HOW TO BUILD GEODEMOGRAPHICS FROM BIG DATA

March 2016

Graham Smith, Associate Director

CACI

WELCOME

BIG DATA & GEODEMS – THE STORY SO FAR

NEW OPPORTUNITIES FOR GEODEMOGRAPHICS

DATA PRIVACY & KEY CONSIDERATIONS

CACI

CACI : 40 YEARS' DATA EXPERIENCE

40

YEARS'EXPERIENCE OF UK CONSUMERS

4bn

DATA POINTS PROCESSED ANNUALLY

100+

DATA SUPPLIERS & PARTNERS

16

DATA PRODUCTS & INSIGHT TOOLS

acorn

LEADING GEODEMOGRAPHIC PRODUCT

650

VARIABLES AGAINST 48M CONSUMERS

CACI

THE NEW FACE OF CACI DATA

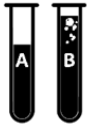
CACI | DATA LAB

PASSION

A team of over 20 data statisticians, planners and visualisation experts working collaboratively in one location.

MISSION

Our mission is to create and provide access to data and insight that excites, inspires and delivers value to our colleagues and clients.



INVESTMENT

An annual budget of over £2m to support the development, innovation and delivery of data-driven solutions.

GOVERNANCE

New enterprise-wide 'Compliance Task force' formalised to provide best practice advice on data compliance.

BIG DATA & GEODEMS – THE STORY SO FAR

CACI

THE INFORMATION EXPLOSION

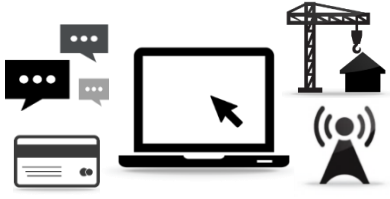
"Every 2 days we create as much information as we did from the beginning of time until 2003."

CACI

“BIG DATA”

VOLUME

VARIETY



VELOCITY

VERACITY

CACI

GEODEMOGRAPHICS

“The analysis of people by where they live”

Sleight, 2004

CACI

INPUTS TO GEODEMS IN THE 1970s

CACI

INPUTS TO GEODEMS IN THE 1970s

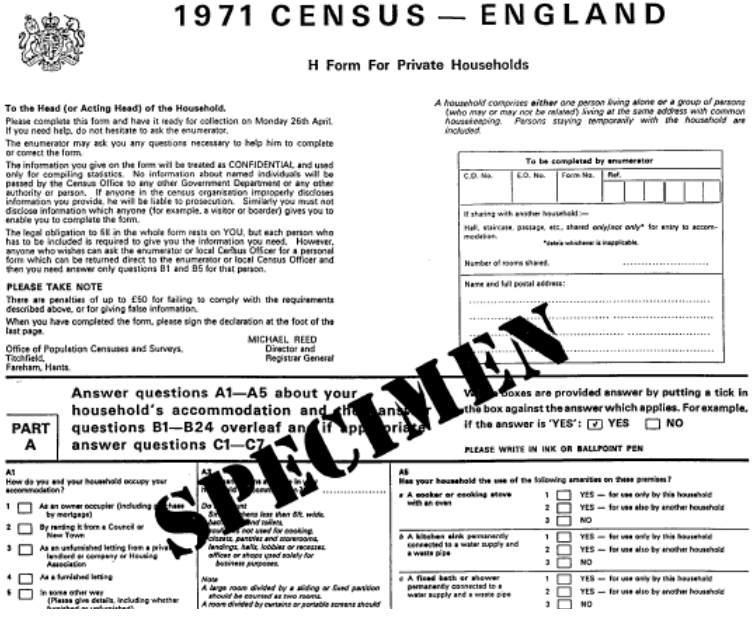
GOVERNMENT / ADMINISTRATIVE

CACI

CENSUS AS BIG DATA

Individual

Age
Marital Status
Employment
Country of birth
Qualifications
Travel to work
Children



1971 CENSUS — ENGLAND

H Form For Private Households

To the Head (or Acting Head) of the Household.

Please complete this form and have it ready for collection on Monday 26th April. If you need help, do not hesitate to ask the enumerator.

The enumerator may ask you any questions necessary to help him to complete or correct the form.

The information you give on the form will be treated as CONFIDENTIAL and used only for compiling statistics. No information about named individuals will be passed by the Census Office to any other Government Department or any other authority or person. If anyone in the census organisation improperly discloses information you provide, he will be liable to prosecution. Similarly you must not disclose information which anyone (for example, a visitor or boarder) gives you to enable you to complete the form.

The legal obligation to fill in the whole form rests on YOU, but each person who has to be included is required to give you the information you need. However, anyone who wishes can ask the enumerator or local Census Officer for a personal form which can be returned direct to the enumerator or local Census Officer and then you need answer only questions B1 and B5 for that person.

