

Using OAC to create new insight from the Expenditure and Food Survey

Demographics User Group Conference
The Royal Society
8th October 2009

Martin Callingham
Visiting Professor
School of Geography
Birkbeck College
University of London

Introduction

Insight

Expenditure and Food Survey

Output area classification and geodemographics

How OAC adds Insight

Conclusions

What is insight?

Insight?

Information



Knowledge



Wisdom

Is there such a thing as a fact?

Factiness



Oughtiness

The role of imagination in insight

Even unimaginative people are very good at it

It is part of the working of the brain

The brain works by interpolating between the facts

That is, informed imagination

Otherwise we would have fact overload.

Too much detail

We have to simplify our world of raw data

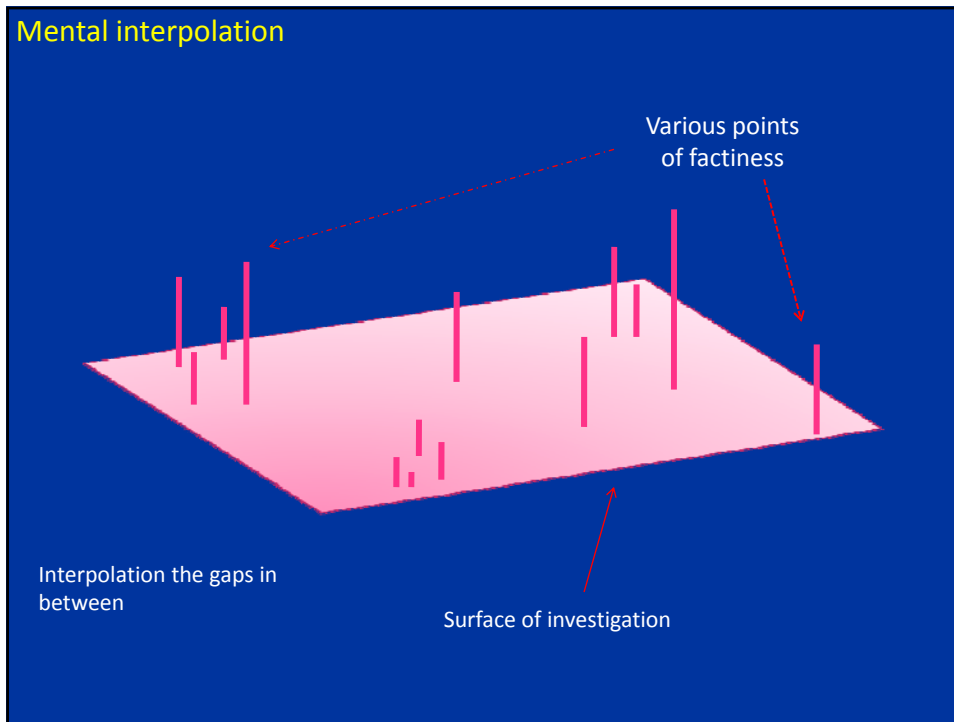
Much of what we see, we make up – based on experience

This is true of how we conceptualise as well

Therefore our world has a large contribution of imagination
in it

and occasionally, we end up thinking things incorrectly





Going down the wrong trouser leg

'Wrong end of the stickiness'

Happens all the time

Its when we get the major point of orientation wrong

And all our subsequent interpolation goes wrong too



Insight is cultural – when we know we are in an Insight rich world

Drive for insight is to help organisations make decisions

- So insight should be shared and be collective
- It is collaborative in the way it is built
- It has ambiguity within it
- It helps discussion
- One can have intelligent arguments, based on fact
- Clarity increases with discussion
- Decisions are made and understood

Insight tools

Cross tabulation and contrast

Triangulation

Self questioning

Contingent thinking - segmentation

Bouncing ideas

Safe cultural environment

A thirst for 'therefores'

What is the EFS?

Expenditure and Food Survey (EFS)

Conducted annually on about 7000 households

Collects

- Basic household information
- Sources of individual income
- Household expenditure in great detail

Uses a diary collection over a defined period of time per household

Therefore expenditure does not give complete picture of the nations household

But does give an excellent picture of the nation's expenditure

Its use is in conjunction with the RPI and National Accounts

Data is reported under a series of expenditure headings which are hierarchical

Categories of expenditure

Food & non-alcoholic beverage
Alcoholic Beverages, Tobacco
Clothing and Footwear
Housing, Water, Electricity
Furnishings, HH Equipment
Health expenditure
Transport costs
Communication
Recreation
Education
Total Restaurants and Hotels
Miscellaneous Goods and Services

Example of EFS components

- 01 Rice
- 02 Bread
- 03 Other breads and cereals
- 04 Buns, crispbread and biscuits
- 05 Cakes and puddings
- 06 Pastry (savoury)
- 07 Pasta products
- 08 Beef (fresh, chilled or frozen)
- 09 Pork (fresh, chilled or frozen)
- 10 Lamb (fresh, chilled or frozen)
- 11 Poultry (fresh, chilled or frozen)
- 12 Sausages
- 13 Bacon and ham
- 14 Offal, pâté etc
- 15 Other preserved or processed meat and meat preparations
- 16 Other fresh, chilled or frozen edible meat
- 17 Fish (fresh, chilled or frozen)
- 18 Seafood (fresh, chilled or frozen)
- 19 Dried, smoked or salted fish and seafood
- 20 Other preserved or processed fish and seafood
- 21 Whole milk
- 22 Low fat milk
- 23 Preserved milk
- 24 Yoghurt
- 25 Cheese and curd
- 26 Other milk products
- 27 Eggs
- 28 Butter
- 29 Margarine and other vegetable fats
- 20 Peanut butter
- 31 Olive oil
- 32 Edible oils and other edible animal fats
- 33 Other edible animal fats
- 34 Citrus fruits (fresh)

