

EU-Labour Force Survey December 2009 release

Setup for importing the Anonymised Yearly Data Sets for 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 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_2008_y.sps

Labels_EULFS_2007-2009.sps

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

- Transformation of EU-LFS 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_2007-2009.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 numbers indicate the region, according to the country codification from 2007.¹ Countries not explicitly assigned to one region by this classification have been classified by GESIS.

The last two digits display the order for the country in alphabetical order as given by the ISO 3166-1-alpha-2 code elements (e.g. Algeria=0501).² The last two digits for countries which have been surveyed remain the same as in the COUNTRY Variable (e.g. Bulgaria=0103). The alphanumerical code from the EU-LFS deviated from the ISO classification in regard to two countries: XS instead of RS for Serbia and UK instead of GB for the United Kingdom.

In Slovenia, the coding of COUNTRYW in the raw data differs from the other countries (see User Guide 2009). 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.

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"
- 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.

¹ http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Coding_lists_explanatory_notes_and_classifications/Country_codification_from_2007_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 2009, p. 63).

– HWUSTRAI, LENTR4W, PURP4W

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

– LEAVTIME, STARTIME

These variables have valid negative values. Thus the missing value category is set to -101 (for "No answer") and -102 (for "Not applicable"). Furthermore there were no cases with the value 999 in the raw data, as mentioned in the User Guide 2009 for "Not applicable". Thus the "Not applicable" category has been derived with help of the filtering variables EXISTPR (for LEAVTIME) and WSTATOR (for STARTIME) (see User Guide 2009, p. 32 f.).

– MARSTAT

In France and Slovenia children under the age 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.

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

– 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)³. Thus they remain unlabelled.

V. Notes on irregularities not mentioned in the User Guide

– AGE

No persons under the age of 15 and over the age of 74 were surveyed in Iceland, Norway and Sweden.

– COEFF

There are a substantive number of cases with a weighting factor of zero in France, Poland and Spain. There are some system-missings in Finland and the UK.

– COURLEN

There is an unusual high frequency of cases with 300 hours in France.

³ 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

- HHSEQNUM
In Iceland and Sweden all cases have the value 1. In Romania one case is system-missing. Its origin is unclear.
- INTWEEK
There are some system-missings in Denmark, France, Norway, Portugal and Sweden. The origin of these missings is unclear.
- NOWKREAS
There are some system-missings in Iceland. The origin of these missings is unclear.

VI. Classifications

- ISCED classification (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
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 2008,
yearly data^a

Country	N (without selection)
01 AT Austria	194.070
02 BE Belgium	105.604
03 BG Bulgaria	34.835
04 CY Cyprus	40.380
05 CZ Czech Republic	241.934
06 DE Germany	48.673
07 DK Denmark	115.508
08 EE Estonia	22.242
09 ES Spain	109.503
10 FI Finland	60.953
11 FR France	348.569
12 GR Greece	293.618
13 HU Hungary	290.431
14 IE Ireland	74.368
15 IS Iceland	12.356
16 IT Italy	671.939
17 LT Lithuania	61.776
18 LU Luxembourg	14.288
19 LV Latvia	38.597
20 NL Netherlands	109.585
21 NO Norway	20.163
22 PL Poland	206.893
23 PT Portugal	167.970
24 RO Romania	240.785
25 SE Sweden	204.042
26 SI Slovenia	65.600
27 SK Slovak Republic	109.840
28 UK United Kingdom	193.630

^a see also "datafileinfo_year_2008.csv" on the CD you received from Eurostat

Table 2: Variables without valid cases in several countries,^a EU LFS 2008, yearly data

	01 AT	02 BE	03 BG	04 CY	05 CZ	06 DE	07 DK	08 EE	09 ES	10 FI	11 FR	12 GR	13 HU	14 IE	15 IS	16 IT	17 LT	18 LU	19 LV	20 NL	21 NO	22 PL	23 PT	24 RO	25 SE	26 SI	27 SK	28 UK
COUNTR1Y															x						x							x
COURFILD					x					x				x	x							x	x		x	x		
COURPURP					x																	x			x			
COURWORH			x		x	x			x													x	x		x			
DEGURBA															x							x						
EDUCLEVL															x													
EVENWK																							x					x
FTPTREAS																										x		x
HATYEAR																						x						
HHSEQNUM							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
HHLINK							x			x					x							x			x			
HHSPOU							x			x					x							x			x			
HHFATH							x			x					x							x			x			
HHMOTH							x			x					x							x			x			
HWOVERP			x																									
HWOVERPU			x																									
INTWAVE						x																						
ISCOPR1D											x																	
ISCOPR3D											x																	
MAINSTAT						x																x						x
NACEPR1D																						x						
NACE1Y1D														x								x						
NA11PRS			x																				x		x	x		
NA11PR1D			x																				x		x	x		
NA11S			x																				x		x	x		
NA111D			x																				x		x	x		
NA111YS			x											x									x		x	x		
NA111Y1D			x											x									x		x	x		
NA112JS			x																				x		x	x		
NA112J1D			x																				x		x	x		
NEEDCARE																												x
NIGHTWK																												x
PRESEEK											x			x														
REGION																						x						
REGIONW																						x						
REGION1Y															x							x	x					x
REGISTER														x								x						
SATWK																												x
SHIFTWK																												x
SIGNISAL											x ^b																	
STAPROPR														x														
STAPRO1Y														x														
SUNWK																												x
TEMPAGCY			x												x													
TEMPREAS																										x		
WSTAT1Y														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.