

Generating Value from the 2011 Census –
How business can benefit from the new free Census data

MRS Census & Geodemographics Group seminar
19 September 2012

Maximising use & value –
Commercial users' needs, and
the Census Offices' plans

Keith Dugmore
Demographic Decisions

Agenda

- The Census – a unique source of information
- Why has the 2001 Census been so valuable to commercial users?
- Do the Census Offices' plans for 2011 meet commercial users' needs?
- Commercial users' priorities for delivery
- Further innovation – some ideas
- Getting ready for the deluge of data

PHEW.....!!!!!!

30-SECOND CHALLENGE

www.30secondchallenge.com

START on the left with the number given and follow the instructions as you go across. Within the 30-second time limit, beginners have to complete their own challenge, intermediates have to

complete their own challenge AND the beginners' challenge, and advanced players have to complete their own challenge AND the intermediate. The answers are on Page 46.

YOUR ANSWERS

Beginner	1/2	x10	TIMES ITSELF	x4	+8	÷9	x6	1/2 OF IT	1/4 OF IT	x20	
Intermediate	29	x20	-1	1/3 OF IT	-1	÷16	TIMES ITSELF	25% OF IT	+19	4/11 OF IT	
Advanced	26	250% OF IT	9/13 OF IT	7/90 OF IT	x1,000	LESS 20%	17 1/2% OF IT	3/7 OF IT	+10%	33 1/3% OF IT	

To order Daily Mail 30 Second Challenge books for £5.99 each, call 0845 155 0720 or visit www.maillife.co.uk/books

Daily Mail
2 March 2011
Page 2

Could 2011 see the last-ever census?

Supermarket loyalty cards may be used instead

THE national census could be abolished – and information gathered through supermarket store cards instead.

A ten-yearly head-count due on March 27 will demand details on everything from ethnic identity to your central heating.

But officials are looking at ways to replace the expensive and controversial count before 2021.

They believe tapping into data already held on store cards such as Tesco Clubcards and by phone and energy companies, banks and even the National Lottery could help provide an alternative.

The idea has been raised in meetings with representatives of private sector firms by the Office for National Statistics, which is responsible for carrying out each count.

It said plans involve looking at 'alternative' sources of information 'including commercial companies'. This year's census

By **Steve Doughty**
Social Affairs Correspondent

will employ 35,000 and cost the taxpayer £500million.

Officials are keen to avoid a repeat of problems encountered in the 2001 census, which missed a million people and led to years of confusion over population figures.

However, the ONS fears the 2011 count could be jeopardised if the public knows of its doubts over the future of the census.

A paper presented by its statisticians last year warned that if the public find out too soon it will have 'an adverse effect on the success of the 2011 census' and the exercise must be kept quiet to 'minimise this risk'.

There is already disquiet over the intrusive nature of the latest count.

It requires a new state database of every address in the country and a 'tracking' operation to ensure that every home fills in a form. The forms have 918 tick-box options over 32 pages.

But all the data sought – and more – are available on private-sector databases.

Store cards contain information on addresses, family members, what they spend and where. Some 15million individuals and families have personal and spending details recorded through Tesco Clubcards alone.

Credit cards have similar information, and energy companies and phone companies have wide-ranging details of family homes and habits.

Privacy watchdogs are critical of attempts to replace the census with private-sector data. Daniel Hamilton, of Big Brother Watch, said: 'Not content with asking intrusive questions about your religious views and the type of central heating you have, the Government now wants to snoop on your store cards.'

'Retailers must hold true to the promises they have given their customers not to pass their information to third parties.'

The possibility of tapping

into private-sector databases is understood to have been raised in talks with the Demographics User Group, headed by a statistician with links to the ONS and major firms.

Members include Tesco, Sainsbury's, John Lewis, Marks & Spencer, the Co-op, Boots, lottery operator Camelot, power firm E.ON, Orange, Barclays and Nationwide.

No one from the DUG was available for comment yesterday.

The ONS said: 'At the moment there is no alternative to the traditional census count.'

'A successful census in 2011 is essential to meet the needs across a wide range of users.'

'The decision to change this approach for the future must therefore be subject to careful consideration.'

'Beyond 2011 [which is running the ONS scheme to replace the census] is evaluating the option of using data already held and is working closely with Government departments and external experts including commercial companies.'

