



Finding Minority Samples

Tony Dent

Chairman

Sample Answers Ltd

Who am I?

“A statistician is defined as somebody who is good with numbers, but lacks the personality to be an accountant.”

- The Economist, March 28, 1998.



Just one of a Crowd?

(an Approximation of statisticians)





My Brief

- How do we approach surveys in niche markets?
- Data Sources and their collection
- Does Targeted Sampling Improve efficiency?
- Can I still infer Findings Across the Population?
- What else can these resources offer?

The Problem

The Euro Controversy:

- **MORI (CBI) say:**
“Business Favours the Euro”
- **ICM (All Businesses) say:**
“Business is Against the Euro”

Why did they conflict?

- Because of the different sample frames!



The Context: Customer Focused IT Solutions

- “The application of information technology to building or strengthening the relationship between a company and its customers”
- Most popular applications are:
 - data warehousing, sales analysis/ target marketing, customer profiles, Market Analysis, etc...
 - source ATT/NCR Consultants Conference Dayton Ohio Oct. 93



Enabling Decisions

- Value of decisions increase with
 - the quality and the quantity of the data accessed
 - and the skill with which it is processed and reported

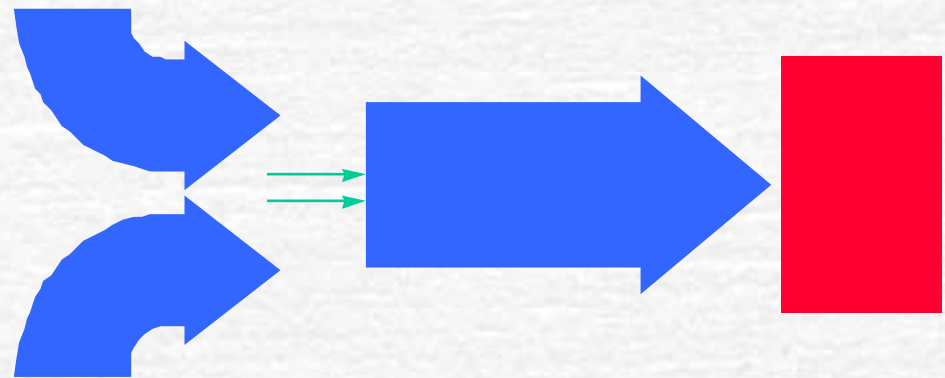


Marketing & Research Needs

- Definition of information required
- Collection of data
- Movement of data
- Analysis

In Summary

- GET *IT*
- MOVE *IT*
 - from place to place
 - from 'data' to 'Information'
- USE *IT*





Defining The Need

Understanding the context by

- Analysing existing data
 - search engines
 - multivariate analysis
- Establishing efficiency
 - determining the cost against the potential value
 - not forgetting the cost of time



Analysing Existing Data

- Re-affirming the context, demanding
 - elimination of bias (weighting)
- Confirming the relevance
- Reporting the importance
 - note: statistical significance is not the same as business importance



Marketer's Needs

- Information in context
 - Fitting with other data
- Speed of delivery
- Focus on opportunity
- Integrity and compliance

The Challenge

- To integrate the research
- into the marketing mix



Example “When Green Means Go for It!”

- Unipath, the Healthcare Division of Unilever, launched ‘Persona’ a computerised version of the Rhythm method of contraception, after 15 years of development.
- Persona is a ‘hand held’ device which monitors hormone levels in urine and indicates on which days women can have intercourse without fear of pregnancy. Green for safe, red for danger!

Example (cont.)

- The target market is women aged 25 to 40 in stable relationships and, apparently, 'traditional categories of consumers - ABC1's etc.. - do not enter into the marketing equation'.
- Research showed that women wanted 'discretion', therefore there was no fanfare launch, i.e. no 'light hearted fun campaign' as originally planned.



Value of Databases (I)

- Potential customers for Persona can be readily identified from lifestyle databases, targeting them specifically by
 - age
 - partnership
 - religion
 - children
 - shopping habits
 - interests



Value of Databases (II)

- these databases can be used for
 - sampling for targeted research
 - analysis to identify key groups
 - test marketing

Accessing the Databases - Direct Marketing

- Is about Research as well as sales
- Hypothesis testing
- Detailed analysis
- Integration of information
 - Ongoing relationships
(suppliers/consumers/manufacturers)
 - Continuous improvement



Some Consumer (Lifestyle) Databases

Key Players:

- Claritas
- Consodata
- Experian
- Wegner Direct (Consumer Surveys)

Plus many other specialist databases

Information Available On Consumers (I)

- Sainsburys Customers * Home PC Users * Shareholders * Business Travelers * Clerical Workers * Boating Holiday makers * Housing Association Tenants * Joggers * Motorcyclists * Divorcees * Eczema Sufferers * Bungalow Dwellers * People with Mortgage Protection * Gamblers * Gordon's Gin Drinkers * Church Goers *

Information Available On Consumers (II)

- Self Employed * Houses with Coal Central Heating * Give to Human Rights Charities * Coin Collectors * Sega Mega Drive Players * People Considering Private Education * Vegetarians * Jazz Fans * High Income Households * Kays Direct Mail Customers * Nescafe Drinkers * National Trust Members * People considering holidays in South America

Information Available On Consumers (III)

- People with telephone bills over £100 per Quarter * Camcorder owners * FT Readers * People who watch over five hours TV per day * People with Halifax Accounts * Saab Drivers * Students * Three Child Households * Golfers * Gold Credit Card Holders * etc...

