

Regarding the construction of a variable's name, the first character always indicates to which file each variable belongs: **D, H, R** or **P**.

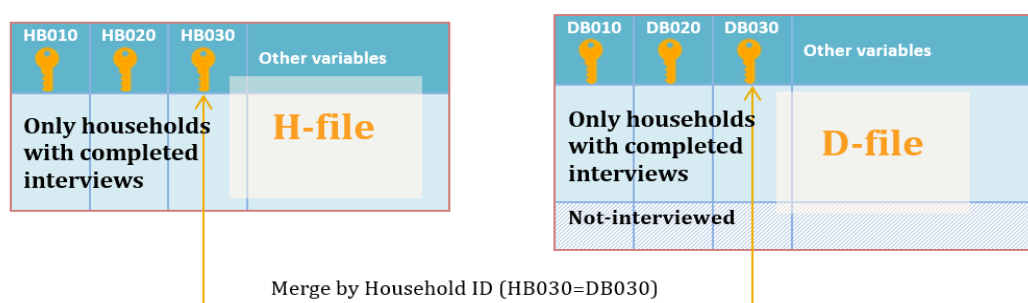
The **Household register file (D-file)** contains every selected household, including those where the address could not be found and also households which could not be interviewed, i.e. all households from the sample (selected + substituted + split off). The topics covered relate, among other things, to the degree of urbanisation, the household questionnaire result and interview acceptance.

By contrast, the **Household data file (H-file)** contains only households which have been contacted, have completed the interview and where at least one member aged 16 or over has completed the personal interview. The topics covered relate, among other things, to total household income (gross and disposable), gross income components at household level, housing and non-housing related areas, non-monetary household deprivation indicators, physical and social environment, dwelling type, tenure status, housing conditions and costs.

Figure 4.2-1 shows how the D-file and H-file should be merged.

H-file \subseteq D-file (case when H-file = D-file is when the response rate = 100%)

Figure 4.2-1: Merging Household data (H-file) with Household register (D-file)



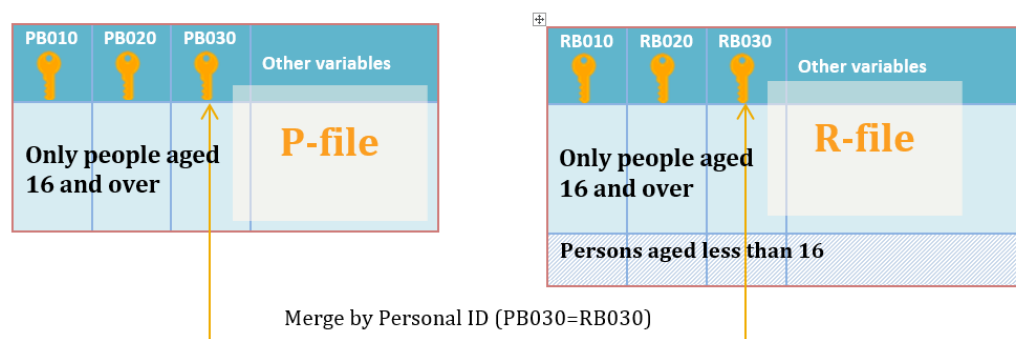
DB010, DB020 and DB030 are used as key variables of part D and HB010, HB020 and HB030 are key variables of part H. These variables create the possibility of merging different files.

Personal register (R-file) contains all the people currently living in the households and those who are temporarily absent. In the case of longitudinal data, it also contains all persons registered in the R-file of the previous year, even if they no longer live in the household (RB110 in 5,6). The information collected is mostly related to basic and demographic data, and childcare.

Personal data (P-file) contains every eligible person (persons aged 16 and over — RB245 = 1, 2 or 3) for whom the information could be collected from interview and/or registers (RB250 = 11, 12, 13 or 14). The topics covered relate, among other things, to education, labour, health and gross personal income (total and components at personal level).

- Figure 4.2-2 shows how the P-file, and the R-file should be merged. P-file \subseteq R-file (P-file = R-file is when the observed sample does not include persons aged less than 16 and all persons respond to the personal interview). So, every person in part P should be listed in part R.

Figure 4.2-2: Merging Personal data (P-file) with Personal register (R-file)



ID numbers have a major role while merging and using the data. However, their correct construction is even more important.

Based on the sample, each household will receive a '**Household number**'. This number is the basis for constructing the Household ID and the Personal ID. It may or may not be a sequential number, but it is unique and should not contain other information. This means that it must not contain any information that conflicts

with confidentiality rules. Also, this number must be unique for all the years of the survey:

- Format of the Household number: 1- 99999999 (maximum eight digits). Normally it is constructed using information from the sample and the fieldwork.

The ID should not contain any confidential information. It should be a source of different data information:

- Household ID = Household number + split number (two-digits)
- Personal ID = Household ID + personal number (two digits)

Split number is related to the split-off households (split could occur from wave two onwards). Here are some main rules when defining the split number:

- The split number for the first wave will always take the value '00';
- If the household remains entire, it will keep the Household number and Split number from one wave to the next;
- In the case of a split-off, the initial household will keep the Household number and Split number from one wave to the next. The other household(s), i.e., the split-off households, will keep the same Household number, but will be assigned the next available unique Split number in sequence;
- In case two sample households merge, if the new household remains at a previous address, it retains the Household number and Split number of the household that was at that address in the previous wave;
- If the new household is at a new address, the Household number and Split number of the household of the sample person who now has the lowest person number in the 'household register' will be retained.

Personal number refers to the number assigned to the person who is recorded as a household member for the first time. For every new person in the household '1' is added to the highest-used person's number (for all the years of the survey and the Household ID). For further information, please see [Figure 4.2-3](#).

Figure 4.2-3: Figure Creation of Household and Personal Number if it is from wave to wave