Rich data source

All this is great for detailed information about the expenditure on different categories and their components

But there are few household classifiers, so cross tabbing is very limited

Mean annual house hold income by category of house

Gross current income of household pa



Ratio of maximum to minimum 1.46

Mean annual house hold income by tenure of house

Gross current income of household pa



Ratio of maximum to minimum 2.89

Not much insight here

But the survey was not designed for this purpose

So how can we get more out of this survey?

How do we get extra mileage out of this survey

To use this survey for generating more insight, we need to be able to cross tab sensibly which means segmenting the respondents.

An easy way of doing this is attach geodemographic categories to the respondents via their postcode

Geodemographics and the Output Area Classification (OAC) in particular

Demographics explain about 70% of behaviour

So knowing the mixture of demographics in a group links to the behaviour of the group

Geodemographics identify small geographical areas of specified demographic mixtures

Therefore the average behaviour in an area can be estimated

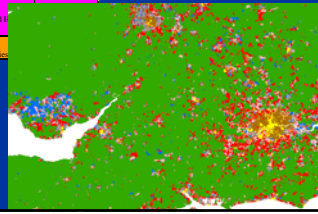
What are geodemographics?

What are geodemographics?

	A	E	F	G	H	I	J
1 OA	Age04	Age14	Age24-4	Age34-4	Age44-54	Age65+	sex
2 00JAN00001	7.20	14.60	21.90	33.00	14.40	0.00	
3 00JAN00002	4.30	16.10	26.30	25.00	18.70	0.00	
4 00JAN00003	4.70	14.50	21.30	15.70	37.50	0.00	
5 00JAN00004	3.40	11.60	27.40	36.30	15.20	0.00	
6 00JAN00005	3.90	9.70	15.10	35.30	20.40	0.00	
7 00JAN00006	6.00	14.90	28.00	29.00	14.20	1.00	
8 00JAN00007	2.40						
9 00JAN00008	2.00						
10 00JAN00009	7.40						
11 00JAN00010	3.10						
12 00JAN00011	7.60						
13 00JAN00012	8.00						
14 00JAN00013	2.30						

	A	B	C	D	E	F	G	H	I	J
1 OA	Agegroup	Orng	Subgroup	Age04	Age14	Age24-4	Age34-4	Age44-54	Age65+	sex
2 00JAN00001	3.30	3d	0.61	0.61	0.70	0.15	0.39	0.00		
3 00JAN00002	3.3a	3d	0.42	0.70	0.74	0.70	0.65	0.00		
4 00JAN00003	1.11	1d1	0.51	0.68	0.70	0.68	0.79	0.00		
5 00JAN00004	3.3a	3d	0.40	0.40	0.35	0.17	0.41	0.00		
6 00JAN00005	4.4b	4d	0.46	0.59	0.62	0.36	0.74	0.00		
7 00JAN00006	4.4b	4d	0.37	0.69	0.76	0.33	0.39	0.11		
8 00JAN00007	3.3a	3d	0.40	0.69						
9 00JAN00008	3.3a	3d1	0.32	0.41						
10 00JAN00009	3.3b	3d1	0.62	0.70						
11 00JAN00010	3.3a	3d	0.46	0.68						
12 00JAN00011	3.3a	3d1	0.62	0.63						
13 00JAN00012	1.11a	1d	0.66	0.76						
14 00JAN00013	2.2a	2d	0.35	0.90						

Super Groups 7	Groups 21	Sub-groups 52
1: Blue Collar Communities	1a: Terraced Blue Collar	3
	1b: Younger Blue Collar	2
	1c: Older Blue Collar	3
2: City Living	2a: Transient Communities	2
	2b: Settled in the City	2
3: Countryside	3a: Village Life	2
	3b: Agricultural	2
	3c: Accessible Countryside	2
4: Prospering Suburbs	4a: Prospering Younger Families	2
	4b: Prospering Older Families	4
	4c: Prospering Semis	3
	4d: Thriving Suburbs	2
5: Constrained by Circumstances	5a: Senior Communities	2
	5b: Older Workers	4
	5c: Public Housing	3
6: Typical Traits	6a: Settled Households	2
	6b: Least Divergent	3
	6c: Young Families in Terraced Homes	2
	6d: Aspiring Households	2
7: Multicultural	7a: Asian Communities	3
	7b: Afro-Caribbean Communities	2



