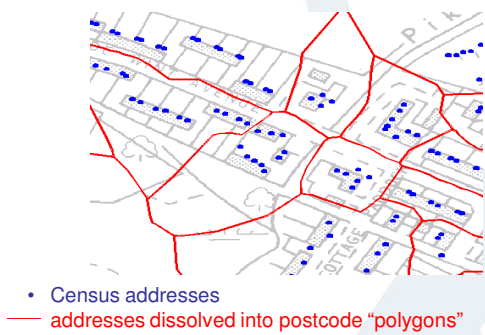
Output Areas – how they were created



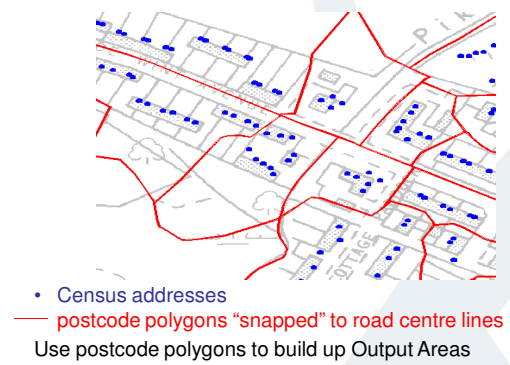
Output Areas – how they were created



Output Areas – how they were created



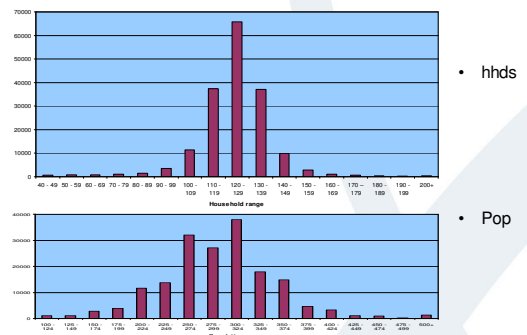
Output Areas – how they were created



Output Areas



Output Areas – achieved size



What are Super Output Areas (SOAs)?

- new geographies for collecting, aggregating and reporting small area statistics
- a layered hierarchy
- groups of Output Areas
- size thresholds and targets

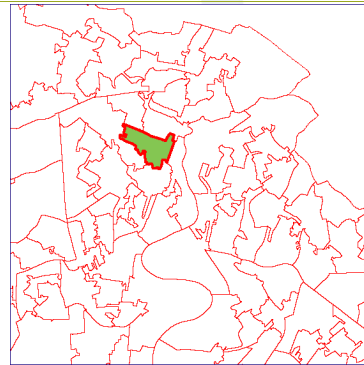
Output Area hierarchy

	Lower threshold	Upper threshold	Population threshold
OAs	100 people	625 people (2 *target)	2.5 * household thresholds
LSOAs	1000 people	3000 people (2 *target)	2.5 * household thresholds
MSOAs	5000 people	15000 people (2 *target)	2.5 * household thresholds

