Using External Data in Market Research

MRS Census & Geodemographics Group
31 March 2003

Future Developments

Keith Dugmore
Demographic Decisions
Agenda

• Demand – the users of external datasets
• Supply – current situation
• Supply – the future big changes
Demand - some serious users of external data

- Abbey, Argos, Barclays, Boots, HBOS, M&S, MSFS, Nationwide, RBS, Saga, Sainsbury, Tesco, Whitbread, Woolworth, Yell
- National businesses
- Database analysis, Area & Site analysis, Market Research
- Which datasets?
# Supply – the shopping list

<table>
<thead>
<tr>
<th>Customer Files (Postcoded)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Area Statistics</td>
</tr>
<tr>
<td>Census - Residents</td>
</tr>
<tr>
<td>Updates</td>
</tr>
<tr>
<td>Projections</td>
</tr>
<tr>
<td>Census - Workplace</td>
</tr>
<tr>
<td>Geodem profiles</td>
</tr>
<tr>
<td>Income estimates</td>
</tr>
<tr>
<td>Market estimates</td>
</tr>
<tr>
<td>Shopping catchments</td>
</tr>
<tr>
<td>Digital Boundaries</td>
</tr>
<tr>
<td>Postal</td>
</tr>
<tr>
<td>Administrative</td>
</tr>
<tr>
<td>TV Areas</td>
</tr>
<tr>
<td>Map Data</td>
</tr>
<tr>
<td>Background - OS, AA</td>
</tr>
<tr>
<td>Road network - drivetimes</td>
</tr>
<tr>
<td>Points &amp; Sites</td>
</tr>
<tr>
<td>Retail Locations</td>
</tr>
<tr>
<td>Goad Plans</td>
</tr>
<tr>
<td>ATMs</td>
</tr>
<tr>
<td>Business Locations</td>
</tr>
<tr>
<td>Planning Applications</td>
</tr>
<tr>
<td>Postcode Directories &amp; Classifications</td>
</tr>
<tr>
<td>Mosaic (various)</td>
</tr>
<tr>
<td>Acorn (various)</td>
</tr>
<tr>
<td>Prizm</td>
</tr>
<tr>
<td>Cameo</td>
</tr>
<tr>
<td>Codepoint</td>
</tr>
<tr>
<td>EDs</td>
</tr>
<tr>
<td>Flood Risk</td>
</tr>
<tr>
<td>Lists</td>
</tr>
<tr>
<td>Electoral Roll</td>
</tr>
<tr>
<td>Lifestyles</td>
</tr>
<tr>
<td>Household classifications</td>
</tr>
<tr>
<td>Person classifications</td>
</tr>
<tr>
<td>Suppression files</td>
</tr>
<tr>
<td>CCJs</td>
</tr>
<tr>
<td>Businesses - CoHo &amp; Yell</td>
</tr>
<tr>
<td>Sample Surveys</td>
</tr>
<tr>
<td>Target Group Index</td>
</tr>
<tr>
<td>FES</td>
</tr>
<tr>
<td>AGB SuperPanel</td>
</tr>
<tr>
<td>Traffic flows</td>
</tr>
<tr>
<td>Pedestrian flows</td>
</tr>
</tbody>
</table>
Future Developments: the strategic changes

- Electoral Roll – restrictions
- 2001 Census – arrival
- Neighbourhood Statistics – development

........Government’s pivotal role as a supplier of raw material
Electoral Roll
Electoral Roll – restrictions

- The Register – name & address of 45m adults
- Commercial Use: opt-out from autumn 2002
- Direct marketing; sample frames; list cleaning; (but not credit checking)
- Public sector users
- Who has opted out?
- Alternative sources
- Restrictions on other registers?
2001 Census
The Census of Population

- Every 10 years – 29 April 2001
- Organised by Government – compulsory – aims for 100%
- England, Wales, Scotland and N Ireland
- Wide range of questions
- Statistics for small areas
2001: The main innovations

- Questions
- Coverage / One Number
- Coding – inc. Social Grade
- Geography
- Outputs
- Access

(beware some differences within the UK)
The Census - core subjects

Sex, age, marital status, relationship, ethnic group, migration, country of birth, long-term illness, occupation, industry, workplace, qualifications, sharing, tenure, rooms, central heating, car ownership... social class...
New Questions