Super Groups 7	Groups 21	Sub-groups 52
1: Blue Collar Communities	1a: Terraced Blue Collar	3
	1b: Younger Blue Collar	2
	1c: Older Blue Collar	3
2: City Living	2a: Transient Communities	2
	2b: Settled in the City	2
3: Countryside	3a: Village Life	2
	3b: Agricultural	2
	3c: Accessible Countryside	2
4: Prospering Suburbs	4a: Prospering Younger Families	2
	4b: Prospering Older Families	4
	4c: Prospering Semis	3
	4d: Thriving Suburbs	2
5: Constrained by Circumstances	5a: Senior Communities	2
	5b: Older Workers	4
	5c: Public Housing	3
6: Typical Traits	6a: Settled Households	2
	6b: Least Divergent	3
	6c: Young Families in Terraced Homes	2
	6d: Aspiring Households	2
7: Multicultural	7a: Asian Communities	3
	7b: Afro-Caribbean Communities	2

What can geodemographics be used for

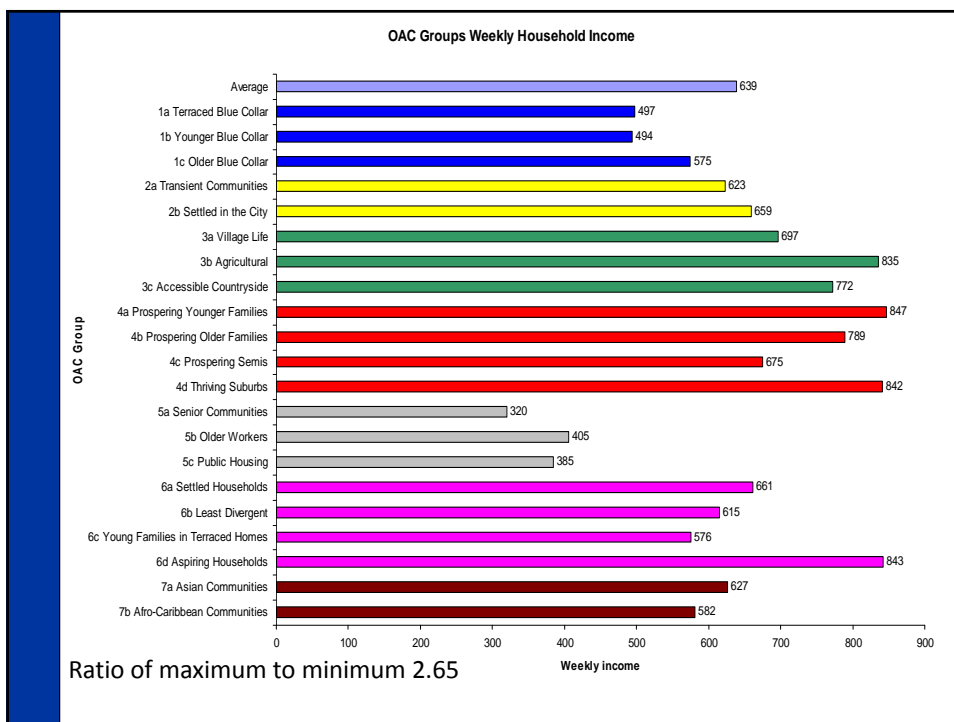
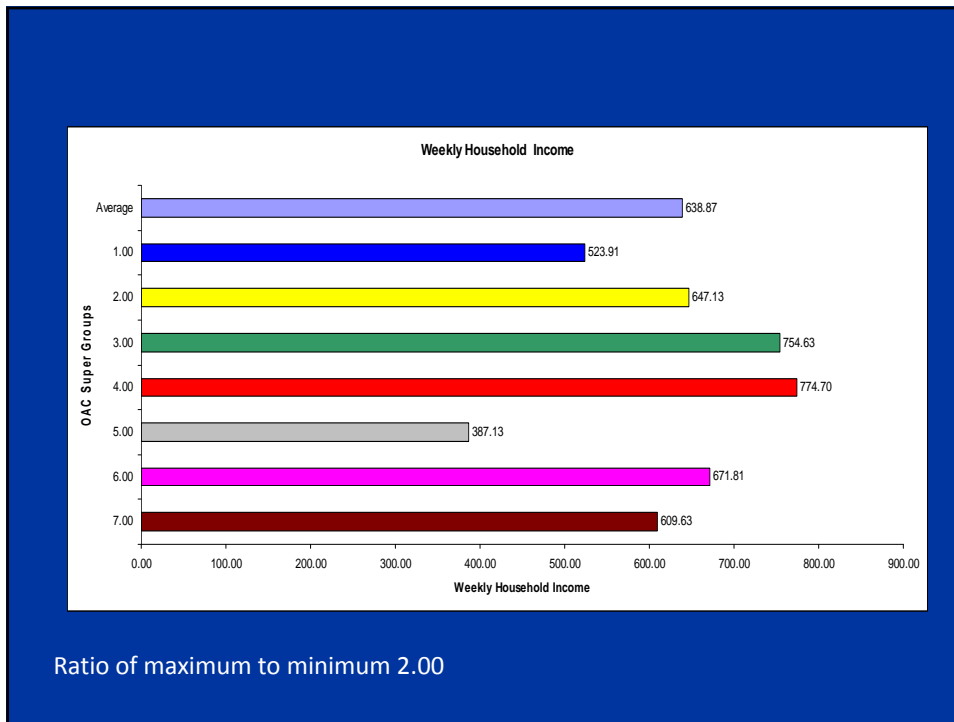
Straight forward mapping of an *area* which helps interpret it
Profiling *users* of a service which helps understand the nature of them
Modelling behaviour or attitude into small areas
Fusing different data sets

