

Experiences with Blaise for Data Editing in the Housing Rental Survey

Harm Melief and Mark Kroon, Statistics Netherlands

Abstract:

The housing rental survey (huurenquête) is conducted by Statistics Netherlands (CBS) and (partly) commissioned by the Dutch Government. In this survey information is collected on rents, service costs and characteristics for some 14000 houses from 1500 rental agencies. The data are collected using (paper) questionnaires or directly from the administration of a number of large agencies. These datasets need to be combined and corrected for errors and omissions. The cleaned micro-data is delivered to the Government and used to calculate the housing rental index.

We have built a system in the Blaise Suite which supports the combining and correction of the data and which calculates the index. Using Manipula we combine the data into a SQL-server database. The records are automatically scanned for errors using a Blaise data model with approximately 150 rules. Corrections are made by calling the respondents and interactively correcting the micro-data using the Data Entry Program (DEP). The index is calculated using Manipula. The entire process is supported by a Maniplus menu and a number of dialogboxes.

In this project we have build a complex data processing tool, which uses Blaise in an unusual manner. This has created a number of unforeseen problems and has revealed some of the limitations of the current package. For instance joining large sets of data, management of classifications and support-data, database interactions and version control, all required either novel applications of the Suite or a melding of Blaise with other tools. In this paper we will describe a number of these problems and solutions. This will help define the scope of the applicability of Blaise for complex systems.