PLEASE TAKE NOTE

There are penalties of up to £50 for failing to comply with the requirements described above, or for giving false information.

When you have completed the form, please sign the declaration at the foot of the last page.

MICHAEL REED
Director and
Registrar General

Office of Population Censuses and Surveys,
Titchfield,
Fareham, Hants.

A household comprises either one person living alone or a group of persons (who may or may not be related) living at the same address with common housekeeping. Persons staying temporarily with the household are included.

To be completed by enumerator

C.D. No. | E.D. No. | Form No. | Ref.

If sharing with another household:—

Hall, staircase, passage, etc., shared only/not only* for entry to accommodation.

*delete whichever is inapplicable.

Number of rooms shared. .....

Name and full postal address:

.....

.....

.....

.....

SPECIMEN

PART A

Answer questions A1—A5 about your household's accommodation and [illegible] questions B1—B24 overleaf and [illegible] answer questions C1—C7.

Where boxes are provided answer by putting a tick in the box against the answer which applies. For example, if the answer is 'YES': ☑ YES ☐ NO

PLEASE WRITE IN INK OR BALLPOINT PEN

A1
How do you and your household occupy your accommodation?

1 ☐ As an owner occupier (including [illegible] by mortgage)
2 ☐ By renting it from a Council or New Town
3 ☐ As an unfurnished letting from a private landlord or company or Housing Association
4 ☐ As a furnished letting
5 ☐ In some other way (Please give details, including whether [illegible])

A3
On the ground floor [illegible] rooms less than 6½ wide, [illegible] bathrooms, toilets, passages not used for cooking, [illegible] cupboards, pantries and storerooms, landings, halls, lobbies or recesses, offices or shops used solely for business purposes.

Note
A large room divided by a sliding or fixed partition should be counted as two rooms. A room divided by curtains or portable screens should

A5
Has your household the use of the following amenities on these premises?

a A cooker or cooking stove with an oven — 1 ☐ YES — for use only by this household; 2 ☐ YES — for use also by another household; 3 ☐ NO

b A kitchen sink permanently connected to a water supply and a waste pipe — 1 ☐ YES — for use only by this household; 2 ☐ YES — for use also by another household; 3 ☐ NO

c A fixed bath or shower permanently connected to a water supply and a waste pipe — 1 ☐ YES — for use only by this household; 2 ☐ YES — for use also by another household; 3 ☐ NO

Household

House Tenure
Communal space
Number of rooms
Car ownership
Household amenities

"IT'S THE BIG FORM WITH THE BIG JOB TO DO"

CACI

INPUTS INTO GEODEMS IN 2016

CACI

INCREASED VOLUME OF DATA SOURCES

- GOVERNMENT
- ADMINISTRATIVE
- OPEN DATA
- CROWD-SOURCED
- COMMERCIAL
- TRANSACTIONAL
- POOLED
- PRIVATE

CACI

INPUTS TO GEODEMS IN 2016

GOVERNMENT / ADMINISTRATIVE

- Census and mid-year estimates
- Land Registry / Registers of Scotland
- Local Authority registers
- Department for Work and Pensions
- Social housing lists

COMMERCIAL

- Name and address data
- Date of birth data
- Age limited housing
- Private rental data
- Lifestyle surveys

RESEARCHED & DERIVED

- High-rise social housing
- Student accommodation
- Communal populations
- Age profiles
- Ethnicity & country of birth

INPUTS TO GEODEMS IN 2016

GOVERNMENT / ADMINISTRATIVE	COMMERCIAL	RESEARCHED & DERIVED
Census and mid-year estimates	Name and address data	High-rise social housing
Land Registry / Registers of Scotland	Date of birth data	Student accommodation
Local Authority registers	Age limited housing	Communal populations
Department for Work and Pensions	Private rental data	Age profiles
Social housing lists	Lifestyle surveys	Ethnicity & country of birth

CACI

GOVT / ADMIN DATA – LAND REGISTRY

Price Paid Dataset

Open Data

24m

Source of house price information consisting of more than 24 million definitive records dating back to January 1995.

Used for a range of housing variables and helps to identify new-build housing

Commercial Ownership Data

Available under licence

3.2m

> 3.2 million title records of freehold and leasehold property in England and Wales.

