DB010: YEAR OF THE SURVEY

Topic and detailed topic: Technical items / Data collection information
Variable type: Annual
Unit: Household
Reference period: Current
Mode of collection: Frame
In use (period): Yes, since the first year of EU-SILC data collection
Series' differences: No changes

VALUES AND FORMAT

Format: Four-digit number, no decimals

FLAGS

-

DESCRIPTION

The year in which the survey-data collection, or most of the collection, is carried out.

DB020: COUNTRY OF RESIDENCE

Topic and detailed topic: Technical items/Data collection information
Variable type: Core variable/Annual
Unit: Household
Reference period: Constant
Mode of collection: Frame
In use (period): Yes, since the first year of EU-SILC data collection
Series' differences: Yes. Before 2012, Greece was abbreviated as GR.
Montenegro, North Macedonia, Serbia, Albania and Kosovo added (North Macedonia in 2010, Montenegro and Serbia in 2013, Albania in 2017, Kosovo in 2018 and Bosnia and Herzegovina in 2022)

VALUES AND FORMAT

SCL GEO code

BE	Belgium	NL	Netherlands
BG	Bulgaria	AT	Austria
CZ	Czechia	PL	Poland
DK	Denmark	PT	Portugal
DE	Germany	RO	Romania
EE	Estonia	SI	Slovenia
IE	Ireland	SK	Slovakia
EL	Greece	FI	Finland
ES	Spain	SE	Sweden
FR	France	UK	United Kingdom
HR	Croatia	CH	Switzerland
IT	Italy	IS	Iceland
CY	Cyprus	NO	Norway
LV	Latvia	ME	Montenegro
LT	Lithuania	MK	North Macedonia
LU	Luxembourg	RS	Serbia
HU	Hungary	TR	Türkiye
MT	Malta	AL	Albania
		XK	Kosovo
		BA	Bosnia and Herzegovina

FLAGS

-

DESCRIPTION

The country of residence is the country in which the usual residence of the person/household is located. The list of countries is defined according to the Eurostat Standard Code list (SCL) GEO which is largely based on the International Organization for Standardization (ISO) 3166 country codes (3166-1 alpha-2).

The SCL GEO is available here:
https://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=CL_GEO&StrLanguageCode=EN&IntPcKey=&StrLayoutCode=HIERARCHIC.

Country code, year and IDs are used as key variables to merge different files.

DB030: HOUSEHOLD ID

Topic and detailed topic: Technical items / Identification
Variable type: Annual
Unit: Household
Reference period: Current
Mode of collection: Frame or interviewer
In use (period): Yes, since the first year of EU-SILC data collection
Series' differences: ---

VALUES AND FORMAT

ID number

Household number 1 - 9999999 (maximum seven digits)

Household ID = Household number + split number (two digits)

FLAGS

-

DESCRIPTION

Household ID (maximum nine digits) = Household number (maximum seven digits) + split number (two digits)

Every household will receive a household number. This number is the base upon which to construct the household ID and the personal ID. It should be a sequential number and should not contain other information. It must NOT contain any information that conflicts with confidentiality rules. This number must be unique for all the years of the survey.

The household identification number (ID) is composed of the household number and the split number. The household number is a unique number assigned to each household in wave 1. The split number for the first wave must always take the value '00'.

When a household stays in the sample, it keeps 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 households, 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 cases when there is more than one split-off, the split number is formed by adding 1 to the highest used split number (for all the years of the survey).

When there is a fusion of two sample households and the new household is located at an address where one of the households was previously interviewed the new household should keep the household number and split number of the household at that address in the previous wave.

When the new household is located 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' should be used.

DB040: REGION OF RESIDENCE

Topic and detailed topic: Technical items/Data collection information
Variable type: Core variable/Annual
Unit: Household
Reference period: Current
Mode of collection: Frame, register or interviewer
In use (period): Yes, since the first year of EU-SILC data collection
Series' differences: Yes, in 2012 and 2015, 2021 (changed the label).

VALUES AND FORMAT

NUTS 2 digits

FLAGS

year >=2021
1 filled according NUTS
-1 missing
year >=2015
1 filled according NUTS-13
-1 missing
year>= 2012 and <2015
1 filled according to NUTS-10
-1 missing
year<2012
1 filled according to NUTS-03
2 filled according to NUTS-08
-1 missing

DESCRIPTION

The region of residence is the region within the country of residence in which the usual residence of the household is located at the date of interview.

From 2021 onwards

Under Regulation (EC) No 1059/2003 of the European Parliament and of the Council all Member States' statistics that are submitted to the Commission (Eurostat) and that are to be broken down by territorial units should use the NUTS classification. Consequently, in order to establish comparable regional statistics, data on the territorial units should be provided in accordance with the NUTS classification.

NUTS (2-digits)

For Member States of the EU, regions are defined on the basis of the Nomenclature of territorial units for statistics (NUTS) as described in Article 3 of Regulation (EC) No 1059/2003 of the European Parliament and the Council on the establishment of a common classification of territorial units for statistics (NUTS) (as amended)[22] and Annex I.

For EFTA and Candidate countries, region refers to the classification of Statistical Regions[23].

Information about the NUTS is available here: http://ec.europa.eu/eurostat/web/nuts/overview.