Daily Mail
2 March 2011
Page 1



EXCLUSIVE
The fabulous
but VERY tight
Oscar dress
that landed
me in hospital
BY JOAN COLLINS

SEE PAGE 23

The Census – a unique source of information

- The questionnaire – you'll remember a wide range of topics for households and individuals
- Every 10 years – a single day in March or April
- Organised by Government – compulsory, aiming for 100% coverage
- England, Wales, Scotland and N Ireland – & often consistent across the whole of the UK
- Scope to produce enormous quantities of detailed statistics for very small areas

Why is Census data so important to commercial companies?

- Decisions, decisions.....
 - What areas are best for our new stores?
 - What should we offer in each outlet?
 - Where should we advertise?
 - Who are our best customers, and prospects?
 - Which areas & people should we survey?
- Investments of £hundreds of millions to be targeted every year



Making Data Work

YOUR M&S



John Lewis

The **co-operative**

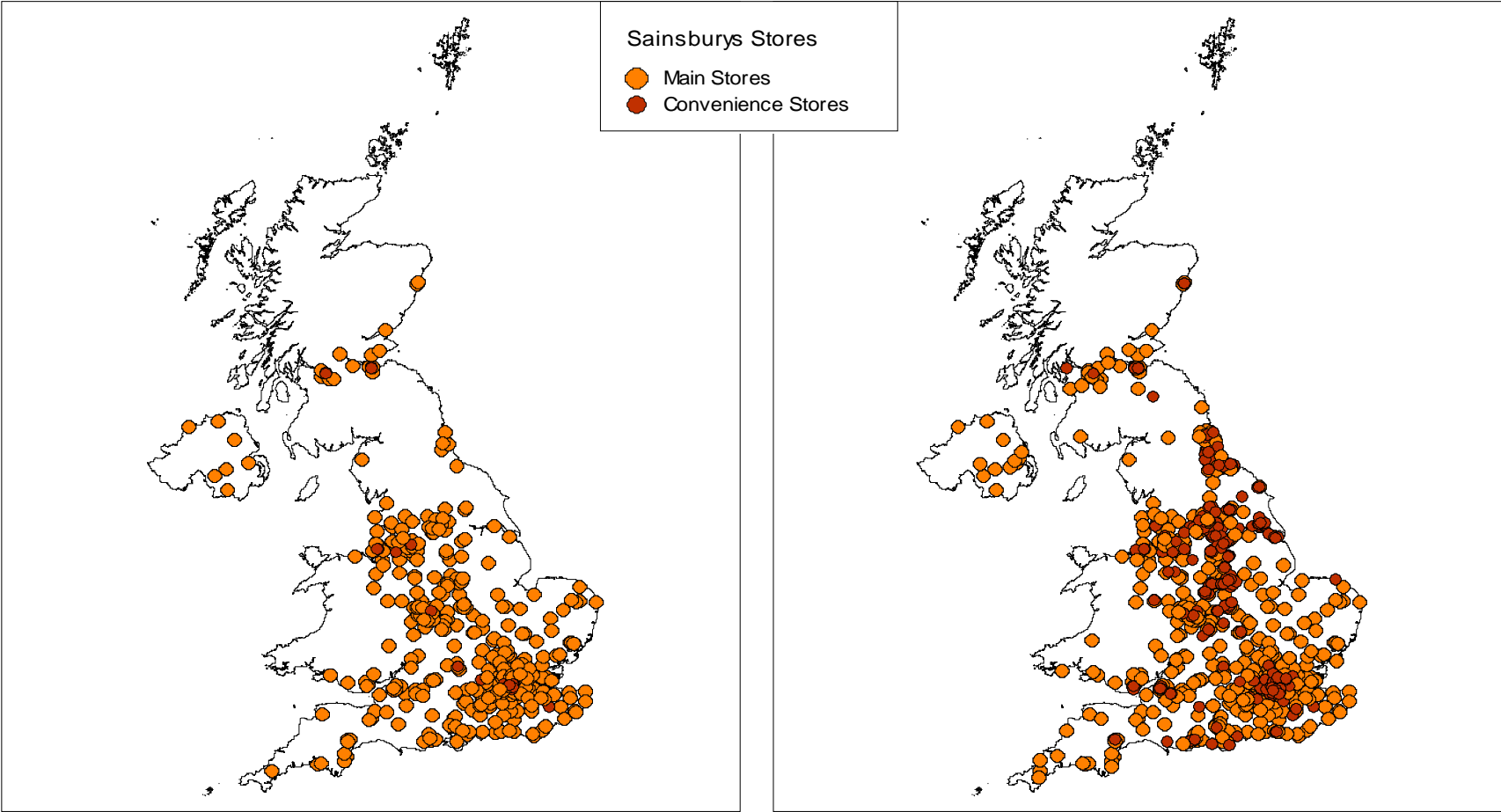
serco



DUG – the Group

- The Members
 - 15 national businesses, from several different sectors
 - Well-established Insight teams
- DUG's Objectives
 - To work with government on behalf of commercial users to ensure the right data, is made available in the right way, at the right time.
 - To encourage DUG members to share their experiences of using various datasets, and also their methods of analysis, and the insights obtained.
 - To act as an incubator for new ideas, which might give DUG members first-mover advantage, and which can be subsequently taken up and developed by others.
