

## EU-Labour Force Survey December 2011 release

### Setup for Importing the Anonymised Yearly Data Sets for 2002–2005

#### Content

##### I. Overview

- II. Explanatory notes on the recoding of alphanumeric variables into numeric variables
- III. Explanatory notes about the recoding of missing values
- IV. Explanatory notes about specific recoding and the definition of variable and value labels
- V. Notes on irregularities not mentioned in the User Guide
- VI. Classifications

##### I. Overview

The following information will help you to import the EU-LFS 2002–2005 yearly data. The data you received from Eurostat are formatted as comma separated values (csv). With the routines provided by GESIS's German Microdata Lab, these data can be converted into SPSS.

The routine to convert the data into SPSS consists of two files:

Setup\_EULFS\_2002-2005\_y.sps  
Labels\_EULFS\_1983-2010.sps

The setup provided by the German Microdata Lab achieves the following objectives:

- Transformation of EU-LFS 2002–2005 yearly data from csv into sav (SPSS)
- Recoding of alphanumeric variables into numeric variables
- Recoding of missing values
- Definition of variable and value labels (by executing Labels\_EULFS\_1983-2010.sps).

For a first check of the converted data, you can compare the number of cases per country with the number of cases listed in table 1. Consistent numbers indicate that the execution of the setup was successful.

## II. Explanatory notes on the recoding of alphanumeric variables into numeric variables

### – COUNTRYW, COUNTR1Y

The first two digits indicate the region, according to the country codification from 2010 (e.g. 05 for North Africa).<sup>1</sup> Countries not explicitly assigned to one region by this classification have been classified by GESIS.

The last two digits usually display the alphabetical order of a country in the corresponding region as given by the ISO code elements (e.g. 01 for Algeria, so that the whole code for Algeria is 0501).<sup>2</sup> The alphanumerical code from the EU-LFS deviated from the ISO classification with regard to the United Kingdom (UK instead of GB).

In Slovenia, the coding of COUNTRYW in the raw data differs from the other countries (see User Guide 2011, p. 63). Thus the recoding within the Setup is different as well. The cases of "EU15" and "Non EU15" in the years 2002 and 2003 have been assigned the values "1501" and "1502".

### – HHNUM, QHHNUM, QUARTER, REGION, REGIONW, REGION1Y

These variables have a large amount of values or – in case of QUARTER – only one value. Thus they remain alphanumeric.

## III. Explanatory notes about the recoding of missing values

Coding standard for missing values:

- -1 for "No answer"
- -2 for "Not applicable"
- -3 for "Not available"
- -4 for "Not specified"
- in exceptional cases -9 for "Otherwise" (only LEAVCLAS) and -101 for "No answer" and -102 for "Not applicable" (only STARTIME and LEAVTIME, see below)

In SPSS these missing values are declared as user-missing values.

The Code -3 for "Not available" is assigned, if a variable has no valid cases in a country. Some of these variables have only system-missings (i.e. blanks) in the raw data of the respective country, others have only cases with "No answer" and/or "Not applicable".

---

<sup>1</sup>[http://circa.europa.eu/irc/dsis/employment/info/data/eu\\_lfs/LFS\\_MAIN/Coding\\_lists\\_explanatory\\_notes\\_and\\_classifications/Country\\_codification\\_from\\_2010\\_onwards.pdf](http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Coding_lists_explanatory_notes_and_classifications/Country_codification_from_2010_onwards.pdf)

<sup>2</sup>[http://www.iso.org/iso/country\\_codes/iso\\_3166\\_code\\_lists/country\\_names\\_and\\_code\\_elements.htm](http://www.iso.org/iso/country_codes/iso_3166_code_lists/country_names_and_code_elements.htm)

Specifics and exceptions:

– COUNTRYB, NATIONAL

The code -1 includes not only cases without a valid answer, but also suppressed cases, cases with the original code "Other or Stateless", and in some countries certain other cases (see User Guide 2011, p. 67).

– MARSTAT

In Slovenia children under the age of 15 had not been asked about their marital status. Thus -1 for "No answer" has been recoded to -2 for "Not applicable" in those cases.

– STARTIME, LEAVTIME

These variables had valid negative values in previous years. Thus the missing value category is set to -101 (for "No answer") and -102 (for "Not applicable").

#### IV. Explanatory notes about specific recoding and the definition of variable and value labels

– HHCOMP

The variable contains the value 50, which is not mentioned in the User Guide. All persons with this value are children with no adult household members present in the dataset. Thus we have labelled the value "No adult household members present in the dataset".

– ISCO3D, ISCOPR3D

A number of cases belonging to the category "armed forces" had been coded incorrectly in the 3-digit ISCO code. This became obvious while crosstabulating the 1-digit with the 3-digit ISCO code. Thus these cases have been recoded.

– REGION, REGIONW, REGION1Y

The values of these variables have country-specific meanings (according to the NUTS classification).<sup>3</sup> Thus they remain unlabelled.

#### V. Notes on irregularities not mentioned in the User Guide

– HHSEQNUM

In countries where the final sampling unit are persons<sup>4</sup> and not households, this variable does not contain any useful values.

