

EU-Labour Force Survey December 2011 release

Setup for Importing the Anonymised Yearly Data Sets for 2009

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 2009 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_2009_y.sps

Labels_EULFS_1998-2010.sps

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

- Transformation of EU-LFS 2009 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).¹ 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).² The alphanumerical code from the EU-LFS deviated from the ISO classification with regard to the United Kingdom (UK instead of GB).

In Malta and 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

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.

In Malta, the grouping differs slightly from the other countries (see User Guide 2011, p. 63), however this is documented in the appropriate value labels.

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

¹http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Coding_lists_explanatory_notes_and_classifications/Country_codification_from_2010_onwards.pdf

² 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).

– HWUSTRAI, LENTR4W, PURP4W, NA11S, NA112JS, NA11PRS, NA111YS, NA111Y1D, NA111D, NA112J1D, NA11PR1D

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.

– LEAVTIME, STARTIME

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

– HATFIELD

The variable contains values, which are not mentioned in the User Guide. These were labelled according to the Classification of Fields of Education and Training.³

– 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".

– ISCOPR3D, ISCO3D

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).⁴ Thus they remain unlabelled.

V. Notes on irregularities not mentioned in the User Guide

– AGERESID, 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").

³http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL_LINEAR&StrNom=CL_FL99&StrLanguageCode=EN&Int

⁴http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Coding_lists_explanatory_notes_and_classifications/NUTS_2006_from_2007_onwards.pdf

- COEFF
There are a number of cases with a weighting factor of zero in France, Poland and Spain. There are a lot of system-missings in Finland. The origin of these missings is unclear.
- DEGURBA
In Switzerland all cases have the value 2 for "Intermediate area".
- HHSEQNUM
In Iceland, Sweden and Switzerland all cases have the value 1. In Denmark all cases have either the value 1 or the value 98. There are some system-missing in Germany and Denmark. The origin of these missings is unclear.
- INTWEEK
There are some system-missings in Denmark, France, Italy, Norway, Portugal and Sweden. The origin of these missings is unclear.
- NOWKREAS
There are some system-missings in Iceland and Malta. The origin of these missings is unclear.

VI. Classifications

- ISCED classification (EDUCLEVEL, HATLEVEL):
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

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

Table 1: Number of cases per country in the EU-Labour Force Survey 2009,
yearly data^a

Country	N (without selection)
01 AT Austria	185.242
02 BE Belgium	104.395
03 BG Bulgaria	35.329
04 CH Switzerland	49.390
05 CY Cyprus	41.422
06 CZ Czech Republic	235.119
07 DE Germany	50.035
08 DK Denmark	118.732
09 EE Estonia	19.856
10 ES Spain	113.337
11 FI Finland	59.445
12 FR France	417.131
13 GR Greece	297.460
14 HU Hungary	289.049
15 IE Ireland	273.130
16 IS Iceland	12.370
17 IT Italy	659.561
18 LT Lithuania	66.803
19 LU Luxembourg	19.229
20 LV Latvia	40.424
21 MT Malta	26.401
22 NL Netherlands	93.491
23 NO Norway	20.422
24 PL Poland	212.420
25 PT Portugal	164.964
26 RO Romania	239.426
27 SE Sweden	201.336
28 SI Slovenia	65.919
29 SK Slovak Republic	105.014
30 UK United Kingdom	95.254

^a see also "datafileinfo_2009-2010.csv" on the DVD you received from Eurostat

Table 2: Variables without valid cases in several countries,^a EU LFS 2009, yearly data

	01 AT	02 BE	03 BG	04 CH	05 CY	06 CZ	07 DE	08 DK	09 EE	10 ES	11 FI	12 FR	13 GR	14 HU	15 IE	16 IS	17 IT	18 LT	19 LU	20 LV	21 MT	22 NL	23 NO	24 PL	25 PT	26 RO	27 SE	28 SI	29 SK	30 UK
COUNTR1Y				x												x							x							
COURFILD				x		x		x			x					x	x					x		x			x	x		
COURLEN				x																										
COURPURP				x		x		x														x		x				x		
COURWORH			x	x		x	x	x		x												x		x		x		x		
DEGURBA																x								x						
EDUCFILD				x		x	x															x		x		x		x	x	
EVENWK																										x				
FTPTRAS																														x
HATLFATH				x				x			x					x								x				x		
HATLMOTH				x				x			x					x								x				x		
HATYEAR																								x						
HHAGE14				x				x			x					x								x				x		
HHAGEYG				x				x			x					x								x				x		
HHCHILDR				x				x			x					x								x				x		
HHCOMP				x				x			x					x								x				x		
HHFATH								x			x					x								x				x		
HHINST	x	x	x	x		x	x	x		x	x		x	x	x		x	x	x	x	x	x	x	x	x	x	x	x	x	x
HHLINK				x				x			x					x								x				x		
HHMOTH								x			x					x								x				x		
HHNB0014				x				x			x					x								x				x		
HHNBCH2				x				x			x					x								x				x		
HHNBCH5				x				x			x					x								x				x		
HHNBCH8				x				x			x					x								x				x		
HHNBCH11				x				x			x					x								x				x		
HHNBCH14				x				x			x					x								x				x		
HHNBCH17				x				x			x					x								x				x		
HHNBCH24				x				x			x					x								x				x		
HHNBCHLD				x				x			x					x								x				x		
HHNBEMPL				x				x			x					x								x				x		
HHNBINAC				x				x			x					x								x				x		
HHNBOLD				x				x			x					x								x				x		
HHNBERS				x				x			x					x								x				x		
HHNBUNEM				x				x			x					x								x				x		
HHNBWORK				x				x			x					x								x				x		
HHPARENT				x				x			x					x								x				x		
HHPARTNR				x				x			x					x								x				x		
HHPERS				x				x			x					x								x				x		
HHSPOU								x			x					x								x				x		
HHWKSTAT				x				x			x					x								x				x		
HWOVERP			x	x																										
HWOVERPU			x	x																										
INCDECIL													x		x	x								x	x	x		x		x
INTWAVE							x																							
ISCOPR1D												x																		
ISCOPR3D												x																		
MAINSTAT				x			x																	x						x
NACE1Y1D				x											x															
NACEPR1D																								x						

	01 AT	02 BE	03 BG	04 CH	05 CY	06 CZ	07 DE	08 DK	09 EE	10 ES	11 FI	12 FR	13 GR	14 HU	15 IE	16 IS	17 IT	18 LT	19 LU	20 LV	21 MT	22 NL	23 NO	24 PL	25 PT	26 RO	27 SE	28 SI	29 SK	30 UK	
PRESEEK												x			x																
REGION					x				x							x		x	x	x	x	x									
REGION1Y				x												x							x								
REGISTER															x								x								
SIGNISAL												x ^b																			
STAPR01Y				x											x																
STAPROPR															x																
TEMPAGCY			x																												
TEMPREAS																														x	
WAYJFOUN				x																x											
WSTAT1Y				x											x																

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

^b In the data of France there were only cases in the two categories "Don't know" and "Not applicable", what indicates that the information is not available.