[22] http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32003R1059

[23] http://ec.europa.eu/eurostat/web/nuts/statistical-regions-outside-eu

DB050: STRATUM

Topic and detailed topic: Technical items/Data collection information

Variable type: Annual

Unit: Household

Reference period: At selection

Mode of collection: Frame, register or sample design

In use (period): Yes, since 2008

Series' differences: Yes (2014) (description and flags applied changes from 2014), the variable label slightly changed in 2021 based on standardised variable names

VALUES AND FORMAT

1 – 99999 Stratum identifier

FLAGS

From 2014 onwards

1 Filled
2 Self-representing primary sampling unit (PSU)
3 Collapsed stratum due to single PSU (only for households selected in the stratum with the single PSU)
-2 Not applicable (no stratification)

Before 2014

1 Filled
-2 Not applicable (no stratification)

DESCRIPTION

DB050 reports on the primary stratum corresponding to each observation unit (individual or household) in case the target population (or a part thereof) is stratified at the first stage of the sample design, providing identification codes for the different strata. The information on the variable should be filled in for all waves, and it should always refer to the situation at the time the concerned unit (individual or household) is selected; it consequently does not refer to the strata used for post-stratification. DB050 collects information about primary strata, the strata used in selecting the sample. The variable label is changed from 'Primary strata' *[Primary strata as used in the selection of the sample]*' used before, to 'Stratum' used from 2021 onwards, harmonised with social standardised variables.

Stratifying a population means dividing it into non-overlapping subpopulations called strata. Independent samples are then selected within each stratum.
The variable contains sampling information that is needed to calculate measures of spread (e.g., variance and standard deviation) and to observe changes over time.
The category 'stratum identifier' provides the identification code of the stratum each observation unit (individual or household) belongs to. Stratum identification codes are to be used in case the target population has been stratified, or in case self-representing primary sampling units (PSUs) have been considered[24].

The category 'not applicable' is to be used in case the target population has not been stratified at the first stage of the sampling process (e.g., when the sample has been drawn by simple random sampling or by cluster random sampling), and self-representing PSUs have not been considered.

Information on the variable should be filled in for all waves (and/or panels), and it should always refer to the situation at the time the concerned unit (individual or household) is selected.

In case the target population is stratified at the first stage of the sampling design (or in case of comparable sample design), all primary strata receive a unique identification code which remains the same for the entire period in which the observation units (individual or household) that receive this identification code remain in

[24] Self-representing PSUs are PSUs selected with a probability of 1, which must be considered to be a stratum rather than a PSU.

the survey. The stratum identification code should be consistent over time for each observation unit (individual or household).

The information in the variable 'stratum' should enable all strata to be identified. Combining 'stratum' with other variables (such as 'region') should not be needed to identify the strata; 'region' should refer to the moment of the interview, while 'stratum' should refer to the moment of the selection.

In cases of variations of territorial units, such as the blending of municipalities, the value of 'stratum' should not change over time; it should always refer to the situation at the time of the selection.

To estimate the measures of spread (e.g., variance and standard deviation), each self-representing primary PSU must be considered to be a stratum rather than a PSU. Therefore, each self-representing PSU receives a separate and unique stratum identification code which remains the same for the entire period in which the observation units (individuals or households) that receive this identification code remain in the survey.

If strata in the sample consist of only one PSU selected among a larger number of PSUs in the stratum population, or if strata in the sample contain only one PSU (among a larger number of selected PSUs) with respondents, the primary strata have to be collapsed so that every stratum consists of at least two PSUs. If done, this grouping will ideally be implemented between strata that are most similar in terms of the variables used for stratification.

The variable PSU should be left blank when a sample of individuals has been drawn and the population has not been clustered.

It is good practice to randomise the stratum identifiers to prevent respondents from being indirectly identified (it makes the linking to particular locations or to a geographical pattern impossible). However, the randomisation process should ensure that the value of 'stratum' for a respondent does not change over time.

The stratum code of the collapsed stratum is equal to the stratum code of the stratum that before collapsing already contained more than one PSU. The households selected in the stratum with the single PSU receive flag code '3'[25].

[25] According to the characteristics and availability of data for different countries, we have used different variables to specify strata and cluster information. In particular, countries have been split into three groups:

1) BE, BG, CZ, IE, EL, ES, FR, IT, LV, HU, NL, PL, PT, RO, SI, UK and HR whose sampling design could be assimilated to a two-stage stratified type using DB050 (primary strata) for strata specification and DB060 (primary sampling unit) for cluster specification;

2) DK, DE, EE, CY, LT, LU, AT, SK, FI and CH whose sampling design could be assimilated to a one-stage stratified type using DB050 for strata specification and DB030 (household ID) for cluster specification;

3) MT, SE, IS and NO, whose sampling design could be assimilated to a simple random sampling, using DB030 for cluster specification and no strata.

DB060: PRIMARY SAMPLING UNITS [PSU]

Topic and detailed topic: Technical items/Data collection information
Variable type: Annual
Unit: Household
Reference period: At selection
Mode of collection: Frame, register or sample design
In use (period): Yes, since the first year of EU-SILC data collection
Series' differences: Yes (2014)

VALUES AND FORMAT

1 – 99999

FLAGS

From 2014 onwards

1 Rotation is implemented at primary sampling unit (PSU) level (the PSU rotates in and out of the sample)
2 Rotation is implemented at secondary sampling unit (SSU) or household level (The PSU remains in the sample for the entire duration of EU-SILC)
-2 Not applicable (no first or second sampling stage)

Before 2014

1 Filled
-2 Not applicable (no first or second sampling stage)

DESCRIPTION

DB060 measures PSUs as used in selecting the sample and it is part of a standardised list of variables.