- Websites
 - DUG: www.demographicsusergroup.co.uk/
 - DUG Network (LinkedIn)

Sainsbury's estate since release of 2001 Census data



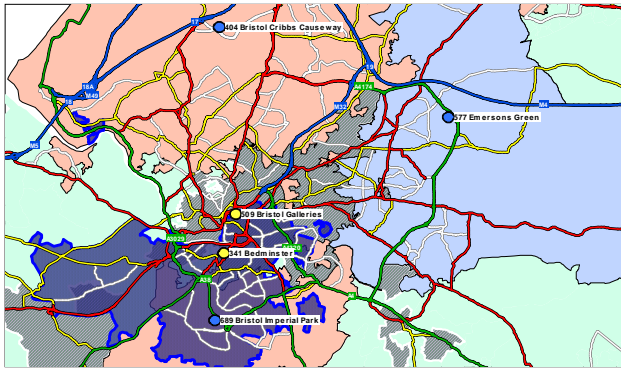
March 2003

March 2010

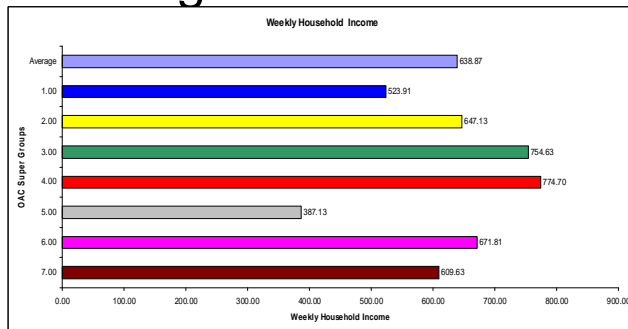
Analysis, and the need for Government data generally

Analyses

1. Local areas



2. Profiling individuals



3. Designing surveys

Data – with national coverage

- Statistics
 - Census-type counts for very small areas
 - Sample surveys
- Map data
 - Background, point locations, road network, boundaries, postcode look-ups
- Lists: big files of individual addresses & sometimes people

Which 2001 Census outputs have been of most value?

- Population counts for Output Areas (c. 120 households)
 - Totals, age, sex, Social Grade, students, car availability, ethnic profiles, single-person households, pensioner households, etc.....
- Workplace counts
- Geodemographic classification, OAC

Together with:

- Digital boundaries – Output Areas
- Postcode / OA directory
- Directory of OAs to higher areas

2001 Census – Pro's and Con's from a DUG member

Pro's:

- Breadth of coverage (UK)
- Depth of coverage (number of variables)
- Small area building blocks
- Basis for geodem classifications
- Workplace stats – daytime populations
- Underpins population estimates & projections
- Free to use!

Con's:

- 10 yearly snapshot
- UK-wide data only freely available by visiting 3 separate suppliers
- Confidentiality rounding makes data complicated to use

& we mustn't forget that many commercial companies use the services of Value Added Resellers, e.g.