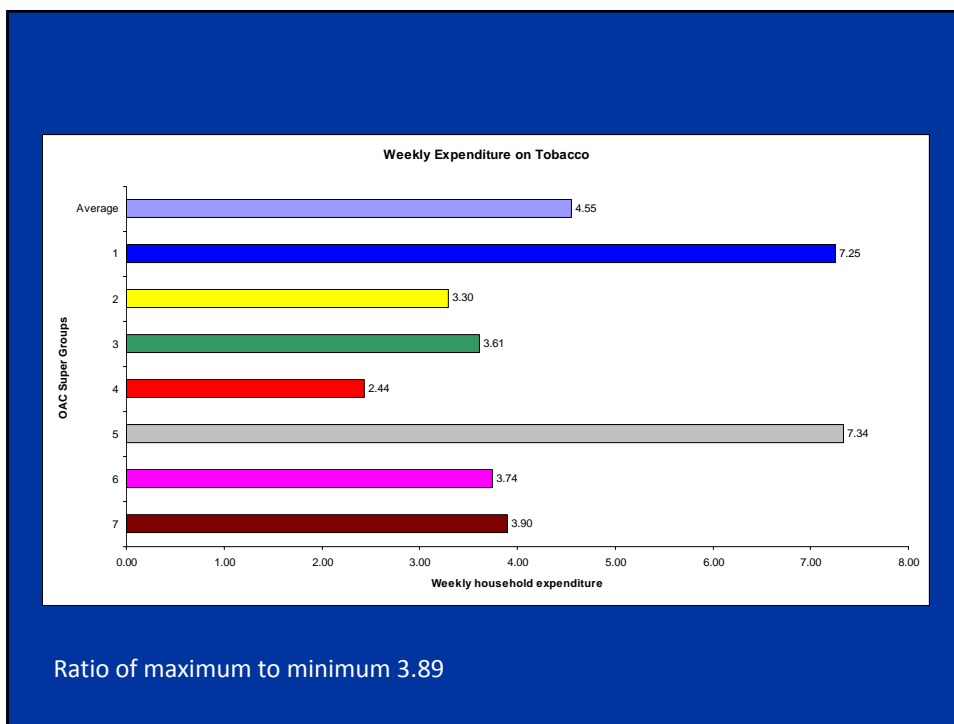
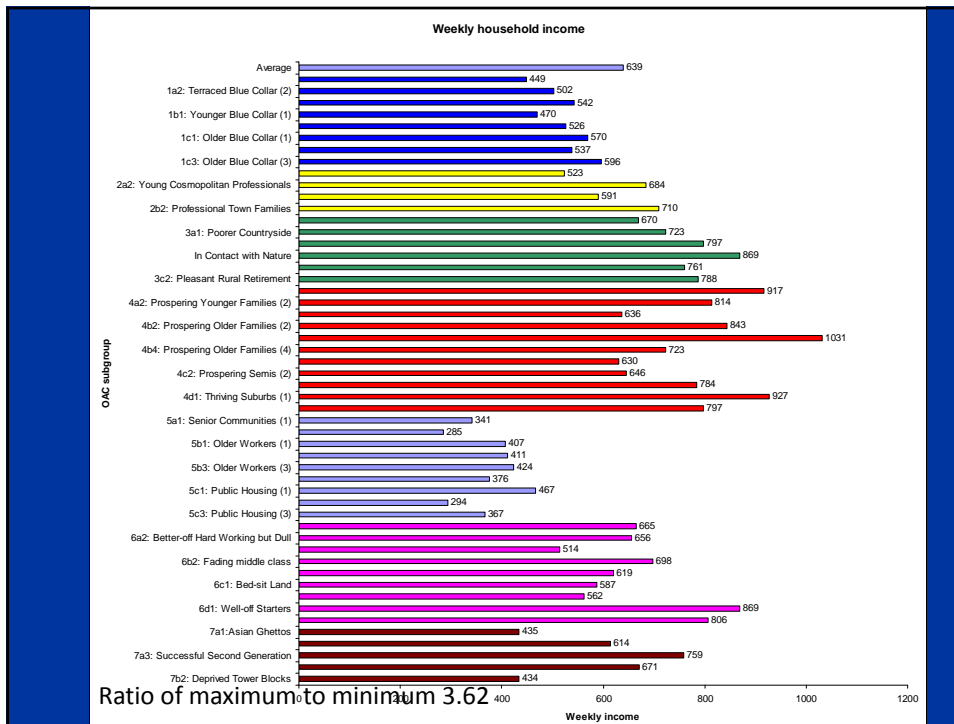
Examples of geodemographic are profiles fro income and expenditure on tobacco

