

Geographically Based Sampling For Household Surveys

Donna Hill

Jodie Thompson

John Henstridge

Data Analysis Australia

Context

- **Customer Relationship Management** requires particularly accurate survey information
 - Segmentation
 - Customer value
- **Utilities (gas, electricity and water) have particular needs**
- **Partially served in past by national surveys**
 - e.g. National Energy Survey 1985-86
 - Now have to be commercial
- **Surveys need “at least as good as ABS” quality**

Cluster Designs

→ Face-to-face interview economy

- Minimises travel
- Higher reliability
- Wider scope of questions

→ Diversity measurement for understanding demand peaks

- Water peak lowers head
- Electricity peak can cause failure

→ Modelling of penetration

- Binomial-logit models

The Need to be Different

→ Utilities now commercial

→ Demand fast turn-around

→ Go beyond factual data to attitudinal

→ Interest in geographical detail

→ Utilities have geographically distributed infrastructure

→ Outsourcing environment

→ Use of commercial field teams

Census Based Cluster Surveys

→ Sample CDs

→ Proportional to size

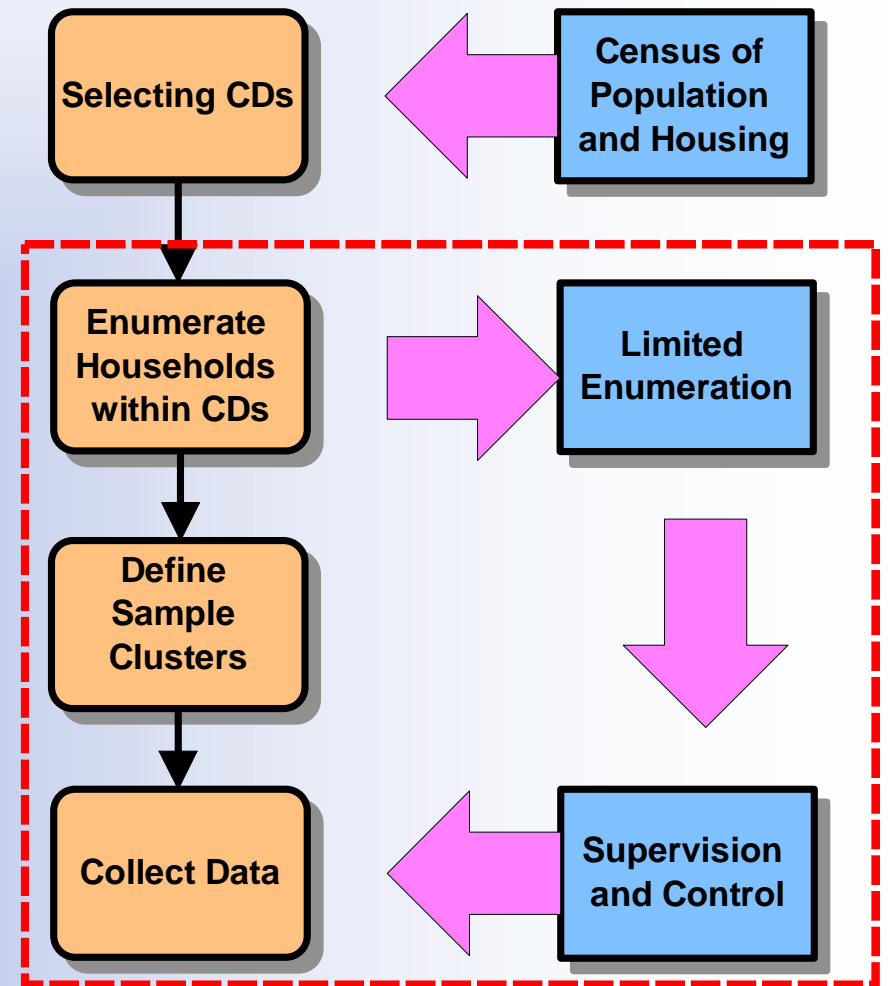
→ Enumerate households within CD

→ Sometimes omitted

→ Define clusters

→ Often not compact within CD

→ Collect data



Problems

→ Model does not work well with outsourced field teams

- Not experienced with enumeration techniques
- Works best in-house

→ Ignores information held by utilities

- Good enumeration available

→ Utilities have special interest in diversity

- Affects infrastructure design
- Compactness or physical size of cluster important