If direct-element sampling is either impossible (lack of a sampling frame) or too expensive to implement (the population is widely distributed geographically), multi-stage selections can be done. Firstly, the population is divided into disjoint sub-populations, called **primary sampling units (PSUs)**. A sample of PSUs is then selected (first–stage sampling). Secondly, each sampled PSU is itself divided into disjoint sub-populations, called **secondary sampling units (SSUs).** SSUs are then independently drawn from each PSU (second-stage sampling) and so on.

The variable reports on the PSU corresponding to each observation unit (individual or household) in case the target population is divided into clusters, providing identification codes for the clusters or PSUs. The information recorded always refers to the situation at the time the concerned unit is selected (individual or household).
A population is divided into clusters (i.e., disjoint sub-populations) in case direct-element sampling is either impossible (due to lack of a sampling frame) or too expensive to implement (the population is widely distributed geographically). A sample of clusters (PSUs) is then selected at the first stage of the sampling process (or an alternative multi-stage selection process).
The variable contains sampling information that is needed to calculate measures of spread (e.g. variance or standard deviation) and to observe changes over time.
The category 'primary sampling unit identifier' provides the identification code of the selected PSU each observation unit (individual or household) belongs to, in case the target population has been divided into clusters in the first stage of sampling.

The category 'not applicable' is to be used when the target population has not been clustered at the first stage of the sampling process, e.g., when the sample has been drawn by simple random sampling or by stratified random sampling.
Information on the variable should be filled for all waves (and/or panels), and it should always refer to the situation at the time the concerned unit (individual or household) is selected.

In case the target population is clustered at the first stage of the sampling design, the selected clusters (PSUs) receive a unique identification code which remains the same for the entire period in which the observation

units (individual or household) that received this identification code remain in the survey. The PSU identification code should be consistent over time for each observation unit (individual or household).

When sampling with replacement is used and the same PSU is selected several times (‘multiple hits’) the PSU receives a unique identification code for every hit.
The variable PSU should be left blank when a sample of individuals has been drawn and the population has not been clustered.

In the situation where dwellings are selected at the first stage of sampling and more than one household shares the same dwelling, dwellings must be regarded as clusters of households and then coded accordingly. Then, if the first[26] stage of the sampling design consists of a selection of dwellings, each dwelling receives a unique code for category 'primary sampling unit identifier' that remains the same for the entire period the households in the considered dwelling remain in the survey sample.

If the first[27] stage of the sample design consists of a selection of households and the final observation unit is the individual, each household receives a unique code for the category 'primary sampling unit identifier' that remains the same for the entire period the household remains in the survey. Split-off households keep their original PSU identifier at the moment they are selected.

It is a good practice to randomise the PSU identifiers to prevent respondents from being indirectly identified (it makes the linking to particular locations or geographical pattern impossible). However, the randomisation process should ensure that the value of 'PSU' for a respondent does not change over time.

[26] This does not necessarily imply that there is a second stage of sampling.

[27] This does not necessarily imply that there is a second stage of sampling.

DB062: SECONDARY SAMPLING UNITS [SSU]

Topic and detailed topic: Technical items/Data collection information

Variable type: Annual

Unit: Household

Reference period: At selection

Mode of collection: Frame, register or sample design

In use (period): Yes, since the first year of EU-SILC data collection

Series' differences: Yes (2014)

VALUES AND FORMAT

1 – 99999

FLAGS

From 2014 onwards

1 Rotation is implemented at primary sampling unit (PSU) level (the PSU rotates in and out of the sample)
2 Rotation is implemented at secondary sampling unit (SSU) or household level (The PSU remains in the sample for the entire duration of EU-SILC)
-2 Not applicable (no first or second sampling stage)

Before 2014

1 Filled
-2 Not applicable (no first or second sampling stage)

DESCRIPTION

If direct-element sampling is either impossible (lack of sampling frame) or too expensive to implement (the population is widely distributed geographically), multi-stage selections can be done. Firstly, the population is divided into disjoint sub-populations, called **primary sampling units (PSUs)**. A sample of PSUs is then selected (first–stage sampling). Secondly, each sampled PSU is divided itself into disjoint sub-populations, called **secondary sampling units (SSUs).** SSUs are then independently drawn from each PSU (second-stage sampling) and so on, see DB060. DB062 provides identification codes for the selected PSUs (SSUs) as used in selecting the sample, secondary sampling units. In the case that the same SSU is selected several times ('multiple hits'), the SSU receives a unique value for every hit.

If the first stage of the sample design consists of a selection of households, households receive a unique code for variable DB060 that remains the same throughout EU-SILC's entire duration. In the latter case split-off households keep their original value at the moment they are selected for variable DB060. In case there is at least a third stage of selection, additional variables DB06$_i$ must be transmitted as identification numbers for the units sampled at stage '*i*'.

In the case of self-representing PSUs (for a definition see variable DB050), SSUs should be treated as if they were PSUs and receive a unique code for variable DB060. If households are selected at the second stage, they receive a unique value for variable DB060 that remains the same throughout EU-SILC's entire duration. In the latter case split-off households keep their original value at the moment of selection for variable DB060. The identification of the self-representing units themselves is implemented in variable DB050.