• Religion
• General health
• Lowest floor level of accommodation
• Provision of unpaid personal care
• Income?? No!!
Maximising Coverage

- Population base - usual residents (no Visitors)
- New form design + Targeting of difficult areas + Post-back
- 94% response (+ 6% estimated)
“One Number”

• 2001: Coverage survey; estimation of the total population; imputation of missing records & values
• Statistics available for the whole country, rather than county-by-county
• Benefits? Costs?
• Insurance!!!
Coding

• Every form has been fully coded - not just 10% for difficult questions
• ....goodbye to sampling error for small areas
• Also, new classifications:
  – NS-SEC (National Statistics Socio-Economic Classification)
  – Social Grade approximation, following extensive research by the MRS
Social Grade

• Social Grade has been a key omission for market researchers and business community – unable to profile population by key discriminator

• Government classifications such as Social Class (RGSC) and Socio-Economic Group (SEG) have not been adequate substitutes

• MRS Census and Geodemographics Group – algorithm to derive social grade based on Census variables
Social Grade approximation

- Particularly crucial since 2001 Census contains no information on income or wealth
- Algorithm has been thoroughly tested to prove robustness and accuracy
- Classification will be hard-coded onto Census dataset and form a standard output variable
- Social Grade tables will be part of the standard tables set – based on both households and individuals
Geography

• Collection (Enumeration) areas built from postcodes
• Output Areas will be quite separate - also built from postcodes
The Starting Point

- ADDRESSPOINT™ locations
- ED boundaries
- Road centrelines
- OS background map

Supplied by Census Division ONS
Postcode Boundaries

- Draw boundaries around individual addresses
- Do not cross ward or parish boundaries
Postcode Boundaries

- Erase lines between addresses with the same postcode
- Usable postcode boundaries?
Postcode Boundaries

- Adjust postcode boundaries to follow nearby ED boundaries and roads
- Use as building blocks
Output Area Design

Initial Random Aggregation of Building Blocks

Iterative Recombination

Design Constraints (Contiguity, Thresholds, Homogeneity, Size, Shape)

2001 Output Areas

Supplied by Census Division ONS
Output Area Design

- Examine all postcodes bordering two OAs
- Move the highlighted postcode from B into A?
- Better homogeneity and population?
Output Areas - Lincoln

Population

Homogeneity

Supplied by Census Division ONS
Output Areas - Lincoln

Output Areas
Postcodes
Addresses
Output Areas – the benefits

- Built from postcodes
- Smaller than 1991 – 125 households
- Designed for homogeneity
- Boundaries available in digital form
- Directory to link postcoded records of survey respondents or customers
- Building bigger areas – postal (sectors & districts) & administrative (wards & above)
Access – Licensing – it’s FREE!

• 1991 – Public services mass purchases; others bought through Census Agencies
• 2001:
  • Public services - “Census Access” - free use of most outputs
  • Now extended to all users – “Click-Use”
  • Distribution Licenses & Approved Suppliers
  • Huge implications for users (& Suppliers)
Access - Technology

- ONS’s enthusiasm for the Web
- Big files - CD, DVD
- Dissemination of data is the priority
- Analysis - experienced users already have their own software tools
- Analysis - creating new markets by offering simple facilities over the Web?
The Census outputs, and the first results
What do we want?
Maximum detail…
(Subjects AND Geography)
but…
“Disclosure Control”
### The statistical Outputs