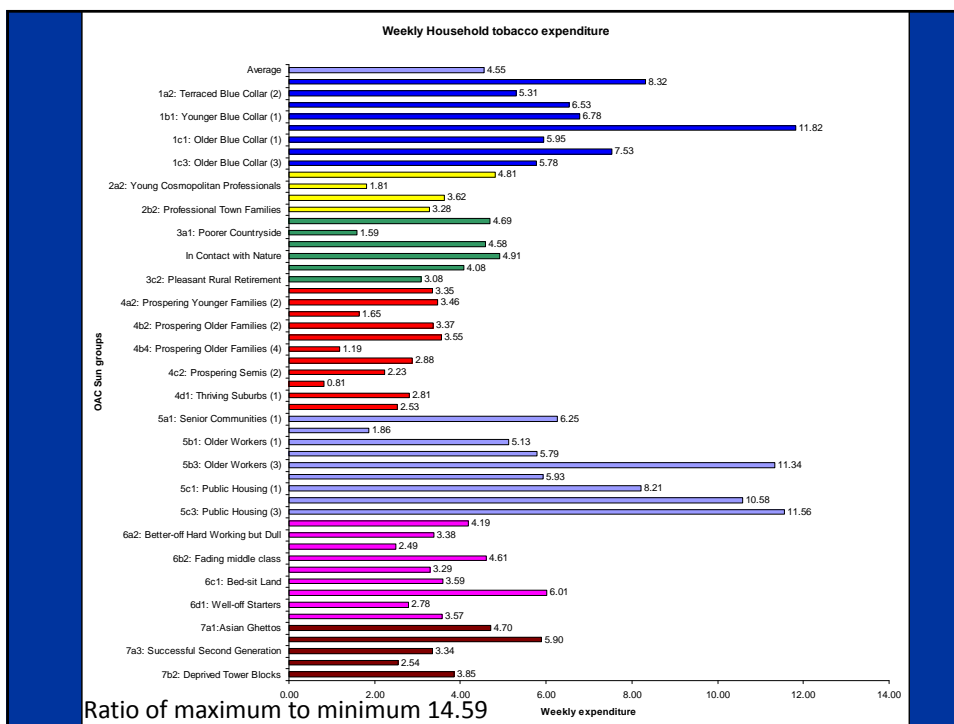
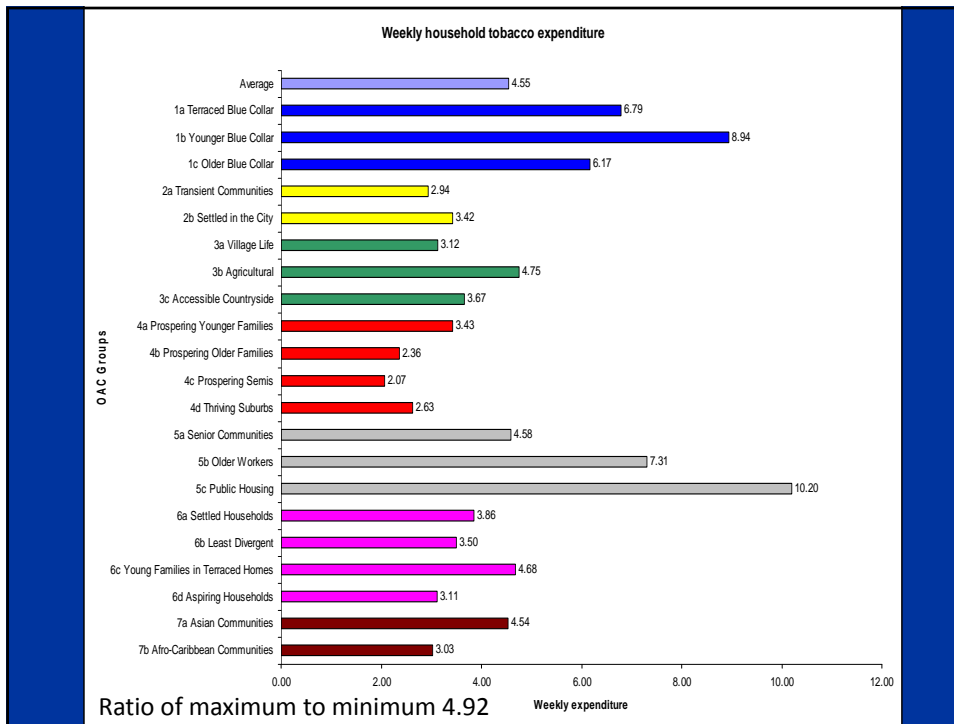
Following are:

the actual household income associated with each OAC demographic mix at the three hierarchical levels

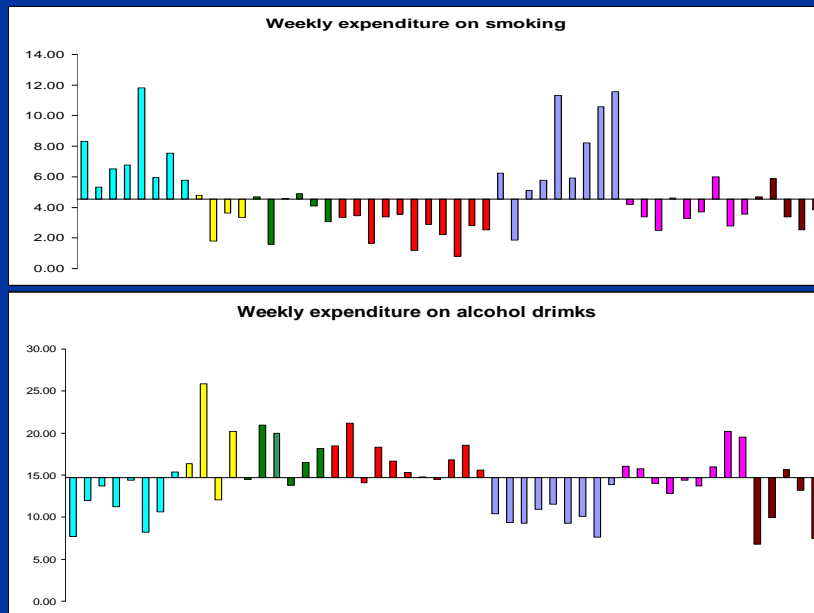
The actual expenditure of tobacco by each of the groups