DB070: ORDER OF SELECTION OF PSU

Topic and detailed topic: Technical items/Data collection information

Variable type: Annual

Unit: Household

Reference period: At selection

Mode of collection: Frame, register or sample design

In use (period): Yes, since the first year of EU-SILC data collection

Series' differences: Yes (2014)

VALUES AND FORMAT

1 – 99999

FLAGS

From 2021 (revised flags)

11	Order on sampling frame is fixed for all EU-SILC survey years and primary sampling units (PSUs) have an equal probability of selection (within explicit strata)
12	Order on sampling frame is fixed for all EU-SILC survey years and PSUs have an unequal probability of selection (within explicit strata)
21	Order on sampling frame may change over time and PSUs have an equal probability of selection (within explicit strata)
22	Order on sampling frame may change over time and PSUs have an unequal probability of selection (within explicit strata)
-2	Not applicable (no systematic selection)

From 2014-2020

-2 Not applicable (no systematic selection)

Or a combination of two digits:

First digit: fixed or changing order of selection

1 order on sampling frame is fixed for all EU-SILC survey years

2 order on sampling frame may change over time

Second digit: probability of selection of PSUs

1 PSUs have an equal probability of selection (within explicit strata)

2 PSUs have an unequal probability of selection (within explicit strata)

e.g. the order of PSUs on the sampling frame remains fixed for the entire duration of EU-SILC and PSUs are selected with a probability equal to their size: the flag is equal to 12

Before 2014

1 filled

-2 not applicable (no systematic selection)

DESCRIPTION

If primary sampling units (or households in case of direct-element sampling) are selected systematically, DB070 contains the rank of selection of those units, order of selection of PSU, and order of selection of PSU as used in the selection of the sample. If PSUs rotate in and out of the sample, this rank should correspond to the rank on the sampling frame, such that PSUs newly selected in the sample could be grouped together on the basis of the order of all PSUs on the sampling frame. The value for DB070 of every selected PSU remains the same for the entire duration of EU-SILC. This information is important for variance estimation purposes because a systematic drawing from a judiciously ordered sampling frame may substantially reduce sampling errors.

If systematic selections have been performed at other sampling stages, additional variables DB070 (i-1), that is the order of the selection of the units of stage *'i'* (i>1), must be transmitted too.

In order to facilitate the computation of the standard errors for i) the common EU indicators, ii) the equivalised disposable income, iii) the unadjusted gender pay gap and iv) a list of income components, countries should[28] fill in this (these) variable(s) (in the case of systematic selection) for all panels and waves in the file, and not

[28] Agreement during the Living Conditions Working Group meeting in June 2009

only the first one of the sub-sample (being the year of the selection of the concerned household). However, the recorded information, always refers to the situation at the time the concerned household is selected.

The above definition also applies to the new entries from the second wave onwards.

DB075: ROTATION GROUP

Topic and detailed topic: Technical items/Data collection information
Variable type: Annual
Unit: Household
Reference period: Constant
Mode of collection: Derived
In use (period): Yes, since the first year of EU-SILC data collection
Series' differences: No changes

VALUES AND FORMAT

1 – 9

FLAGS

1	Filled
-2	Not applicable (no rotational design used)

DESCRIPTION

This variable must be filled only for the countries using a rotational design.

Rotational design

Refers to any sample selection which is based on a fixed number of sub-samples, called replications, each one representative of the target population at the time of their selection. Each year, one sub-sample rotates out and a new one is drawn as a substitute.

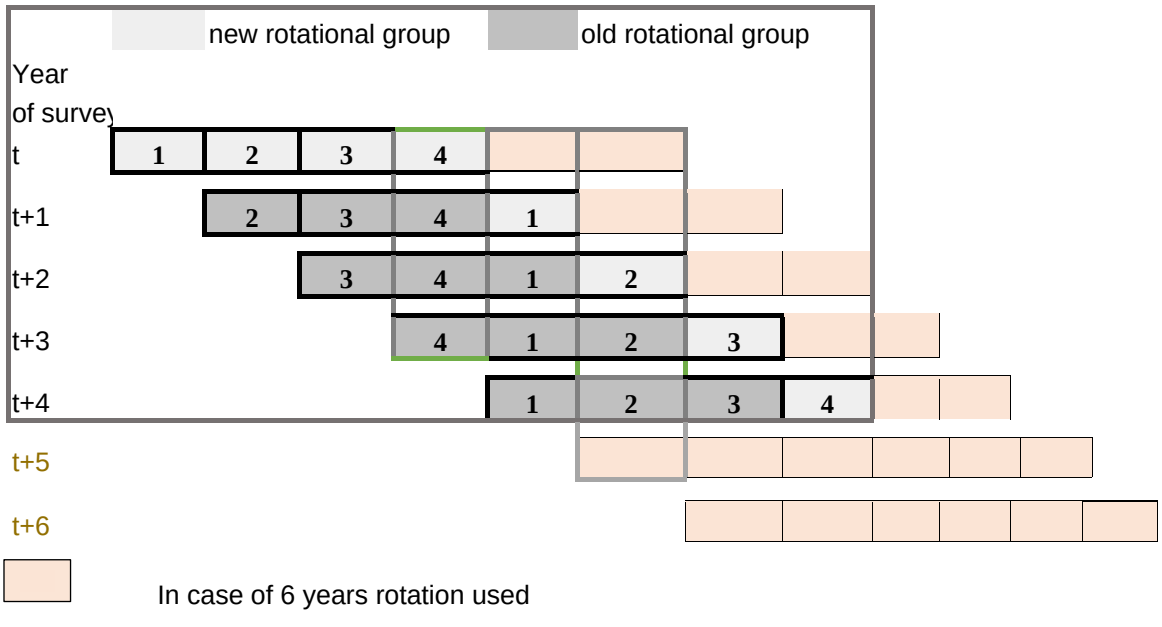
In the case of a rotational design based on four replications with a rotation of one replication per year, one of the replications must be dropped immediately after the first year, the second must be retained for 2 years, the third for 3 years, and the fourth for 4 years. From the second year onwards, at the start of each new year one replication must be introduced and retained for 4 years.