2011 Census –
commercial users' needs, &
the Census Offices' actions

2011 Census – meeting commercial users' needs? Questions & Coverage

DUG members' views

Questions on a wide range of topics, continuing previous Censuses

New Questions:

- Second residence (& alternative population definitions); Language
- Income – TOP priority, & asked in many other countries

Target effort to achieve coverage of >90% in all areas, with estimates of the full 100%

Census Offices' actions

Most topics have been retained

- Both included
- Rejected, after much debate

Creation of a definitive address register, targeting of difficult areas, + methods for estimating non-response
(Better response than 2001)

2011 Census – meeting commercial users' needs? Outputs: Geography, & Confidentiality

DUG members' views

Retain 2001 Output Areas:

- Small (120 households)
- Linked to postcodes
- Stable 2001-2011

Confidentiality: “Statistical Disclosure Control” (SDC). Learn from the mistakes of 2001, deciding on a method that produces consistent statistics within and between tables

Census Offices' actions

Formal public consultation, with the welcome decision to retain existing OAs wherever possible

After considerable research, the chosen pre-tabular method is very welcome, giving consistent tables

2011 Census – meeting commercial users’ needs?

Outputs: Statistics

DUG members’ views

Many detailed tables, + univariate counts at OA level are fundamental.

They should also include:

- ONS’s OA geodemographic classification
- Statistics for Workplace populations at fine level

More population bases, e.g. non-term time; 2nd homes

Census Offices’ actions

The Census Offices’ consultations & plans suggest that we won’t lose many significant outputs c.f. 2001, and will gain with the new topics and population bases.

Workplace Zones will give smaller areas in city centres, BUT disclosure control on the basis of businesses (rather than workers) will be more destructive than in 2001.

“Daytime Population” for Output Areas would be a good solution

2011 Census – meeting commercial users' needs?

Outputs: UK-wide

DUG members' views

Consistency across UK's 4 countries

For many variables there is universal interest.

Others may be local (e.g. ethnic or religious classifications in London c.f. Northern Ireland), but there is demand for comparable super-sets

Census Offices' actions

The heads of the 3 Census Offices issued a joint statement of their intention to cooperate – good news

The Census Offices have put much work into comparing tables, & also creating a Census database of records for the whole UK

2011 Census – meeting commercial users' needs?

Outputs: Accompanying products

DUG members' views

As well as statistics, it is essential to have:

- Digital boundaries for Output Areas
- Digital map background
- Postcode / Output Area directory

Census Offices' actions

The Census Offices will have to negotiate agreements with other government agencies, which have no similar statutory commitment to the “public good” – but with the Transparency agenda the outlook is encouraging

2011 Census – meeting commercial users' needs?

Delivery: Licensing & timing

DUG members' views

Licensing, and free at the point of use, repeating the success of 2001, & maximising use

Quicker? Users have differing views on this – “we need it as soon as possible” versus “we’ve waited 10 years, so let’s get it right”

Timetable. Agreed by all users that a timetable should be published & delivered: delays cause big problems of staff resources.

Census Offices' actions

Access will be free at the point of use again

Timescales are a little quicker than 2001 – headline results at Local Authority level were published in July 2012 (after 15 months, rather than 18). **(But Scotland's timetable is 6 months later)**

The Census Offices appreciate the importance of a guaranteed timetable

2011 Census – meeting commercial users' needs?

Delivery: Easy access

DUG members' views

Good delivery by the Census Offices (= good access for Users) is vital:

- Simple flat (c.f. hierarchical) files
- Popular formats (Excel, csv, etc)
- Downloadable from the web in most cases
- An API for integrating Census stats into other applications
- Ideally from one website or portal for all UK data
- Email alerts to let users know of new releases