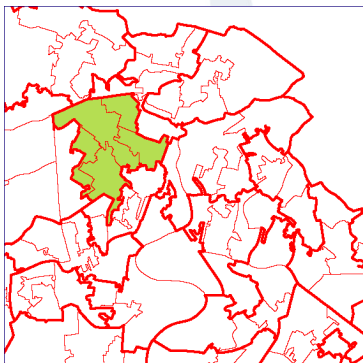
A group of Output Areas



A Lower Layer Super Output Area



A Middle Layer Super Output Area

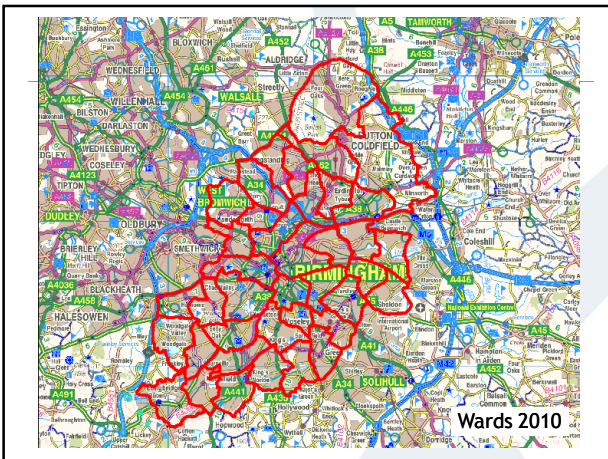
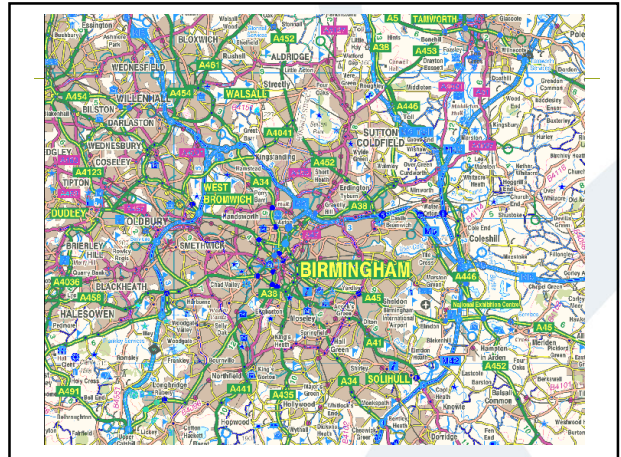


OA/SOAs hierarchy

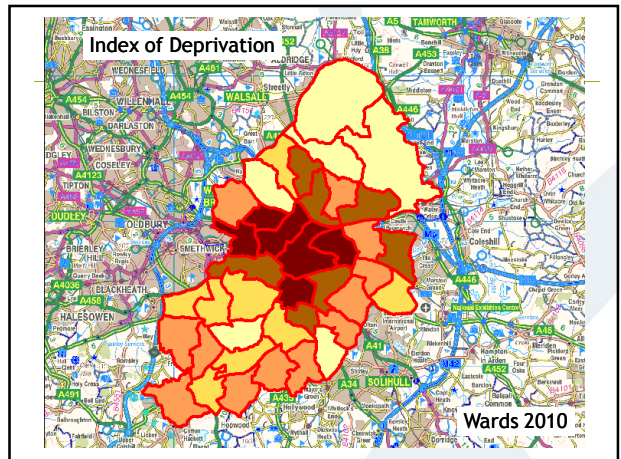
	Lower threshold	Upper threshold	Population threshold
OAs	100 people	625 people (2 *target)	2.5 * household thresholds
LSOAs	1000 people	3000 people (2 *target)	2.5 * household thresholds
MSOAs	5000 people	15000 people (2 *target)	2.5 * household thresholds

Key benefits of Super Output Areas

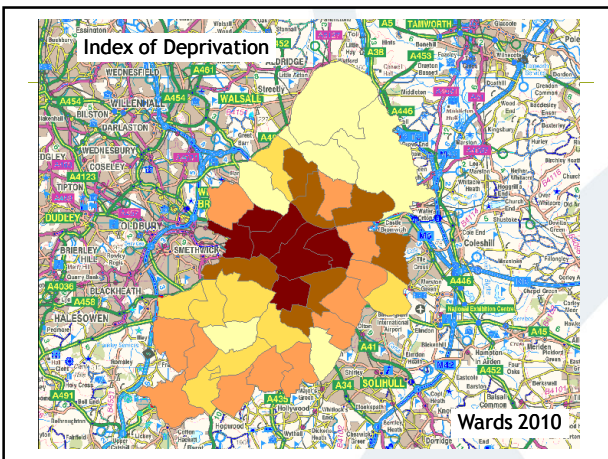
- small
- stable
- independent
- nationally consistent size and method
- a choice of level at which to publish data