Rotation group

Each replication is called a rotational group and the information on the group to which the household belongs is especially useful for controlling the implementation of the sample over time.

Regarding the numbering of the rotation groups over time, it is recommended that each rotation group keeps the same number throughout the period of the survey (see figure hereafter):

PATTERN FOR ENUMERATION OF ROTATIONAL GROUPS



DB076: INTERVIEW WAVE

Topic and detailed topic: Technical items/Data collection information
Variable type: Annual
Unit: Household
Reference period: Current
Mode of collection: Derived
In use (period): New, from 2021
Series' differences: No changes

VALUES AND FORMAT

1 - 9 Wave number

FLAGS

1	Filled
-7	Not applicable (DB010 < 2021)

DESCRIPTION

The wave corresponds to the number of years the respective household stays in the survey. If the household is in the data sample for the first time, the value 1 should be selected. If the household is in the sample for the second time, the value 2 should be selected, value 3 for the third time, value 4 for the fourth time, etc. The same approach should be used when using a longer panel.

A split-off household keeps the same wave as the initial household.

DB080: HOUSEHOLD DESIGN WEIGHT

Topic and detailed topic: Technical items/Weights
Variable type: Annual
Unit: Household
Reference period: Current
Mode of collection: Derived
In use (period): Yes, since 2008 (except 2014)
Series' differences: Yes, 2014

VALUES AND FORMAT

0 (format 2,5)	Weight
Required format	Household cross-sectional weights will be coded with at least one integer and five decimals.

FLAGS

1 Filled
-2 Not applicable (not in the first year in the survey or split off)

DESCRIPTION

Weighting for the first year of each sub-sample

Design weights (household weights DB080 and 'Selected respondent' weights PB070)

The design weights need to be defined for all selected units, not only for responding units. The household design weights aim to draw inference on the household population from the household sample. The design weight of a household (h) is the inverse of its inclusion probability. DB080 is computed as follows:

In the case that households are sampled (or addresses or other units containing households):

$$DB080_h = 1 \text{ / (probability of selection of h)}$$

In the case that persons are sampled:

$$DB080_h = 1 \text{ / } \Sigma \text{ (probabilities of selection of eligible persons in h)}$$

'Eligible persons' are people who are given a non-zero probability in the selection procedure, such as people aged 16+. When the probability of selection is the same for all eligible people in each household, the denominator is simply the number of such people in the household multiplied by the probability of selection.

From the second year onwards (case of a rotational panel)

For the households which are not interviewed for the first time, no values are to be given for the design weight and these are to be flagged '-2' (not applicable).

In the weight section (8), a more extensive explanation on the weighting procedures is given.

Further adjustments taking into consideration non-response or calibration are done DB080 (N).

The design weights have to be inflated by the inverse of the response propensities in order to compensate for the loss of units in the sample. A classical procedure consists of modifying the design weights by a factor inversely proportional to the response rate within each 'homogeneous group', in which the response probabilities are assumed to be equal:

$$DB080_h^{(N)} = DB080_h \cdot \frac{1}{R_K}$$

Where R_K denotes s the (weighted) response rate in the group (*k)* the household (*h)* belongs to:

$$R_K = \frac{sum\ of\ design\ weights\ of\ responding\ units\ in\ cel\ k}{sum\ of\ design\ weights\ of\ selected\ units\ in\ cell\ k}$$

See construction of weights in the first part of section 8: weights.

DB090: HOUSEHOLD CROSS-SECTIONAL WEIGHT

Topic and detailed topic: Technical items/Weights
Variable type: Annual
Unit: Household
Reference period: Current
Mode of collection: Derived
In use (period): Yes, since the first year of EU-SILC data collection
Series' differences: Yes, 2014

VALUES AND FORMAT

0 (format 2,5) Weight

Required format Household cross-sectional weights will be coded with at least one integer and five decimals. In the regular transmission (reconciled file format) these variables should only be filled in for the records related to the last year of operation.

FLAGS

From 2014 onwards

1 Filled
-7 Not applicable (DB010 ≠ last year of operation)

Before 2014

1 Filled

DESCRIPTION

The household cross-sectional weights are the final estimation weights. Only the households that are accepted into the database (DB135 = 1) have a cross-sectional weight; the others are assigned a weight of 0. The calibration is done taking all rotational groups together.

Units of the sub-samples appear in the EU-SILC sample only in situations where a rotational design is used. In this case, sample weights are called 'cross-sectional' weights. These weights are required for all the types of units considered in the EU-SILC survey.

The household cross-sectional weights (target variable DB090) will be used to draw inference on the population of private households at national and European levels *(see the weight for DB080 and then further).*

After adjustments for non-response and to external sources (calibration) of the household design weight, the cross-sectional household weight (DB090) is calculated. *DB090 is used to weigh household data and indicators produced at household level.*

More precisely, suppose that there are 'J' auxiliary variables $x_1 \ldots x_j \ldots x_J$, called calibration variables, with known population totals (for the numerical variables) or marginal counts (for the categorical variables). Without loss of generality, we can assume that all the calibration variables are numerical (otherwise, we consider the 0/1 variables for each category).