What are the problems?

- Consumer Databases arise from self-completion questionnaires, they therefore have certain biases
 - under represent young single people
 - fewer of the relatively poor households
 - less ethnic minorities
 - more (well off) retired persons
- All problems familiar to researchers!

What are the Advantages?

- Large numbers of Records (ranging from 1.25 million to over 5 million).
- Careful adjustments for the bias can give information of equal reliability to most comparable research.
- The exception is whenever a 'proper' random face to face sample is used.
- They both suffer from the inherent bias of the willing respondent!

International Business Databases

- **Dun & Bradstreet** (Over 250 countries “covered”)
- **Kompass** (58 countries)
- **Frankform Int’l** (24 Million “Key Executives”)
- **Harte Hanks** (IT only)
- **Yellow Pages** (Euro)

None of them standardised across countries, coverage varies enormously!

Sampling Definition

“Obtaining information about a universe from a portion of that universe.

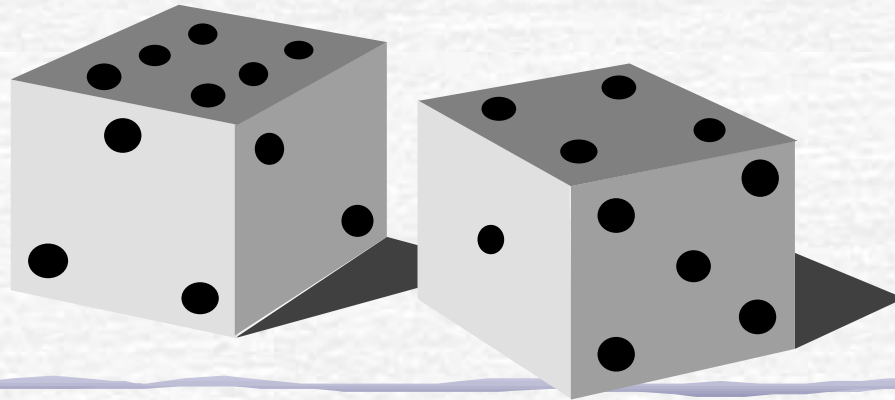
Samples provide estimates of the universe.”



Probability Samples

A probability sample is selected by a random procedure that gives every member of the population a known non-zero probability of selection.

Usually, for Business samples, the probability of selection will not be equal.





Non-Probability Samples

With targeting methods, there may be a loss of statistical projectability because not all elements of the universe have the opportunity to be selected, and the varying probabilities cannot be directly measured.

Therefore, use targeted samples when you do not need to measure a market with a known reliability.

Sampling Frames & Errors

- Sampling Frame: List of members of a population, which can be accessed with known probabilities of selection.
- Sampling Frame Error: Degree to which the frame fails to account for all elements of the population.
- Sampling Error: Refers to all error in a survey, which can occur because a sample is used.



Choosing a Sample Frame

A) The Research Need defines:

- Coverage of the Market
- Information Needs

B) Cost Efficiency requires

- Useful control information
- Data Quality

But: B) should be less important than A)



Research Objectives

- Business Opinion
- New Product Development
- Market Sizing

It is the objectives of the research which define the sample frame required and the database (or databases) which should be used.



Efficiency of A Sample

is a Combination of...

- Cost
- Time
- Reliability
- Value

Cost efficiency means..

- *Compromise* - statistical 'values' are distorted by:
 - sample frame (method) inaccuracy
 - response (rate) bias
 - recording error



So, how about my brief? (1):

- How do we approach surveys in niche markets?
 - By understanding the business need
 - Defining the context of the requirement in relation to the total market place
 - Identifying a suitable sample frame



So, how about my brief? (2):

- Data Sources and their collection
 - Varied
 - Horses for courses
 - Depends on choice of method!

So, how about my brief? (3):

- Does Targeted Sampling Improve efficiency?
 - *Yes, at least in terms of data collection costs and speed.*
- Can I still infer Findings Across the Population?
 - *When handled with care and understanding of the context, yes. But checks for bias should always be made (multivariate analysis).*
- What else can these resources offer?
 - *Many things, but I think others will have covered this!*

References (I)

- “On-line database solutions and the demands of BT’s customer satisfaction study”
 - (Bright D. & Simpson J. - IMRA Conference Proceedings 1992)
- “The impact of I.T. on M.R.: The last 10 years and the next 10 years”
 - (Roughton G. - JMRS Vol.. 34 No.2 April 1992)
- “Developing the ultimate target marketing machine”
 - (Oppenshaw Prof.. S. - AEMRI/ABMRC Conference 1992)



References (II)

- “From myopic maximisation to panoramic optimisation”
 - Clancy K. & Krieg K. - ESOMAR seminar 1990
- “Is the tail of market research wagging the dog”
 - van Herk J. - ESOMAR congress 1994