Used to identify locations of private housing owned by social housing providers

CACI

COMMERCIAL DATA – NAMES & ADDRESSES

Edited Electoral Roll

Lifestyle Data Sources

Transactional Databases

Niche Databases

48m

Consumer Register

Used to provide the "spine" and validation file for individual and household level variables

CACI

FILLING THE GAPS THROUGH DERIVED DATA

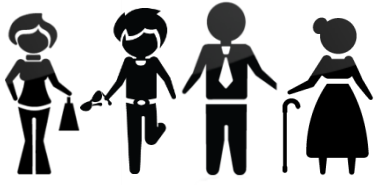
Names & Addresses

Date of birth data sources

Modelled ages, based on forename and other known attributes

Full age profile for every residential household and postcode

% by Age	20-24	35-39	70+
Margaret	0%	1%	48%
George	5%	4%	39%
Prakash	3%	17%	2%
Max	37%	7%	4%



Age Profiles by Name and Gender

CACI

BUILDING GEODEMOGRAPHIC CLASSIFICATIONS

Separate processes for different sources

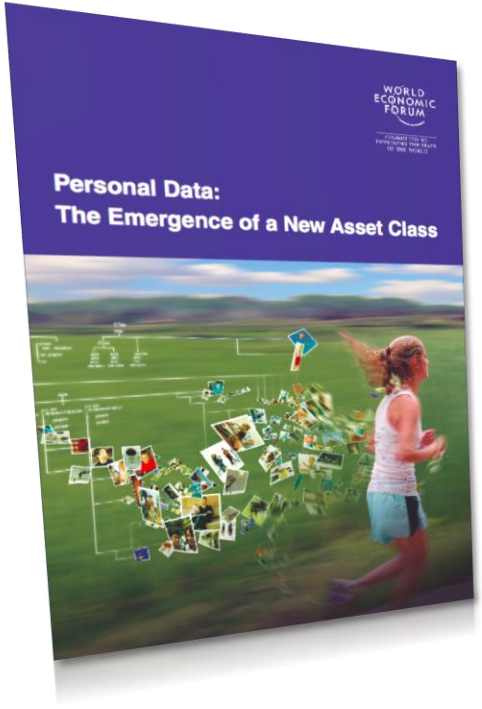
Sources are updated regularly, and many are available at postcode, household or individual level

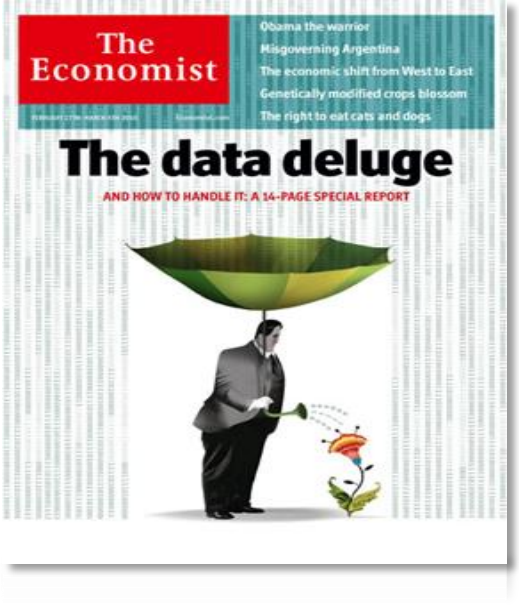
Census remains important for calibration and detailed demographics