<table>
<thead>
<tr>
<th>The major Outputs (ONS for England &amp; Wales)</th>
<th>Month</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key Statistics.</strong> (c400 counts). LA level</td>
<td>February</td>
</tr>
<tr>
<td><strong>OAs - Boundaries; P'code Dir; Lookup to higher geogs</strong></td>
<td>late March?</td>
</tr>
<tr>
<td><strong>Key Statistics.</strong> OAs, &amp; Wards (as at Dec 02)</td>
<td>May</td>
</tr>
<tr>
<td><strong>Census Area Stats.</strong> (c8,000) OAs &amp; Wards</td>
<td>May</td>
</tr>
<tr>
<td><strong>Standard Tables.</strong> (c25,000) Wards &amp; above</td>
<td>May</td>
</tr>
<tr>
<td><strong>Profiles.</strong> (c100 counts) Postal Sectors (best fit)</td>
<td>July</td>
</tr>
<tr>
<td><strong>Key Stats &amp; CAS.</strong> Postal Sectors (best fit)</td>
<td>July</td>
</tr>
<tr>
<td><strong>Special Workplace Stats.</strong> (Origin / Dest) OAs &amp; above</td>
<td>Sept / Oct</td>
</tr>
<tr>
<td><strong>Special Migration Stats.</strong> (Origin / Dest) Wards &amp; above</td>
<td>Sept / Oct</td>
</tr>
<tr>
<td><strong>ONS Geodem Classification.</strong> OAs in UK</td>
<td>December</td>
</tr>
</tbody>
</table>
**Additional bespoke products**

- **Sample of Anonymised Records (SARs)**
  - An invaluable innovation from 1991
  - Allows extensive profiling of particular subpopulations
  - Three SARs for 2001
    - Individuals SAR – 3% sample with geographic threshold of 90,000
    - Household SAR – 1% sample of households with Government Office Region geography
    - SAM – individual SAR with more detailed variables at GOR level of geography
- **Special tables** – ONS’s service
The early results

• National trends
  – Total populations by country
  – Over 60’s now exceed under 16’s
  – Fewer men than women for all ages >21

• Local news
  – Total populations – some surprises
  – Contrasts across the UK
How will the results be used?
The growing pyramid

Numbers / Sophistication

Time

The growth of knowledge / company commitment

Numbers / Sophistication

Time

- Team with diverse toolkit
- Simple Toolkit OR Licensed System
- Ad hoc Projects
- Site Reports, List, Profiling
Current planning for arrival of the data

- Users’ needs
  - analysing customers & prospects / assessing areas & sites / planning sample surveys
- Getting hold of the data – data volumes and packaging
- New geodemographic classifications
- New models
- Dissemination within companies
- Integration – many other datasets
Plus….NeSS
(Never Expected Such Statistics)
Neighbourhood Statistics – the early arrivals

<table>
<thead>
<tr>
<th>Topic</th>
<th>Datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crime</td>
<td>Notifiable Offences</td>
</tr>
<tr>
<td>Economic Deprivation</td>
<td>Family Credit claimants</td>
</tr>
<tr>
<td></td>
<td>Income Support claimants</td>
</tr>
<tr>
<td>Education Skills &amp; Training</td>
<td>University admissions by residence</td>
</tr>
<tr>
<td></td>
<td>Pupils at Key Stage 2</td>
</tr>
<tr>
<td>Health</td>
<td>Attendance Allowance claimants</td>
</tr>
<tr>
<td></td>
<td>Disability Living Allowance claimants</td>
</tr>
<tr>
<td></td>
<td>Drug mis-users</td>
</tr>
<tr>
<td></td>
<td>Hospital episodes</td>
</tr>
<tr>
<td>Housing</td>
<td>Dwellings by Council Tax band</td>
</tr>
<tr>
<td></td>
<td>Changes of ownership by dwelling price</td>
</tr>
<tr>
<td>Work Deprivation</td>
<td>Jobseekers Allowance claimants</td>
</tr>
<tr>
<td></td>
<td>Incapacity Benefit claimants</td>
</tr>
<tr>
<td></td>
<td>VAT registered enterprises</td>
</tr>
<tr>
<td></td>
<td>Employee jobs</td>
</tr>
<tr>
<td></td>
<td>Unemployment claimants</td>
</tr>
</tbody>
</table>
House prices
NeSS Development

- Policy Action Team 18 – Better Information
- England & Wales; Scotland; N Ireland
- Websites
- Initial datasets – Wards (as in 1998)
- Future datasets – many, including Income estimates
- Future geography – as for the Census
The Census / NeSS will change the future

- Free data
- Common geography
- Boundaries & directories to link other datasets
- The permutations are infinite
- Existing organisations – use will spread & deepen
- New users in large numbers
- Some Resellers will rise to the challenge
Keith Dugmore
Demographic Decisions Ltd.

Tel: 020 7834 0966
Fax: 020 7834 4304
Email: dugmore@demographic.co.uk
Web: www.demographic.co.uk