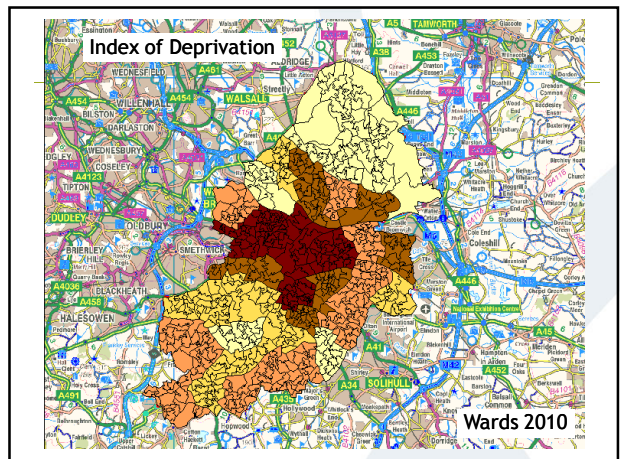
Wards 2010



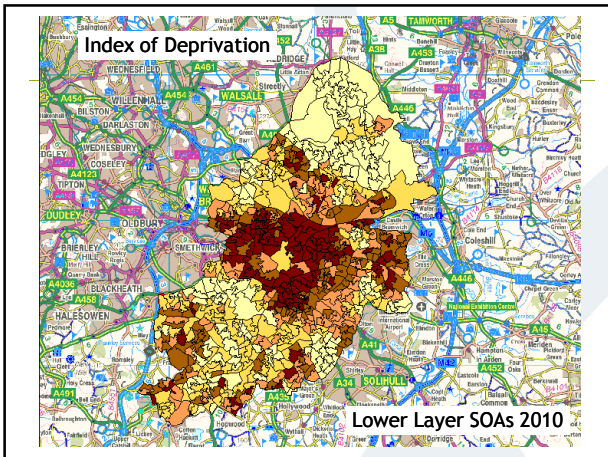
Wards 2010



Wards 2010

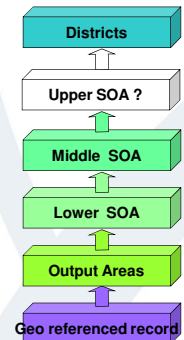


Wards 2010



Geography policy for statistics

- Emphasis on stability
- improving our ability to measure change over time
- Use of standard geographic units
- Statistics built from these blocks - or by estimation for other areas



Changes since 2001 - population

- Population growth, especially migration
- More and smaller households
- Newly built properties
 - Greenfield/new land
 - Brownfield/in-filling
- Sub-division of existing properties
- Changing socio-economic characteristics of areas

Changes since 2001 - geography

- Postcodes
- Census address register
- Ward/parish changes since 2003
- Administrative re-organisation

Requirement – Population Growth

- OAs and SOAs designed as a stable geography for the use of time-series statistics

However

- Some of population and household thresholds used in 2001 have been breached
- Will need to be redesigned to avoid disclosure (under threshold) and to allow comparability (over threshold).

Estimated population change

Output Areas

	2005 below	2005 within	2005 above	2001 totals
2001 below	221	228	1	450
2001 within	147	173553	682	174382
2001 above	0	78	506	584
2005 totals	368	173859	1189	175416

99.1%

2011 Census

- Total change to OA/SOAs limited to 5% to ensure stability
- Change only where:
 - 2001 thresholds have been breached
 - OA/SOAs weren't socially homogeneous in 2001
- Use of the National Address Register will improve accuracy of Output Areas

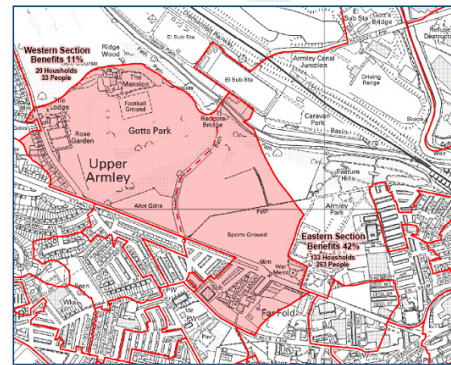
Workplace Zones

- OAs based on where people live not work – can be unsuitable for workplace statistics
- Some OAs contain no/few businesses; some contain many businesses or large employer, e.g. business parks, City of London
- Look at splitting/merging OAs for a new geography nesting with OAs
- User Group established
- Pilot zones to be created by end 2010

2010 Output Geography consultation

- Need for an Upper layer SOA ?
- Workplace Zone requirements
- Provide instances of OAs/SOAs that are unfit as a statistical geography
 - Priority instances
 - Not useful for analysis due to their design
 - ONS panel to consider redesign

Social Homogeneity?



“Freeing” UK geographic data

- “Making Public Data Public” – Data Transparency
- data.gov.uk - “linked data”
- OS OpenData launched 1st April 2010
 - Core UK data free including commercial re-use
 - NS Postcode Directory now Open
 - 2001 Output Areas now Open
- A Public Sector Mapping Agreement from 1/4/2011
- Simplified OS pricing and licensing models

Finally.....

- Definitive
- accessible
- free
 - > standard
 - > joined-up
 - > consistent