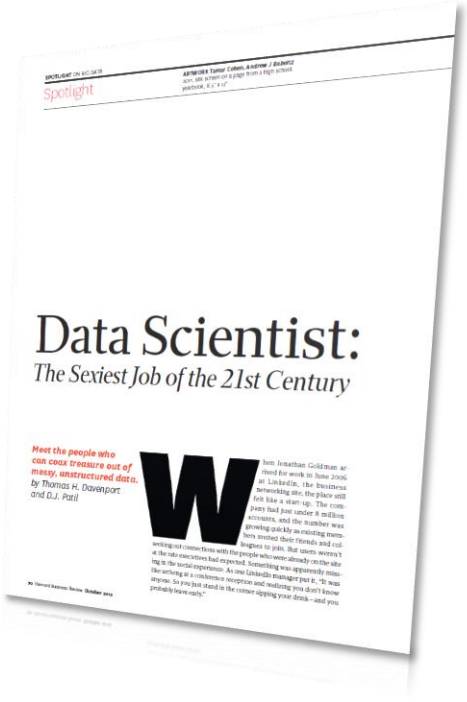
CACI

NEW OPPORTUNITIES FOR GEODEMOGRAPHICS

BUSINESSES ARE HOOKED ON DATA



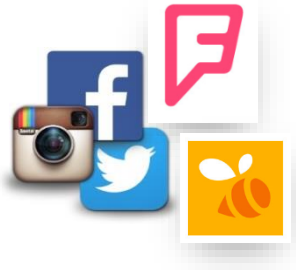




CACI

BIG DATA OPPORTUNITIES

SOCIAL MEDIA



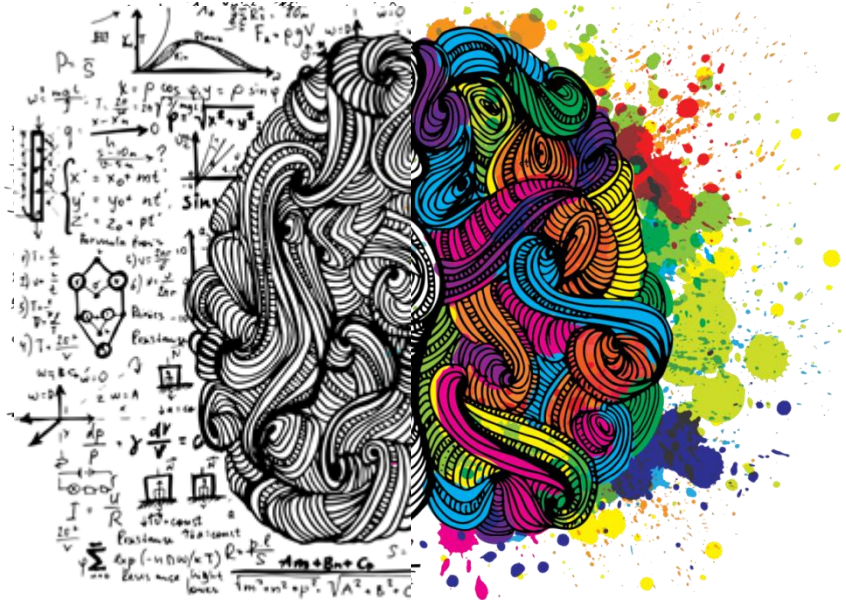
CARD TRANSACTIONS

MOBILES, APPS,
SMART METERS

- ✓ INTERESTS & ATTITUDES
- ✓ BROWSING BEHAVIOUR
- ✓ SPENDING PATTERNS
- ✓ CONSUMPTION
- ✓ LOCATION
- ✓ GEO-TEMPORAL MOVEMENT

HOW DO WE EXTRACT THE GEO AND THE DEMOGRAPHICS ?

CUSTOMER DATA AND GEODEMS



- Customer data is only part of the story
- Purchasing decisions are still predicated on affluence and lifestage
- In many cases customer data only provides basic demographics

GEODEMOGRAPHICS

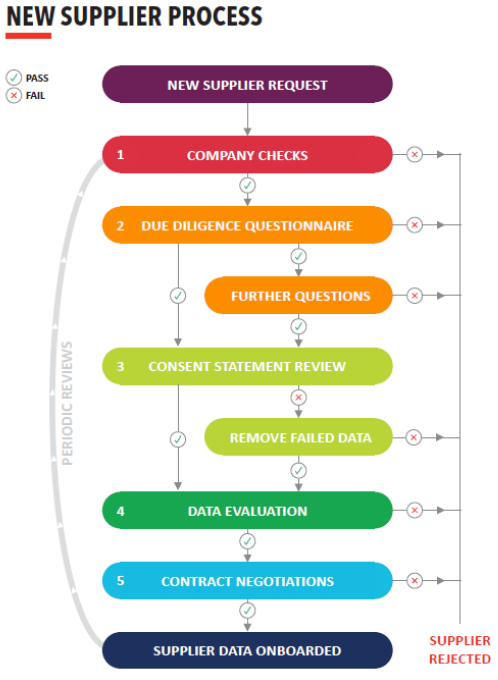
ADD COLOUR AND CONTEXT TO CUSTOMER SEGMENTATIONS AND ANALYSIS

CACI

DATA PRIVACY & KEY CONSIDERATIONS

PRIVACY AND DATA PROTECTION

- AWARENESS
- COMPLIANCE
- SUPPLY



SIX KEY CONSIDERATIONS FOR USING BIG DATA

CACI

SIX KEY CONSIDERATIONS FOR USING BIG DATA

1. DO WE HAVE CORRECT PERMISSIONS FOR USE?
2. ARE THERE BARRIERS TO OBTAINING OR PROCESSING THE DATA?
3. CAN WE LINK DATASETS? WHAT ARE THE CONNECTORS?
4. CAN WE LINK TO AN ADDRESS?
5. ARE THERE SUFFICIENT BASIC DEMOGRAPHIC VARIABLES?
6. IS THERE INHERENT BIAS? HOW COMPLETE A VIEW IS IT?

SUMMARY – AND SOME PARTING THOUGHTS

- Geodemographics are still, and will continue to be, as relevant now as they were in the 1970s
- Further democratisation of data and expansion of Open Data is expected
- The volume of data collected by private companies will increase at an ever faster pace
- New opportunities for Big Data (and small data) to input into customer analytics and geodemographics should be embraced
- Privacy and compliance concerns may bring about more demand for anonymised aggregated data and sample-based surveys
- A return to focus on postcode geodemographics rather than individuals?

THANK YOU

CACI

CACI