Challenges in Converting the National Crime Victimization Survey to Blaise

Charlie Carter
Roberto Picha
Technologies Management Office
Introduction

• The National Crime Victimization Survey is one of many household based surveys developed at the Census Bureau.

• The survey collects crime data for every household member who is age 12 and older.
Introduction

• When compared to other household based surveys administered at the Bureau...
  - CE 1 Respondent
  - AHS 1 Respondent
  - NHIS 3 Respondents
  - NCVS Every eligible household member

• Because of the number of eligible household members, whenever the interviewer exits and re-enters the instrument to complete someone’s crime report, we are required to retrieve any partial data already collected for that person.
Introduction

• NCVS has a sample size of about 56,000 households that are interviewed a total of 7 times during a 3 year period.
Introduction … Survey Challenges

• Challenge # 1
  - Developing an approach for collecting multiple crime reports for every eligible household member.

• Challenge # 2
  - Developing an approach for comparing all the reported crimes across all household members to ensure that there are no duplicates.
Introduction … Survey Challenges

- Challenge # 3
  - Integrating an Editing and Coding Instrument into the regular data collection instrument.
Was your credit card illegally used in the last 6 months?
Challenge #1: Collecting Multiple Crime Reports

- Screener questions asked about number of incidents within the reference period.
- Respondent can report up to 30 incidents for each of the 10 screener questions “crime categories” (such as personal theft, auto theft, assault, and so on).
- Detailed crime reports are then collected for all crimes mentioned in the Screener questions.
Collecting Multiple Crime Reports: Initial “Brute-Force” Approach

- Arrayed block for up to 30 Household members.
- Each person-level block contained an arrayed block for the 10 crime categories.
- Each crime category block contained an arrayed block for up to 30 crimes.
- 9,000 nested blocks (30 * 10 * 30).
Collecting Multiple Crime Reports: Results of “Brute-Force” design

- Exceeded 16,300 Blaise page limit – failed to compile;
- Reducing nested blocks
  - Load time of 45 seconds;
  - Navigation lag of 2 seconds;
  - Too many arrayed blocks on route;
Collecting Multiple Crime Reports: Solution

Data Storage Array

Data collected in one straight block

Data Collection Block

Data collection block is EMPTY, and questions are re-asked

Flags and counters are used to track and link the storage arrays to enable retrieval and processing

USCENUSBUREAU
Was any item taken from your pocket?
Challenge #2: Un-duplication

- Comparing crime reports that were collected during interview.
- Two kinds of Un-duplication
  - Un-duplication of current crime reports.
  - Un-duplication with previous crime reports.
Un-duplication Challenges

• Retrieving crime report data from a storage array.
  ❑ “Length of data for one block exceeds 65519 positions.”
• Formatting the data so reports are readable to the interviewer.
  ❑ Side-by-side comparison.
Un-duplication Process and Structure

- Compare the incident data below. Do not discuss incidents with the respondent.
- Is the incident on the left a duplicate of the incident on the right?

**INCIDENT REPORTED DURING CURRENT ENUMERATION**

**INCIDENT REPORTED DURING CURRENT ENUMERATION**

**Enum: 02 Ln No: 1 Inc. No: 2**

**Enum: 02 Ln No: 2 Inc. No: 4**

**MONTH: 2**
WHERE: In own dwelling/att garage.
Respondent and other HHLD member(s) present
WEAPON: Blunt object,
THEATEN: Weapon present,
NOTHING TAKEN AND NO ATTEMPT TO TAKE ANYTHI
SUMMARY: Offender tried to break into house
, was warned off, police called, offender fou
nd 2 days later, replaced door

**MONTH: 2**
WHERE: On street.
Respondent was present.
WEAPON: Knife,
ATTACK: Stabbed/cut w/ knife,
THEFT: Cash, Wallet, Credit cards,
SUMMARY: Rugged in parking lot of mall, cut o
n arm defending self, stitched at hospital, w
allet found hours later

Rows are all crimes reported by all respondents.
Columns are crimes reported by the current respondent.
Un-duplication Solutions

• Transferring crime data.
  ❑ Instead of parameters, create the report display at a higher block level.
    ▪ Did not exceed block limits.
    ▪ Able to re-use text fills.

• Displaying the data for comparison.
  ❑ Concatenated data and a string buffer up to 45 characters per line of each report.
Challenge #3: Editing & Coding

Editing and Coding (E&C) is the process of reviewing data that was collected during interviewing for quality assurance. Criteria which determined whether the E&C instrument would be called depended on...
1.- When at least one member had a race category equal “6- other” during the first time the case was interviewed . . .
2.- When at least one member had responded “Yes” to the called police crime or not called police crime questions regarding a potential crime.
Editing & Coding

3.- There were one or more crime incidents reported in the household at time of interviewing either [CATI] and/or [CAPI], screener information at the person level.

List of Crime Incident Reports for the household

... table that arrayed the maximum 30 crime report incidents that matched incidents to the person number and line number
Editing & Coding: Challenge

• Using NCVS as E&C presented serious issues for routing questions
• Populating E&C blocks with same data from the demographics and crime report.
  ❑ “Exceeded the record length size allowed of 32767.”
  ❑ Nested blocks
• “Load time was greater than 60 seconds.”
  ❑ When reducing amount of nested blocks.
Editing & Coding: Solution

- Updating E&C block (separate E&C block).
  - Screens to check race for HH and call police
  - Arrayed as Crime Report Storage data to include only those variables needed from reviewing data
    - Retrieval of Crime report for reviewing and editing was simpler
- Populate E&C component via Manipula script.
  - Most efficient mode to update E&C blocks without tying up resources
Conclusion: Lessons Learned

- Separate the data collection activities from the data storage when there are multiple respondents that are being asked the exact same series of questions.
- Develop the E&C instrument as a separate instrument instead of integrating it into the regular data collection instrument.
Dan reported missing...