New household weights (DB090) are 'as close as possible' (as determined by a certain distance function) to the initial weights $DB080^{(N)}$. These new weights are calibrated on the totals X_j of the 'J' auxiliary variables; in other words, they verify the calibration equations:

$$\forall j = 1 \ldots J \qquad \sum_{k \in s} DB090_k \cdot x_{jk} \quad = \quad X_j.$$

Where $DB090_k = g_k \times DB080^{(N)}$

This process can be done using different statistical software and different methods (e.g., logit regressions through the SAS macro, CALMAR).

When using CALMAR, it is recommended to use a bounded method and to impose lower and upper bounds LO and UP on the weight adjustment factors 'g_k', usually referred to as g-weights. In practice, one has to bear in mind that the choice of bounds is not free and directly depends on the calibration variables that are chosen: the limits must be adjusted, taking into account the differences between the estimates based on the 'old' initial weights and the benchmark totals that the new weights are to reproduce, so CALMAR can find a solution within the constraints applied to the problem. In practice, those limits are determined by some 'guess and check': we start with a small interval [LO, UP] and we enlarge it until CALMAR finds a solution. Applying calibration bounds prevents negative and extreme weights.

It can use different household variables and individual variables aggregated in the household level.

In the weight section, a more extensive explanation on the weighting procedures is given.

DB095: HOUSEHOLD LONGITUDINAL WEIGHT

Topic and detailed topic: Technical items/Weights information
Variable type: Annual
Unit: Household
Reference period: Current
Mode of collection: Derived
In use (period): Yes, since 2014
Series' differences: No changes

VALUES AND FORMAT

0 (format 2,5)	Weight
Required format	These weights had to be coded with at least one integer and five decimals.

FLAGS

1	Filled
-2	Not applicable ((DB110 = 9) and (DB010 = last year of operation))

DESCRIPTION

The household longitudinal weights are the final estimation weights. Only the households that are accepted into the database (DB135 = 1) have a longitudinal weight; the others are assigned a weight of 0. The calibration is done taking all rotational groups separately.

The household longitudinal weight DB095 is a product of the first calibration of the population. In year X, after the non-response adjustment that is applied to each rotational group separately, each rotational group should be calibrated separately to the cross-sectional population referring to the end of the year X-1. The calibration should be done using the integrative calibration approach to ensure that each member of the same household receive the same weight.

Household longitudinal weights are defined for all households accepted into the database. To combine the calibrated weights, they should be scaled based on the number of rotational groups in the longitudinal sample of the survey. A particular rotational group should be surveyed at least 2 consecutive years before being part of the longitudinal component. Each wave of the longitudinal component contains different numbers of rotational groups (3, 2 and 1). For this reason, household longitudinal weights should be scaled with different scaling factors for each year. The table[29] below illustrates the scaling factors and the approximate target population.

These weights will be part of D-file as well as part of the longitudinal data set delivered each year. The set consists of at least three panels of duration: 2, 3 and 4 years.

The variable DB095 should be adjusted backwards with every data transmission because the number of surveyed rotational groups change with the evolution of the panel.

The sum of all non-zero DB095 for year X should be approximately equal to sum of non-zero DB090 for that year.

In the weights section, a more extensive explanation of the weighting procedures is given.

[29] The table illustrates a classical EU-SILC rotational design with 4 rotational groups. If a country uses rotational design with different numbers of rotational groups the scaling factors should be adjusted accordingly.

DB075/ DB010	X-3	X-3 Flag	X-2	X-2 Flag	X-1	X-1 Flag	X	X Flag
4	$\approx \frac{DB090\ (x-3)}{1}$	Filled	$\approx \frac{DB090\ (x-2)}{2}$	Filled	$\approx \frac{DB090\ (x-1)}{3}$	Filled	$\approx \frac{DB090\ (x)}{3}$	Filled
3	-	-	$\approx \frac{DB090\ (x-2)}{2}$	Filled	$\approx \frac{DB090\ (x-1)}{3}$	Filled	$\approx \frac{DB090\ (x)}{3}$	Filled
2	-	-	-	-	$\approx \frac{DB090\ (x-1)}{3}$	Filled	$\approx \frac{DB090\ (x)}{3}$	Filled
1	-	-	-	-	-	-	-	Not applicable ((DB0110 = 9) and (DB010 = last year of operation))

DB100: DEGREE OF URBANISATION

Topic and detailed topic: Technical items/Localisation

Variable type: Core variable/ Annual

Unit: Household

Reference period: Current

Mode of collection: Derived

In use (period): Yes, since the first year of EU-SILC data collection

Series' differences: Yes (2012 and 2021 modalities were changed)

VALUES AND FORMAT

From 2021 onwards

1	Cities
2	Towns and suburbs
3	Rural areas

Before 2021

1.	Densely-populated area
2.	Intermediate area
3.	Thinly-populated area

FLAGS

1	Filled
-1	Missing (allowed only from wave 2 onwards)

DESCRIPTION

This variable reports on the degree of urbanisation in the area where the usual residence of the person or the household is located.

From 2021 onwards the modalities were changed according to the standardised social variables.

This variable must be filled in for every household in wave 1. From wave 2 onwards, a missing value (flag -1) is allowed in exceptional cases (like moving house).