Smoking



Enrichment

Simply adding a variable to the respondent which reflects their neighbourhood demographic mix has added a powerful way of segmenting the data and is a route to insight

Triangulation using OAC

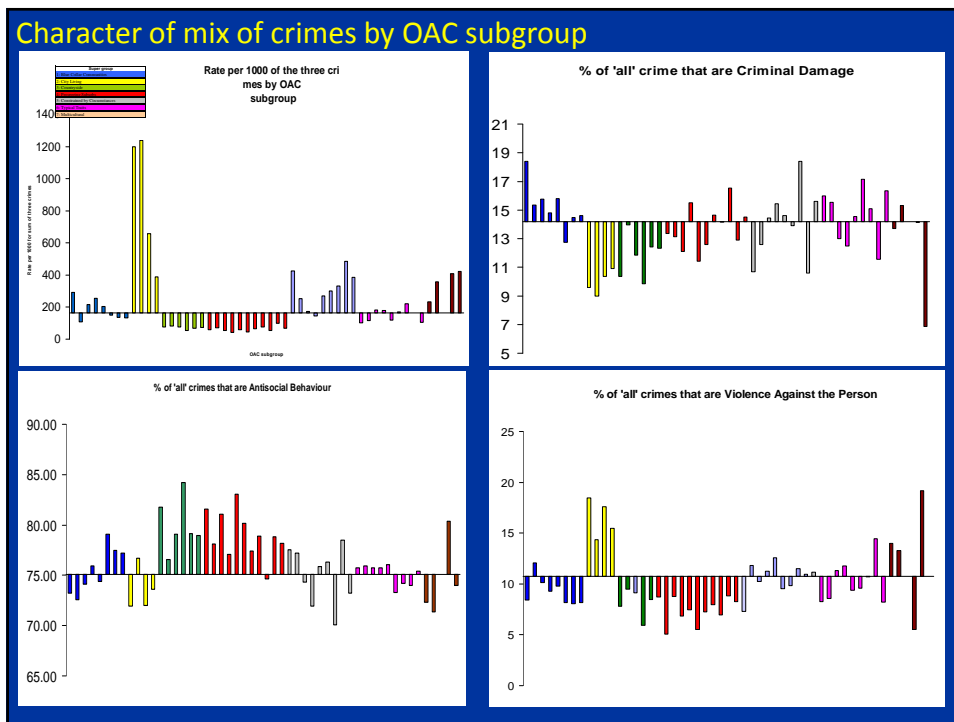
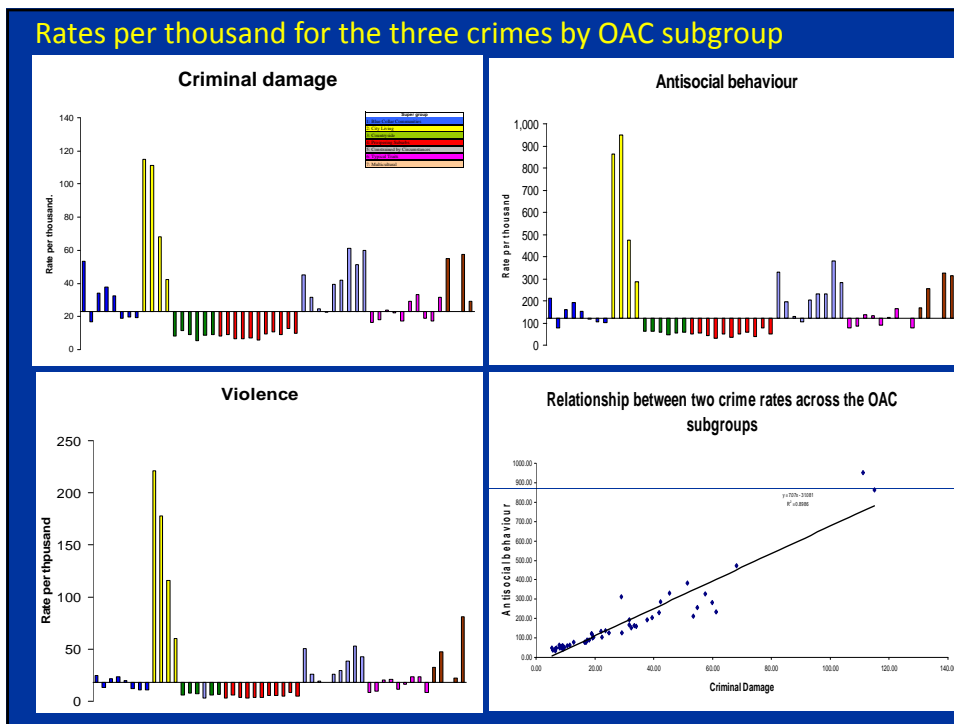
Tri-angulation