– HHSEQNUM, INTWEEK, INTWAVE, DEGURBA, METHODDA – METHODDM

These variables contain system-missings in some countries. The origin of these missings is unclear.

---

<sup>3</sup>[http://circa.europa.eu/irc/dsis/employment/info/data/eu\\_lfs/LFS\\_MAIN/Coding\\_lists\\_explanatory\\_notes\\_and\\_classifications/NUTS\\_2006\\_from\\_2007\\_onwards.pdf](http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Coding_lists_explanatory_notes_and_classifications/NUTS_2006_from_2007_onwards.pdf)

<sup>4</sup> See section "Characteristics of the national surveys" at [http://circa.europa.eu/irc/dsis/employment/info/data/eu\\_lfs/lfs\\_main/lfs/lfs\\_publications.htm](http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/lfs_main/lfs/lfs_publications.htm)

– YEARESID

In 2004 and 2005, there are only blanks in Iceland for persons who have not been born in Iceland. Thus there are only two categories (0 "Born in the country" and -1 "No answer"). In 2004, the same is true for Romania.

– COEFF

There is a large number of cases with a weighting factor of zero in Poland. There are a lot of system-missings in Finland (except 2002). The origin of these missings is unclear.

– Data of Greece for 2005 (gr2005\_y)

There are seven cases with defective raw data. Due to a supernumerary comma for the variable COURPURP in the raw data, the values of all following variables are mismatched.

To exclude these cases from the analysis, they can be identified as follows:

QHHNUM=Q124203 and HHSEQNUM=1, QHHNUM=Q125947 and HHSEQNUM=4,

QHHNUM=Q129646 and HHSEQNUM=1, QHHNUM=Q148871 and HHSEQNUM=3,

QHHNUM=Q152446 and HHSEQNUM=3, QHHNUM=Q152446 and HHSEQNUM=4,

QHHNUM=Q154574 and HHSEQNUM=4

## VI. Classifications

– ISCED classification (EDUCLEVEL, HATLEVEL):

[http://circa.europa.eu/irc/dsis/employment/info/data/eu\\_lfs/LFS\\_MAIN/Related\\_documents/ISCED\\_EN.htm](http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Related_documents/ISCED_EN.htm)

– other classifications used in the EU LFS:

[http://circa.europa.eu/irc/dsis/employment/info/data/eu\\_lfs/LFS\\_MAIN/Coding\\_lists\\_explanatory\\_notes\\_and\\_classifications/Index\\_EU\\_LFS\\_coding\\_%20lists\\_explanatory\\_notes\\_and\\_classifications.htm](http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Coding_lists_explanatory_notes_and_classifications/Index_EU_LFS_coding_%20lists_explanatory_notes_and_classifications.htm)

---

GESIS - Leibniz-Institute for the Social Sciences

German Microdata Lab

Square B2, 1

D-68159 Mannheim

Tel: 0621-1246-265

Fax: 0621-1246-100

<http://www.gesis.org/gml>

E-Mail: [gml@gesis.org](mailto:gml@gesis.org)

Table 1: Number of cases per country in the EU-Labour Force Survey  
2002-2005, yearly data<sup>a</sup>

Country	N (without selection)			
	2002	2003	2004	2005
01 AT Austria	58.717	56.268	45.185	205.058
02 BE Belgium	27.834	27.565	27.739	109.125
03 BG Bulgaria	57.549	38.524	36.540	140.035
04 CH Switzerland	41.263	57.679	54.229	51.791
05 CY Cyprus	10.667	10.779	10.615	39.514
06 CZ Czech Republic	62.091	60.973	62.807	249.737
07 DE Germany	328.074	330.297	327.088	477.171
08 DK Denmark	16.081	15.471	15.445	59.792
09 EE Estonia	4.948	4.735	4.533	17.904
10 ES Spain	172.552	175.209	175.159	610.851
11 FI Finland	42.743	40.782	39.680	39.500
12 FR France	175.939	86.889	87.774	342.099
13 GR Greece	77.451	73.017	80.992	314.576
14 HU Hungary	82.904	88.212	82.310	312.190
15 IE Ireland	105.569	101.500	86.545	91.174
16 IS Iceland	3.646	3.186	3.110	12.424
17 IT Italy	194.041	192.359	172.264	704.372
18 LT Lithuania	12.993	12.219	12.149	47.947
19 LU Luxembourg	13.429	16.394	21.189	90.024
20 LV Latvia	5.943	6.022	6.020	21.919
21 MT Malta	-	-	-	-
22 NL Netherlands	97.594	96.740	112.913	469.514
23 NO Norway	20.838	20.931	21.306	85.331
24 PL Poland	58.623	58.303	57.433	225.325
25 PT Portugal	45.617	46.385	50.714	190.128
26 RO Romania	41.757	41.556	68.596	269.408
27 SE Sweden	55.699	56.316	52.185	148.124
28 SI Slovenia	19.766	19.881	18.871	71.406
29 SK Slovak Republic	29.420	28.889	28.772	116.250
30 UK United Kingdom	136.155	130.417	125.610	123.141

<sup>a</sup> see also "datafileinfo\_2002-2005.csv" on the DVD you received from Eurostat