From 2021 onwards the variable classifies LAU2 into three types of area:

1. *'Cities'* - densely-populated areas where at least 50% of the population live in an urban centre.
2. *'Towns and suburbs'* - intermediate density areas where at least 50% of the population live in urban clusters, but which are not 'cities'.
3. *'Rural areas'* - thinly populated areas where more than 50% of the population live in rural grid cells.

This classification is based on a combination of criteria of geographical contiguity and minimum population threshold applied to 1 km^2 population grid cells.

The LAU2 list including the degree of urbanisation is published by Eurostat on the RAMON server: https://ec.europa.eu/eurostat/ramon/miscellaneous/index.cfm?TargetUrl=DSP_DEGURBA.

For further methodological details on the classification of LAU2s by degree of urbanisation, see http://ec.europa.eu/eurostat/web/degree-of-urbanisation/methodology.

For further information, see:

Local Administrative Units (LAU) - NUTS - Nomenclature of territorial units for statistics - Eurostat (europa.eu)

Modalities used before:

- Densely populated area: contiguous grid cells of 1km^2 with a density of at least 1,500 inhabitants per km^2 and a minimum population of 50,000.
- Intermediate area: clusters of contiguous grid cells of 1km^2 with a density of at least 300 inhabitants per km^2 and a minimum population of 5,000.
- Thinly-populated area: Grid cells outside urban clusters.

DB110: HOUSEHOLD STATUS

Topic and detailed topic: Technical items/Data collection information

Variable type: Annual

Unit: Household

Reference period: Current

Mode of collection: Interviewer

In use (period): Yes, since 2005

Series' differences: No changes

VALUES AND FORMAT

Household from previous wave
- *1* At the same address as last interview
- *2* Entire household moved to a private household within the country

Household no longer in-scope
- *3* Entire household moved to a collective household or institution within the country
- *4* Household moved outside the country
- *5* Entire household died
- *6* Household does not contain sample person

Address non-contacted
- *7* Household cannot be accessed (due to, for example, climatic conditions)
- *11* Lost household (no information on record about what happened to the household)

New household for this wave
- *8* Split-off household
- *9* New address added to the sample of this wave or first wave

Fusion
- 10 Fusion household

FLAGS

- 1 Filled

DESCRIPTION

The household is at the same address as the last interview: This situation appears when at least one of the sample persons lives at the same address as in the previous wave.

Entire household moved to a private household within the country: This situation appears when no sample persons live at the same address as in the previous wave, but when the household moved to a private household within the country.

Entire household moved to a collective household or institution within the country: This situation appears when all the sample persons moved to a collective household or institution within the country.

Household moved outside the country or to territories not covered by the survey: This situation appears when all the sample persons moved outside the country or to territories not covered by the survey.

Entire household deceased: This situation occurs when all the sample persons have died.

Household does not contain sample person: This situation occurs when sample persons are no longer in scope for a variety of reasons.

Address non-contacted: This category distinguishes between households that cannot be accessed and households that are 'lost', i.e. no information on record as to what happened to the household.

Split-off households: New household which has split off from the initial household since the previous wave and now considered as a separate household.

New address added to the sample in this wave or first wave: This situation occurs where for the first time this household is recorded in the survey and is not a split-off household (first wave, new household introduced into the sample in order to maintain the sample size because of loss due to non-response).

Fusion: This situation occurs when sample persons from different sample households from the previous wave join together to form a new household. The household that disappears will be coded as 'fusion household' while the other one is considered as the main household and keeps the identification number. The household that keeps the identification number will be coded as: '1' (if it remains at the same address as in

the last interview), as '2' (if the entire household moved to a private household within the country).

Initial contact with the address

It is recommended that a letter be sent prior to visiting the household to inform its members that they have been selected to participate in the survey, as well as to provide the main characteristics of the survey and to request a visit to conduct an interview.

Where possible, it is recommended that an appointment by phone be made in order to reduce non-contact in case the person from the home is absent, sick, etc.

If the interviewer has any difficulty in finding an address there are several sources which may be helpful such as the local post office, the City Hall, the police, etc.

Where known, the telephone number of the household should be included in the instructions for the interviewer.

Second, third and subsequent contacts

A major risk of attrition in a longitudinal survey is linked to either individuals or the entire household moving between waves. Special procedures need to be established to trace all moving/split-off households. These procedures can be established by the interviewer's organisation or the central organisation:

In order to trace people or households who move between waves, the interviewer can take several measures, such as: (a) asking at each interview about any intention or expectation of a move before the next interview; (b) contacting households by mail or phone in the intervening period between waves; (c) requesting that the household informs the interviewer if a move takes place providing for such information appropriate financial incentives; (d) where a move is likely to occur, asking for the name and address of a friend or relative who could inform the interviewer about the new location.

For those countries using a sample of addresses or households, the first task at each interview is to get all the information needed to identify the household members as well as information on any changes in the household composition. It is important to obtain the date of move, the reason for the move and the new address of the household.

The interviewer should try to contact neighbours, population registers or another source of information in order to discover the situation of sample persons who moved since last interview and who have not left their new address. Once the new address has been obtained, the interviewer should try to locate the household at the new address if within the interviewer's area. Where the address is outside of the interviewer's area, the interviewer should advise their supervisor of the change of address.

