



EUROPEAN COMMISSION  
EUROSTAT

Directorate F: Social statistics  
Unit F-3: Labour market and lifelong learning

Luxembourg, December 2023  
ESTAT/F-3

**Subject: Information note on the use of the 2023 release of EU LFS scientific use files**

Dear Sir/Madam,

Please find enclosed the 2023 release of the anonymised EU Labour Force Survey (LFS) scientific use files for the reference years 1983-2022.

Data cover all EU Member States as well as Iceland, Norway, Switzerland and the United Kingdom subject to data availability. Particular country datasets may be excluded as a result of the approval procedure for research project proposals. Besides core LFS data, the datasets also include the corresponding modules for the reference years 2005 to 2022. No ad-hoc module was conducted in 2015.

The encrypted and extractable core LFS data files are organised by country. For each country, there are two zip files, one with all yearly datasets and one with all quarterly datasets. The yearly datasets contain the respective module variables for a given reference year and all household level data. Where subsamples are used for specific sets of variables, these subsamples can be identified by selecting the observations with the respective weights greater than zero. More detail is explained in the general notes following this letter.

The coding of the variables follows the standard introduced with the new legal basis for the EU LFS from 2021. Data up to 2020 have been recoded accordingly to the extent possible. The EU LFS User Guide provides detailed information on the changes introduced in 2021 and on the respective recoding.

The following supporting documentation is provided:

- The EU LFS User Guide describing the data underlying the 2023 release. Chapter 7 describes the anonymisation criteria for the scientific use files.
- The EU LFS explanatory notes valid from 2021 (reporting guidelines for the Member States).
- A list of data revisions received since the last release.
- Comma-Separated Value (CSV) metadata files listing all datasets included in this release; these metadata files also include information on the sample size per data file.
- Comma-Separated Value (CSV) metadata files providing the lists of available quarterly and yearly variables by reference year.

In addition to age in 5-year bands, data on age in single years is provided for most countries. It should be noted, however, that the data on age in single years must not be disclosed in tables published. For more details, please refer to the [terms of use](#) and [guidelines for publication](#).

The variable INCDECIL is not yet available from reference year 2021 onwards. While countries reported this variable to Eurostat up to 2020, they are now required to report detailed data on the gross monthly pay from the main job in the variable INCGROSS, which is not provided in the scientific use files for reasons of confidentiality. Eurostat intends to derive income deciles from the source data in the future. A detailed quality analysis of the INCGROSS data is currently ongoing, which does not allow yet the inclusion of a derived INCDECIL in the scientific use files.

Yours sincerely,

(Signed)  
Didier Dupré  
Head of Unit

#### GENERAL NOTES ON THE DATA INCLUDED IN THE RELEASE

The structure of the LFS changed in 2006 to reduce the burden on respondents. Since then, all variables have been collected on a yearly basis, but only a selection of them on a quarterly basis. In addition, there are biennial variables starting with reference year 2021. The yearly, biennial and module variables are also referred to as structural variables. To comply with this structure, the LFS data are divided into two databases:

- Files ending in q1, q2, q3 and q4 contain only the quarterly variables.
- Files ending in \_y contain all variables of the LFS survey, the quarterly and the structural ones.

Data for years prior to 2006 are treated in the same way as data from 2006 onwards, i.e. a similar split into quarterly and yearly datasets is carried out. In consequence, the quarterly files for reference years 1983 to 2005 also contain only the set of quarterly variables. Please note that the split into quarterly and yearly variables applied in this release is the one valid from 2021 onwards.

Data for all four quarters of a year are in principle progressively available starting between 1998 and 2004 for all countries, except for Germany for which quarterly data are available from 2005 and Croatia for which this is the case from 2007 (until then, Croatia transmitted two semi-annual files per year). For comparison reasons, the reference sample for yearly files hence corresponds to one reference quarter in spring until 2004, and to an annual sample covering all weeks of the year from 2005.

For the years from 2005 onwards this annual sample consists of the full set of the 4 quarterly samples in all countries with the following exceptions:

- A sub-sample representing the whole year is used for ES, FR, NL, NO from 2006, DE from 2006 to 2011 and again from 2020, FI from 2007, BG, UK from 2008, CH from 2010, CZ from 2011, LU from 2015, BE and LV from 2017, IE, EL, PT and SE from 2021.
- Quarter 2 is used for IE and FI from 2005 to 2006, UK from 2005 to 2007, CH from 2005 to 2009.

These special annual sub-samples are usually smaller than the ones of a single quarter – the exact sample size depends on the rotation scheme in place in a given country. For further information on subsampling from reference year 2021, please refer to chapter 3.3 of the EU LFS User Guide.

While the quarterly datasets contain only the quarterly variables, the yearly datasets contain all variables of the LFS survey, i.e. the quarterly, annual, biennial (from 2021) and module variables. Household level data including derived variables are in the yearly datasets as well from 2006. Where households are generally surveyed, the full annual datasets are relevant for household level analysis. Some countries survey individuals in general and use an annual subsample of households. These household subsamples (for DK and FI from 2006, SE from 2009, LU from 2015 and NL from 2021) are included in the yearly datasets as well.

The (sub)samples to be used for analysis of quarterly, annual, biennial and module variables, and of household data, can be identified by selecting the observations with weighting coefficients greater than zero. The weighting coefficients are:

#### Quarterly datasets

Quarterly variables: COEFFQ

#### Yearly datasets

Annual variables: COEFFY

Biennial variables: COEFF2Y (from 2021)

Module variables: COEFFMOD

Household variables: COEFFHH, COEFFHHA VG

Analyses at individual level, using the household background, should use the (individual) household weight COEFFHH. The average weight of all household members (COEFFHHA VG) is to be used for analyses at household level.

From reference year 2021 onwards, subsamples can also be identified with the technical variable INTQUEST. It is more detailed than before and indicates precisely the assignment of each observation to the (sub)samples of the survey. Where household subsampling is used, INTQUEST codes 09 and 10 identify the additional household members of the initially selected individual. Code 10 is only used where a module requires household background information and a country decided to use a different household subsample for module background information than for the core data (e.g. in NL). The cases with INTQUEST 10 do not have any weighting factors as they only provide background information on the initially selected individual in the module subsample. Here, only the COEFFMOD of this individual is relevant for the analysis of the module data.

Data for EU countries are usually available depending on their EU accession date, and from 2002 for all EU countries except Malta. Data for Malta are included from 2009. For the United Kingdom, due to Brexit, quarterly data are covered up to 2020 Q3 and yearly data up to 2019. Quarterly data for Germany of 2020 are excluded because of quality constraints.

The United Kingdom is included in the “EU28” aggregate in the country variables up to 2019. From reference year 2020, the aggregate “EU27\_2020” is used for the European Union; the United Kingdom is included in the aggregate of “Other Europe”.

More details on the structure and content of the database and the anonymisation criteria applied can be found in the EU LFS User Guide. Additional useful information about the EU LFS is also available at

<https://ec.europa.eu/eurostat/web/lfs/overview>

and

[https://ec.europa.eu/eurostat/statistics-explained/index.php/EU\\_labour\\_force\\_survey](https://ec.europa.eu/eurostat/statistics-explained/index.php/EU_labour_force_survey)