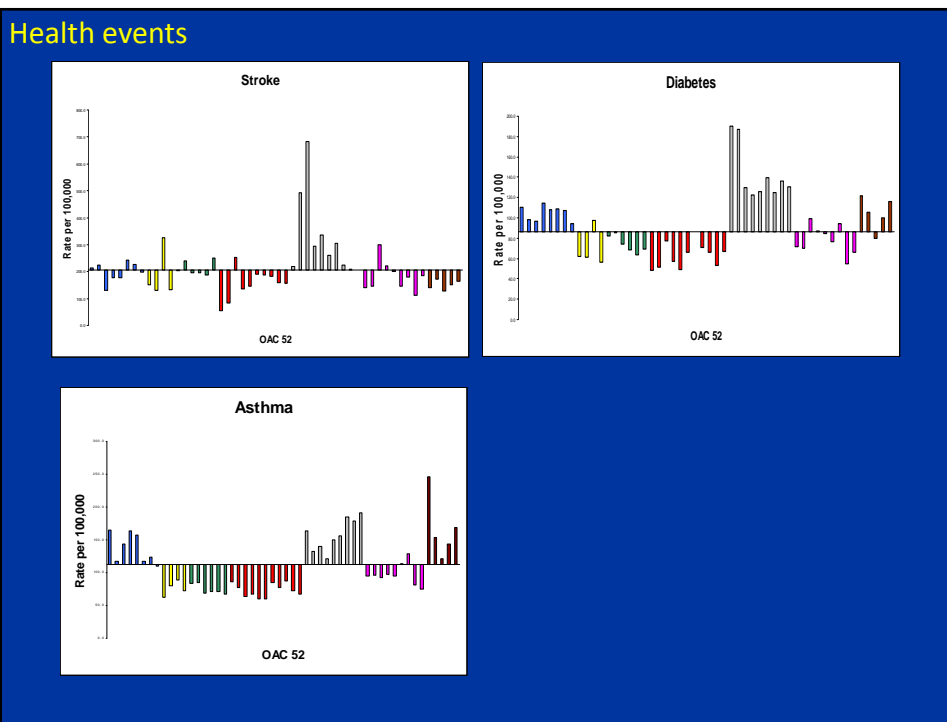
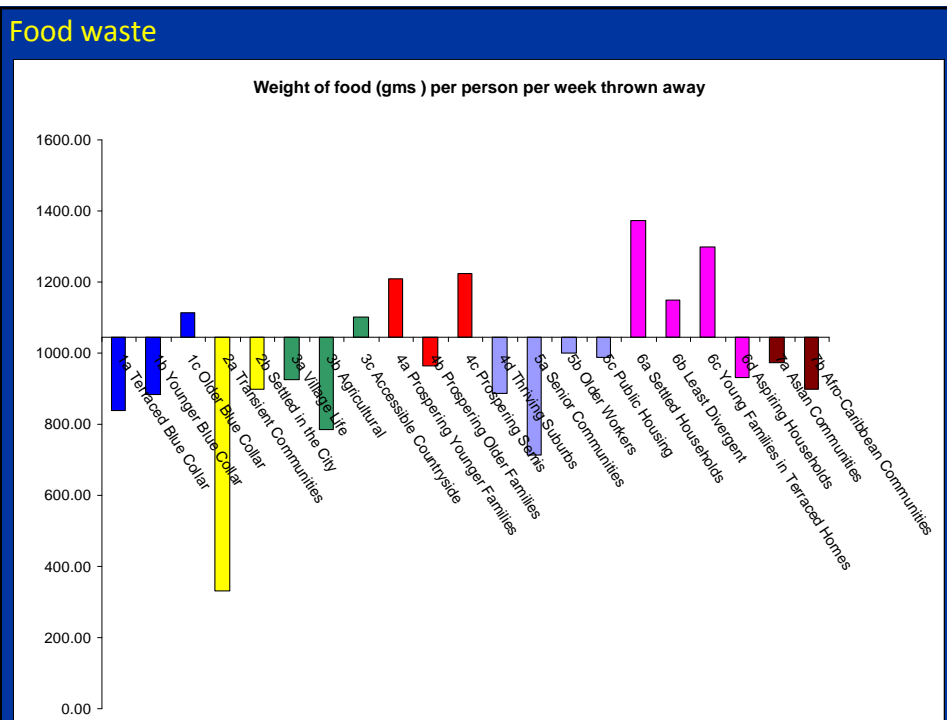
Triangulation is a process of examining different data source to see if it builds confidence in a conclusion

OAC is powerful at doing this as other surveys or data sets can be OAC coded

Following are examples from:

- Crime
- Household waste
- Health





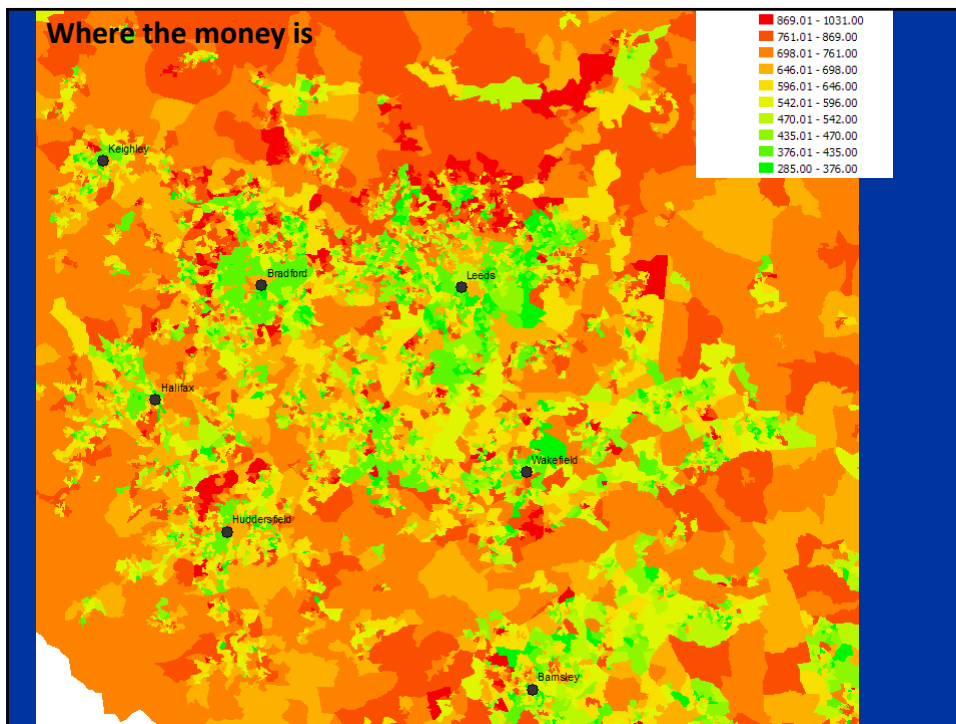
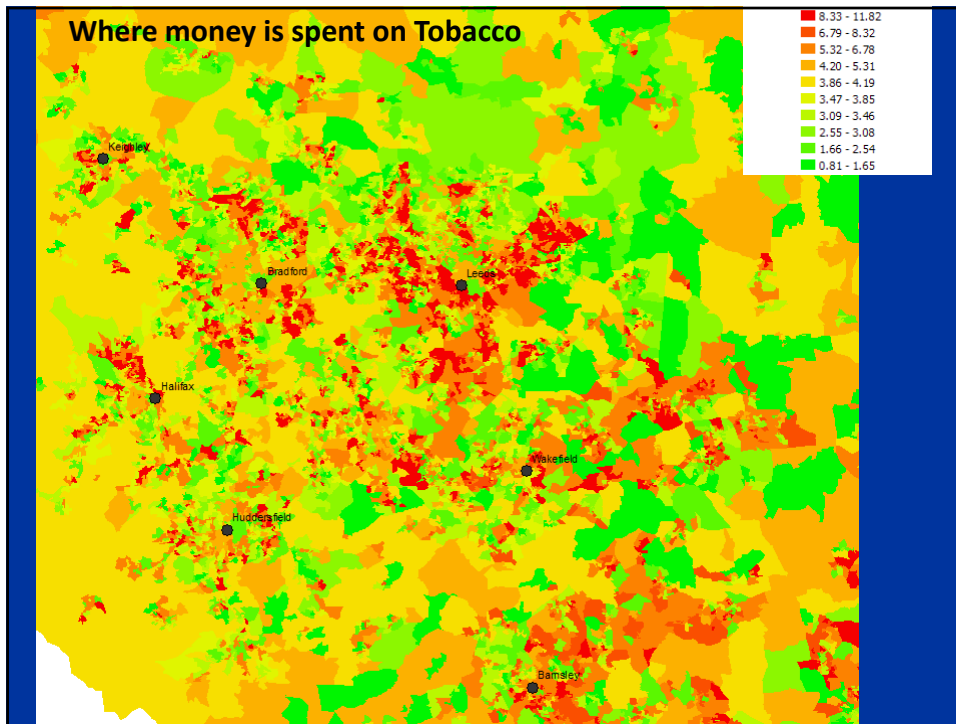
Locating the insight

Mapping modelled markets

As 230,000 output areas are classified, we now have an estimate of income and smoking in each

Therefore is possible to visualise it with a map

All the previously shown graphs can be converted to a map



More OAC coded surveys

As more waves of the EFS come available, the authority of the data will improve

A number of ONS surveys are to be OAC coded

This will increasingly provide better insight at very low cost

Hopefully, other location event data could also be OAC coded for example NeSS

How OAC is an important tool for insight

Cross tabulation and contrast

Triangulation

Self questioning

Contingent thinking - segmentation

Bouncing ideas

Safe cultural environment

A thirst for 'therefores'

The last 100 yards

We can amplify the value of this survey very cheaply

And in this form it can really contribute to insight by providing the real data points as launch points for our imagination and give excellent food for thought.

The public sector is facing 10% cuts so current assets must be made to work harder to help rebuild national prosperity

Choices between 'nice to have' and 'essential' expenditures have to be made.

So if an activity is worth having it should be made really good

If not, it should be axed.