

## EU-Labour Force Survey December 2011 release

### Setup for Importing the Anonymised Yearly Data Sets for 2006–2008

#### 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 2006–2008 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\_2006-2008\_y.sps  
Labels\_EULFS\_1998-2010.sps

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

- Transformation of EU-LFS 2006–2008 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\_1998-2010.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.

### – 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.

### – YEARESID

Since 2008 the 5-year groups (11-14, 15-19 etc.) had been coded alphanumerically and had to be recoded into numeric values in line with the standard aggregation of AGE. Until 2007, there is only one group for "more than 10 years".

## 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.

---

<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)

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".

For detailed information on which variables have no valid cases in which countries and have thus been recoded to -3 see table 2.

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).

– HWUSTRAI, LENTR4W, PURP4W, INCDECIL

These variables have no valid cases in the anonymised data sets of all countries. Thus they have been recoded to -3.

– 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").

– PROXY

Children under the age 15 had not been asked about the nature of participation. Thus -1 for "No answer" has been recoded to -2 for "Not applicable" in those cases.

#### 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.

---

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

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

### – YEARESID

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").

### – COEFF

There is a large number of cases with a weighting factor of zero in France. A number of cases with a weighting factor of zero appear in Germany (in 2008 only), Spain and Poland. There are a lot of system-missings in Finland and in the United Kingdom (in 2008 only). A small number of system-missings appear in Belgium (in 2006 only) and Greece (in 2007 only). The origin of these missings is unclear.

### – HHSEQNUM

In Iceland, Sweden and Switzerland all cases have the value 1. In Denmark all (in 2007 and 2008) or nearly all (in 2006) cases have system-missing values. There are some system-missing in Romania too. The origin of these missings is unclear.

### – INTWEEK

There are some system-missings in Denmark, France and Norway, Portugal and Sweden. The origin of these missings is unclear.

## VI. Classifications

### – ISCED classification (EDUCLEVL, 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  
2006-2008, yearly data<sup>a</sup>

Country	N (without selection)		
	2006	2007	2008
01 AT Austria	201.083	201.771	194.070
02 BE Belgium	114.278	111.978	105.604
03 BG Bulgaria	134.298	130.964	34.835
04 CH Switzerland	48.262	48.485	47.899
05 CY Cyprus	38.227	38.462	40.380
06 CZ Czech Republic	253.254	250.602	241.934
07 DE Germany	49.685	49.178	47.915
08 DK Denmark	58.899	117.997	115.508
09 EE Estonia	20.189	23.342	22.242
10 ES Spain	106.279	107.224	109.503
11 FI Finland	38.353	60.517	60.953
12 FR France	339.775	350.757	348.568
13 GR Greece	303.621	296.383	293.618
14 HU Hungary	311.397	305.048	290.431
15 IE Ireland	85.734	85.133	74.368
16 IS Iceland	12.550	12.435	12.356
17 IT Italy	684.303	677.746	671.939
18 LT Lithuania	44.958	63.022	61.776
19 LU Luxembourg	85.080	21.178	14.288
20 LV Latvia	18.640	37.822	38.597
21 MT Malta	-	-	-
22 NL Netherlands	108.545	106.472	109.585
23 NO Norway	20.663	20.785	20.163
24 PL Poland	216.840	206.875	206.893
25 PT Portugal	179.534	171.960	167.970
26 RO Romania	257.982	249.521	240.785
27 SE Sweden	209.871	208.058	205.262
28 SI Slovenia	69.813	67.877	65.600
29 SK Slovak Republic	113.417	110.377	109.840
30 UK United Kingdom	120.651	120.379	193.630

<sup>a</sup> see also "datafileinfo\_2006-2008.csv" on the DVD you received from Eurostat

Table 2: Variables without valid cases in several countries,<sup>a</sup> EU LFS 2006–2008, yearly data

	2006	2007	2008
COUNTR1Y	x	x	x
COURFILD	x	x	x
COURLEN	x	x	x
COURPURP	x	x	x
COURWORTH	x	x	x
DEGURBA	x	x	x
EDUCFILD	x	x	x
EDUCLEVL	x	x	x
EVENWK	x	x	x
FTPTREAS	x	x	x
HATFILED		x	
HATYEAR	x	x	x
HHAGE14	x	x	x
HHAGEYG	x	x	x
HHCHILDR	x	x	x
HHCOMP	x	x	x
HHFATH	x	x	x
HHINST	x	x	x
HHLINK	x	x	x
HHMOTH	x	x	x
HHNB0014	x	x	x
HHNBCH2	x	x	x
HHNBCH5	x	x	x
HHNBCH8	x	x	x
HHNBCH11	x	x	x
HHNBCH14	x	x	x
HHNBCH17	x	x	x
HHNBCH24	x	x	x
HHNBCHLD	x	x	x
HHNBEMPL	x	x	x
HHNBINAC	x	x	x
HHNBOLD	x	x	x
HHNBPERS	x	x	x
HHNBUNEM	x	x	x
HHNBWORK	x	x	x
HHPARENT	x	x	x
HHPARTNR	x	x	x
HHPERS	x	x	x
HHSPOU	x	x	x
HWOVERP	x	x	x
HWOVERPU	x	x	x
INTWAVE	x	x	x
ISCOPR1D	x	x	x
ISCOPR3D	x	x	x

	2006	2007	2008
MAINSTAT	x	x	x
NACEPR1D	x	x	x
NACE1Y1D	x	x	x
NA11PRS			x
NA11PR1D			x
NA11S			x
NA111D			x
NA111YS	x	x	x
NA111Y1D	x	x	x
NA112JS			x
NA112J1D			x
NEEDCARE	x	x	x
NIGHTWK			x
PRESEEK	x	x	x
REGION	x	x	x
REGION1Y	x	x	x
REGISTER	x	x	x
SATWK			x
SEEKREAS	x	x	
SHIFTWK			
SIGNISAL	x	x	x
STAPROPR	x	x	x
STAPRO1Y	x	x	x
SUNWK			x
SUPVISOR	x	x	
TEMPAGCY	x	x	x
TEMPREAS	x	x	x
WAYJFOUN	x	x	x
WSTAT1Y	x	x	x

<sup>a</sup> Variables with either only system-missings (i.e. blanks) in the raw data or only cases with "No answer" and/or "Not applicable".