Geographically Based Alternative

→ Complete enumeration

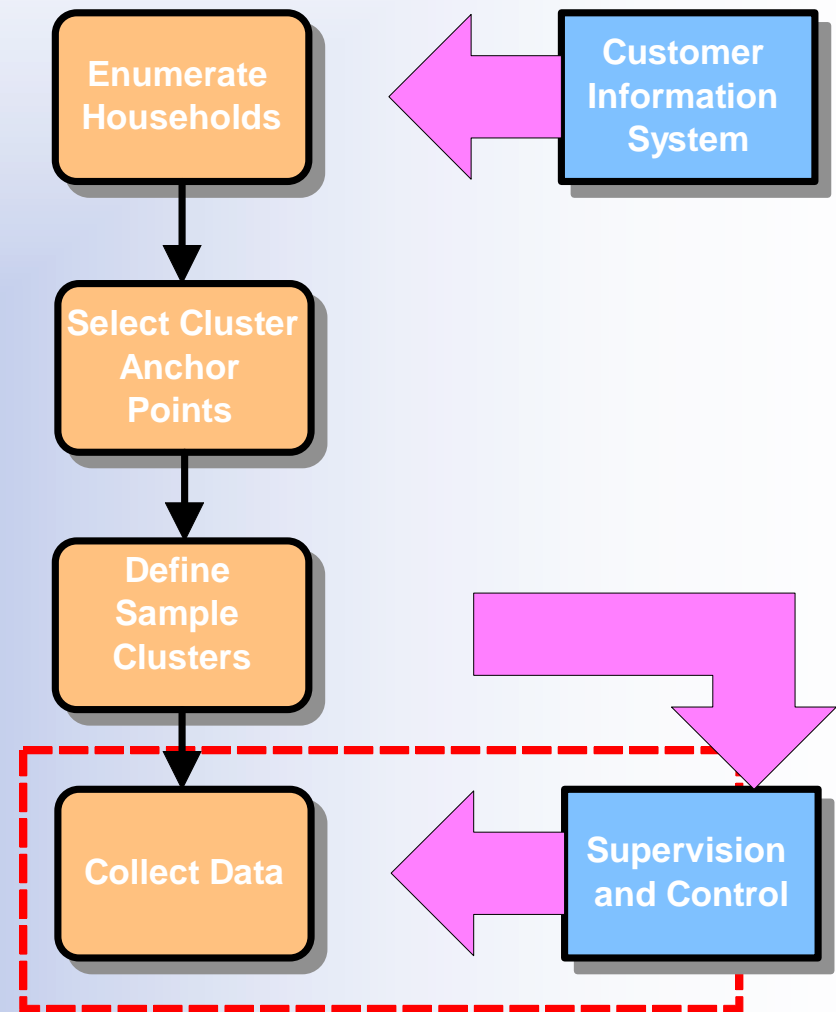
→ Customer Information System

→ Sample to get anchors

→ Define compact clusters around anchor

→ Requires geographical ordering

→ Collect data



Advantages

→ Most design can be done in-house

- Limited scope for fieldwork error
- Use geographical information in design
 - Geocoding of addresses
 - Sequencing of addresses (meter reading)
 - Census boundaries
- Minimises risk of changing fieldwork team

→ Cross linking with Customer Information System data

- Consumption

→ Up-to-date in rapidly changing areas

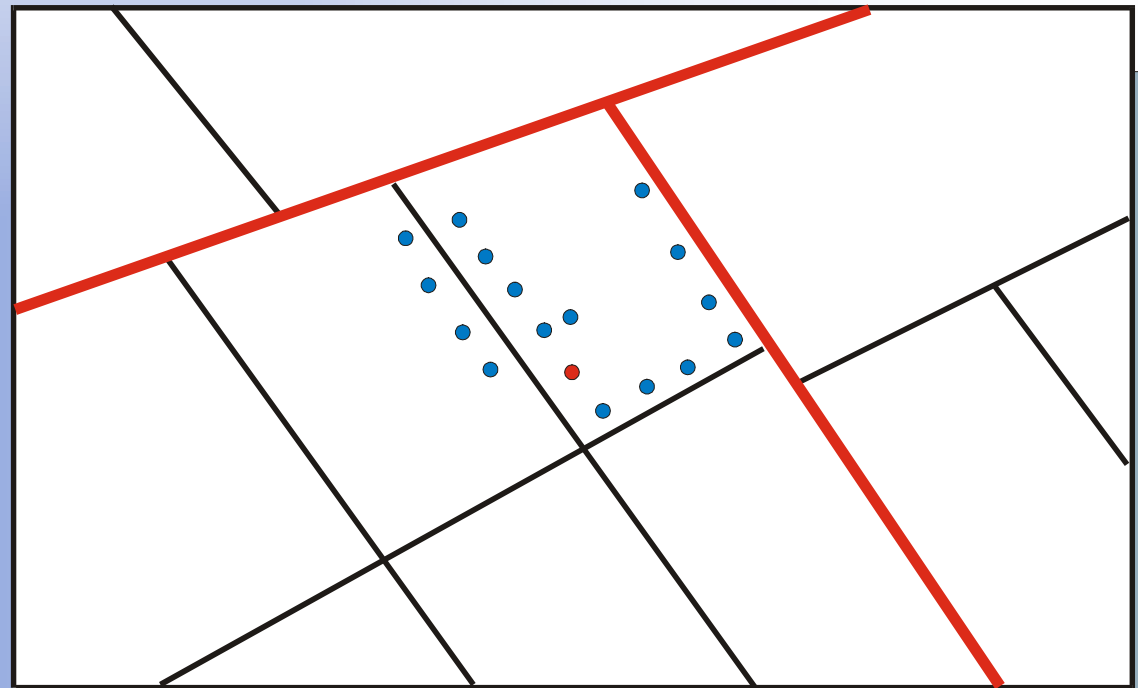
Cluster preparation

→ Checking validity of households

→ Customer types

→ Connectivity of cluster

→ GIS mapping



High Response Rates

→ Quality population frame

- Addresses checked to remove shared meters
- All addresses valid

→ Initial contact made by mail

- One week before

→ Demand high response in each cluster

- Typically 12 responses out of 17
- Follow up visits to all households where contact attempted

→ Cross checking

- Record meter numbers

Suburban Fringe Important

→ High demand

- Water - high usage, establishing gardens
- Energy - purchasing new appliances
- Distinct market segment

→ Census information not as accurate in these regions

- Changes rapidly date Census information

Conclusions

→ Highly successful methodology

- Applied to energy and water issues
- Reliable - consistent over several surveys

→ Successfully used for modelling

- Geographically based models of demand and appliance ownership