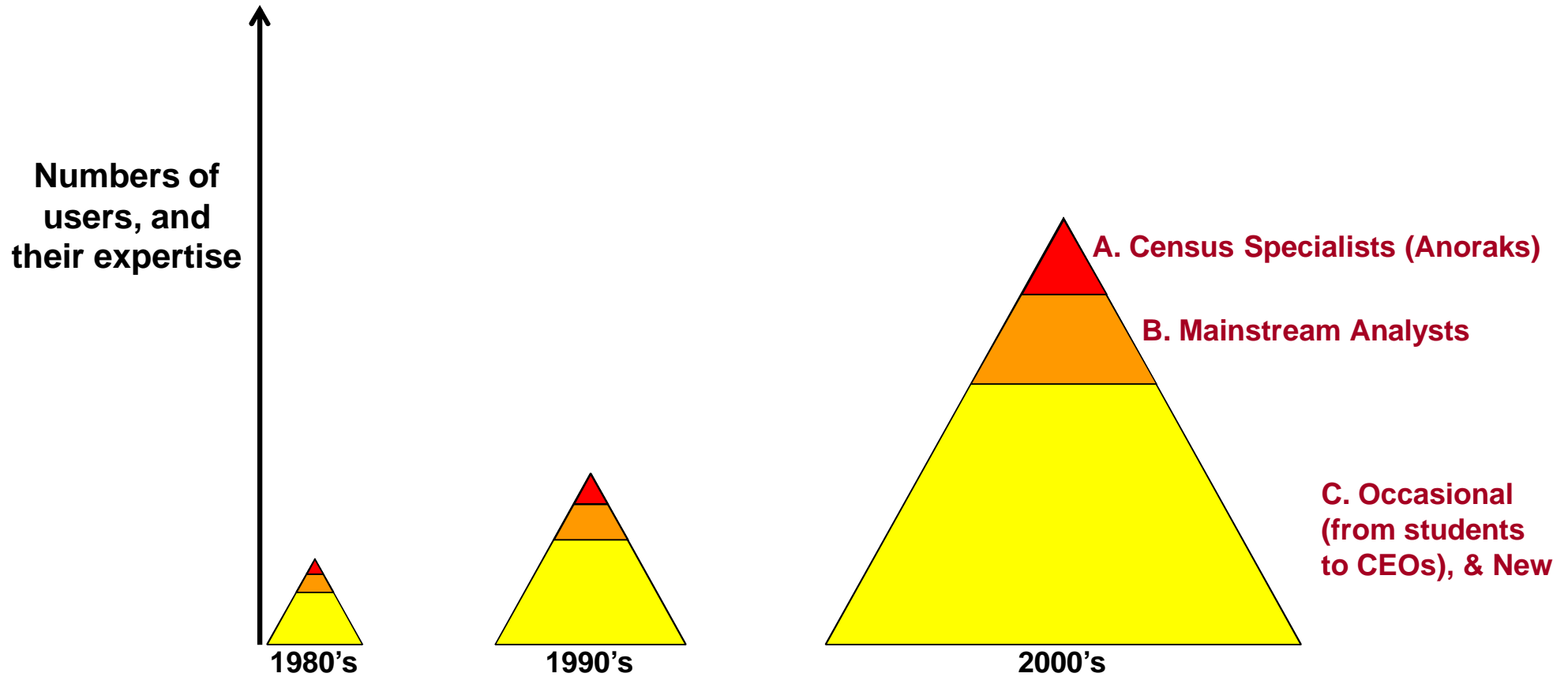
Census Offices' actions

Not all planning is finalised, but now that the availability of csv has been confirmed, there aren't currently any outstanding concerns

[Have been reassured that Scotland's decision to use SuperTable will not repeat the problems of 2001]

[Doubtful about the benefits of flexible tables /hypercubes]

Census – user segments



2011 Census – potential products

- Published tables (*for users A, B, C*)
 - Key Statistics, Census Area Statistics, Standard Tables, & area classifications – together with supporting digital boundaries & directories
- Commissioned Tables (*A, & some B*)
 - Ordered by experienced users who require more detail than is available in the published tables (2001 cost c.£150 a time)
- Flexible table generation (*A, & some B*)
 - Plans for 2011 to offer users the opportunity to create their own tables for the first time, using hypercubes of aggregated statistics
- Specialist datasets, particularly used by academics, often for central government policy projects (*A*)
 - Origin / Destination tables (Special Workplace Statistics, Special Migration Statistics)
 - Microdata files: SARs (various); and the Longitudinal Study

DUG's Priorities for delivery

First.....

- Key stats at OA level – which will also be input for geodem classifications
- UK-wide
- + boundaries & postcode directory

Then.....

- Workplace / Daytime population
- Detailed tables
- Ad hoc commissioned table service
- Samples of Anonymised Records
- Origin/destination tables – commuting, and migration

The need for innovation in access – the last 100 yards

- Only 3% of Census budget was used for Outputs in 2001
- The Census offices must maximise use / grow the pyramid / gain more fans
- Improve access to increase benefits / utility.....