DB120: CONTACT AT ADDRESS

Topic and detailed topic: Technical items basic data/ Data collection information

Variable type: Annual

Unit: Household

Reference period: Current

Mode of collection: Interviewer

In use (period): Yes, since 2008

Series' differences: No changes

VALUES AND FORMAT

11 Address/phone contacted

Address/phone non-contacted:

21 Address/phone non-contacted: not located/not found

22 Address/phone non-contacted: unable to access

23 Address/phone non-contacted: non-existent/non-residential or non-private/unoccupied/not principal residence

FLAGS

1 Filled

-2 Not applicable (DB110 not equal to 2, 8 or 9)

DESCRIPTION

This variable needs to be filled in only when DB110 = '2', '8', or '9'. Otherwise, no values are to be given for this variable and it must be flagged '-2' (not applicable).

11 Address contacted/phone should be used when:

- a household has its main residence at the sampled address which can be located and accessed.

- an entire household, or a part of the household (split-off household) has moved to a private household within the country and the new address can be accessed.

- Computer-assisted telephone interviewing (CATI) is used, and the phone number is available.

21 Address/phone non-contacted: not located/not found should be recorded when:

- it is not possible to locate the address despite special efforts being made to do so.

- the CATI method is used and the phone number of that household could not be found.

22 Address/phone non-contacted: unable to access should be recorded when:

- it is not possible to access the address due to weather factors (like flooding, etc.) or geographic factors, for example, there may be no access roads. The factor causing the lack of access should not be only temporary.

- the CATI method is used and the number cannot be reached.

23 Address/phone non-contacted: non-existent/non-residential or non-private/unoccupied/not principal residence should be recorded when:

- an address is non-residential if it is used for different purposes than those of a principal residence (business, shop, vacation home, etc.).
- an address does not exist if the building has been demolished.
- an address is unoccupied or empty if nobody is currently living there.
- a CATI interview is used and a phone number does not exist or the phone belongs to a company, shop, hotel, etc.

For those countries selecting a sample of persons, such as one from a population register, the goal is to contact these people wherever they live. The status 'address does not exist, is non-residential, is unoccupied or is not principal residence' should not be relevant. In the case of death of the selected respondent, the household can be classified as 'address is unoccupied'.

DB130: HOUSEHOLD QUESTIONNAIRE RESULT

Topic and detailed topic: Technical items Basic data/Data collection information

Variable type: Annual

Unit: Household

Reference period: Current

Mode of collection: Interviewer

In use (period): Yes, since 2008

Series' differences: No changes

VALUES AND FORMAT

11	Household questionnaire completed
21	Refusal to cooperate
22	Entire household temporarily away for duration of fieldwork
23	Household unable to respond (illness, incapacity…)
24	Other reason

FLAGS

1	Filled
-1	Missing
-2	Not applicable (DB120 not 11 and DB110 not equal to 1)

DESCRIPTION

This variable needs to be filled in only when DB120 = 11 or when DB110 = 1. Otherwise, no values are to be given for this variable and it must be flagged '-2' (not applicable). The information is collected for each address contacted if the household questionnaire is completed, refused, entire household is temporary away, household is unable to respond or the questionnaire is not completed for other reasons.

11 Household questionnaire completed: A household questionnaire is considered as completed if most of its variables have been filled in.

21-24 Interview not completed: In the case of an interview survey, at least three call-backs must be made before a household is considered as being non-responsive (interview not completed), unless there are clear reasons (such as a definite refusal to cooperate, circumstances endangering the safety of the interviewer, etc.) why this household cannot be interviewed.

The interview has not been completed for some of the following reasons:

21 Refusal to cooperate

The household refuses to give information either from the beginning of the interview or having initially agreed to provide the information, refuses to do so at a later date (for any reason).

The interviewers should do their best to get cooperation not only by explaining the survey, but also by calling back again when an appointment is broken.

22 Entire household temporarily away for duration of fieldwork

An entire household is temporarily away when all its household members are temporarily away and will not return during the period when fieldwork is being carried out in the area.

Before declaring a household as being temporarily away, the interviewer must ensure that all necessary attempts to contact the household have been made (visiting the household at different hours, different days, calling by phone, etc.).

23 Household unable to respond (illness, incapacity, etc.)

This situation appears when either all household members are unable to respond to the interview or are unable to self-administer the questionnaires due to incapacity, illness etc.

In the case where a one-person household is unable to respond to or self-administer the questionnaire due to incapacity or illness, the interviewer should, if feasible, try to contact someone outside the household who is able to help or provide the information on behalf of the incapacitated person.

The interviewer must only assign the code 'household unable to respond' when the illness or incapacity is not temporary; where the illness or incapacity is temporary, the interviewer should return to the household during the period when fieldwork is being carried out in the area.

24 Other reasons

This item refers to any situation that is not referred to above, either because all the household members do not know the language, all of them are illiterate, etc., and nobody outside the household can provide the information.

DB135: HOUSEHOLD INTERVIEW ACCEPTANCE

Topic and detailed topic: Technical items Basic data/Data collection information
Variable type: Annual
Unit: Household
Reference period: Current
Mode of collection: Derived
In use (period): Yes, since 2008
Series' differences: No changes

VALUES AND FORMAT

1	Interview accepted for database
2	Interview rejected

FLAGS

1	Filled
-1	Missing
-2	Not applicable (DB130 not equal to 11)

DESCRIPTION

The interview must be accepted for the database if at least one of the personal interviews is completed.

In those countries for which a selected respondent is chosen for the personal interview, the interview of the selected respondent must be completed.