RESPONDENTS AND NON RESPONDENTS TO POPULATION AND HOUSING CENSUS.
SOME STRATEGIES FOR DATA COLLECTION DESIGN IN THE ERA OF LOW RESPONSE RATE AND HIGH RESPONSE BURDEN.

Manuela Bussola, Novella Cecconi, Elena Donati, Linda Porciani
Istat | Direzione centrale per la raccolta dati
AIMS OF THE ANALYSIS

Finding some of the **reasons** behind the decrease of (CAWI) response rate through the description of the **profile of respondents / non respondents** to acquire elements for designing **adaptive** organizational and communication **strategies** to respondents

in order **to increase the response rate**, especially web response rate
A Decision tree model for the profile of respondents/non respondents

Why a decision tree model?

The method allows to have a meaningful classification through homogenous predictors starting from a population very ethereogenous, such as the PPHC sampled households.

How the decision tree model was applied?
1. PPHC sampled households was divided into three macro aggregation (sampled, respondents and non respondents)
2. Each subpopulation is the object of a specific decision tree model
3. The classification algorithm is CHi-squared Automatic Interaction Detection (CHAID)

CHAID is a multiple segmentation techniques based on $\chi^2$ test.
A contingency table ($\chi^2$ and $p$-value) was calculated for each explicative variable (by modalities) and dependent variable. The comparison between the minimum $p$-value ($p$-min) and the stop value ($\alpha$, that is the maximum dimension of the tree or the maximum number of levels or the minimum frequency in a node) of a specific attribute $X_i$ could be:
- $p_{min} < \alpha$; $X_i$ is an included modality;
- $p_{min} > \alpha$; $X_i$ is an attribute of a leaf.
A Decision tree model for the profile of respondents/non respondents

**Sampled households**
(n. 939.588)

**Respondents households**
(n. 855.295)

**Non Respondents households**
(n. 84.293)

* Excluding off-target households (death households, moving households, etc.)
A Decision tree model for the profile of respondents/non respondents

Other cross variables: Internet use | Institutions trustness | efficiency of postal solicits
The tree model of sampled households

The most significant variable

Household by citizenship

Respondents (%)

Yes 91%
No 9%
The tree model of respondents

The most significant variable

Education level

Data collection techniques

CAWI 48.7%
CATI 9.6%
CAPI 41.7%
The **tree model** of non respondents

**THE MOST SIGNIFICANT VARIABLE**

**INNER AREAS**

- BELT + INTERM.
- SINGLE
- MULTI
- REMOTE + ULTRAREMOTE

**Contacted by interviewer**

- YES 23.3%
- NO 66.7%
**Adaptive respondents/non respondents strategies** in the survey process

<table>
<thead>
<tr>
<th>Profiles</th>
<th>Possible Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non Respondents</strong></td>
<td>Improving the collaboration with local foreigner associations</td>
</tr>
<tr>
<td>• 1 member Non Italian</td>
<td>Starting the data collection from 1 member households and non Italian households</td>
</tr>
<tr>
<td>• Living in Single-Municipality Service Center</td>
<td>Reducing the field workload through the improvement of self-enumeration</td>
</tr>
<tr>
<td><strong>Respondents</strong></td>
<td></td>
</tr>
<tr>
<td>• Italian households</td>
<td></td>
</tr>
<tr>
<td>• Low field workload</td>
<td></td>
</tr>
<tr>
<td><strong>Interviewer</strong></td>
<td>Improving the communication campaign in remote areas focused to the self enumeration</td>
</tr>
<tr>
<td>• Low educational level</td>
<td>Improving an additional technique, as smart questionnaire (i.e. through a QR Code in the informative letter)</td>
</tr>
<tr>
<td>• Living in remote areas</td>
<td></td>
</tr>
<tr>
<td><strong>Self Enumeration</strong></td>
<td></td>
</tr>
<tr>
<td>• High educational level</td>
<td></td>
</tr>
<tr>
<td>• Living in Single-Municipality Service Center</td>
<td></td>
</tr>
<tr>
<td>• High professional status</td>
<td></td>
</tr>
<tr>
<td><strong>Not Contacted</strong></td>
<td></td>
</tr>
<tr>
<td>• Living in Single and Multi Single-Municipality Service Center</td>
<td></td>
</tr>
<tr>
<td>• High field workload</td>
<td></td>
</tr>
<tr>
<td><strong>Contacted</strong></td>
<td>Reducing the field workload through the improvement of self-Enumeration</td>
</tr>
<tr>
<td>• Living in Belt areas</td>
<td>Making available the mobile phone numbers</td>
</tr>
<tr>
<td>• Low field workload</td>
<td>Involving municipality census offices to customize the communication campaign at the local level</td>
</tr>
</tbody>
</table>
thanks

MANUELA BUSSOLA | bussola@istat.it
NOVELLA CECCONI | ncecconi@istat.it
ELENA DONATI | eldonati@istat.it
LINDA PORCIANI | porciani@istat.it