Utility

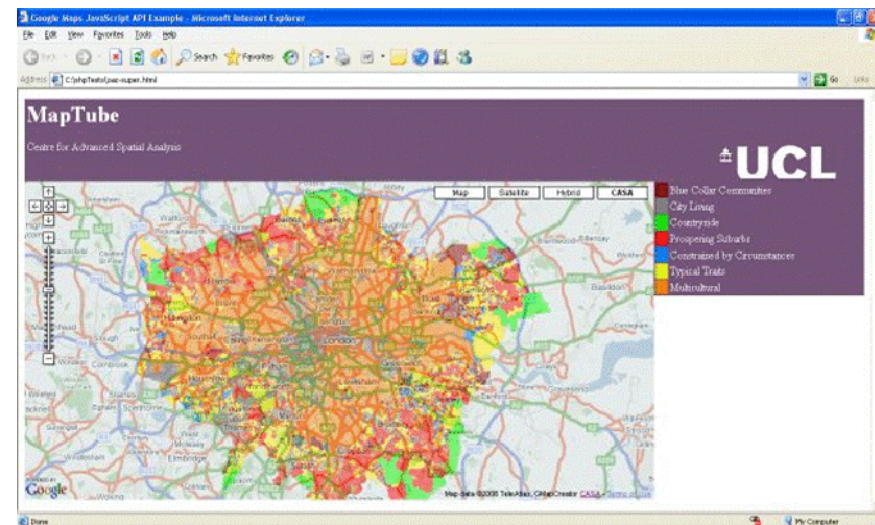
"That property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness.....It is the greatest good to the greatest number of people which is the measure of right and wrong."

Jeremy Bentham.
Introduction to the
Principles of Morals and
Legislation (1789)



2011 innovation – encouraging Apps & Mashups

- Making big datasets available to mass audiences
- Apps:
 - Mobile phone
 - Ipad
 - Netbook
 - Laptop
- Mashups with other datasets:
 - E.g. <http://www.maptube.org/>



2011 Innovations

- Encourage new markets – ubiquitous Search websites (offering simple summaries)
 - Google, BBC, newspaper websites, Up My Street, etc.
- Encourage new markets – business websites (offering immediate access to detailed information)
 - Existing Census Distributors, + Bloomberg, Dun & Bradstreet, etc?
- And a thought triggered by Spotify – selecting just those items of interest.....

Much easier remote commissioning of tables

- Have the Amazon or Trainline websites in mind.....
- Menus for selections
 - Choose Geographical level, Geographical area; Topics and their variables, Output Excel / CSV / map, etc. Go to checkout. Submit.
- “Quick Table”
 - 2,3 (or more)-way
 - The complete table / population
 - But automated checking may flinch at some sparse cells due to confidentiality / SDC concerns

There should be a way to resolve this paradox of tables.....

- “Quick Count”
 - Just counts / cells of interest – not the entire table
 - Local concentrations of particular populations, not the empty areas

We need to squeeze out every drop of useful information.....

2001 Census Area Statistics		Area Identifier -					
Theme Tables		00BGGH0018					
St. Dunstan's and Stepney Green		Tower Hamlets					
		CROWN					
		COPYRIGH					
		T					
PRODUCED USING SASPAC		00BGGH0018 RESERVED					
Table CT003 THEME TABLE ON							
ETHNIC GROUP - PEOPLE							
Table population: All people							
		White			Asian or Asian British		
	All People			Other			
		British	Irish	White	Indian	Pakistani	Bangladeshi
Total Persons	437	34	0	6	16	0	369
Males	218	19	0	3	9	0	184
Females	219	15	0	3	7	0	185
Age							
0-4	45	3	0	0	0	0	42
5-15	105	4	0	0	6	0	95
16-29	147	15	0	0	3	0	129
30-49	78	5	0	3	4	0	66
50-pensionable age	17	0	0	0	0	0	17
Pensionable age to 74	21	3	0	0	0	0	18
75 and over	6	3	0	0	0	0	3

& getting ready for the deluge of data.....

- Antipasti of Local Authority estimates in July, but next month.....
- OA boundaries & statistics – big datasets
- Experience of previous Censuses.....
 - Census specialists need to help all the occasional users
 - Is the IT in place?
 - Anticipating users' needs and questions?
 - Support? Training?
 - Services from VARs
- Maximise use & value – we need to plan!

Keith Dugmore
Demographic Decisions Ltd.

Tel: (0044) 020 7834 0966

Email: dugmore@demographic.co.uk

Web: www.demographic.co.uk