EU-Labour Force Survey
December 2011 release

Setup for Importing the Anonymised Yearly Data Sets for 1998-2001

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 1998-2001 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_1998-2001_y.sps
Labels_EULFS_1983-2010.sps

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

– Transformation of EU-LFS 1998-2001 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).[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 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" 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", -102 for "Not applicable" and -103 for "Not available" (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".

---

[1] 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
[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).

– MARSTAT

In Bulgaria and 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"), -102 (for "Not applicable") and -103 "Not available".

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).[3] Thus they remain unlabelled.

V. Notes on irregularities not mentioned in the User Guide

– HHSEQNUM

In countries were the final sampling unit are persons[4] and not households, this variable does not contain any useful values.

– HHSEQNUM, HHLINK, INTWEEK, INTWAVE, DEGURBA, METHODA – METHODM

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

---

[3] http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/lfs_main/lfs/lfs_statistical_classifications.htm
[4] 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

– ILOSTAT

In Cyprus in 1999, there are no cases with value 2 for "Unemployed" and no cases with value 3 for "Inactive". Instead, there are a lot of cases with system-missing. The origin of these missings is unclear.

– COEFF

There is a large number of cases with a weighting factor of zero in Poland in 2001.

VI. Classifications

– ISCED classification (EDUCLEVL, 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

Table 1: Number of cases per country in the EU-Labour Force Survey 1998-2001, yearly data[a]

Country	N (without selection)			
	1998	1999	2000	2001
01 AT Austria	61.902	61.618	59.146	59.296
02 BE Belgium	80.066	27.479	27.446	26.391
03 BG Bulgaria	–	–	53.404	59.991
04 CH Switzerland	16.306	17.720	17.733	18.738
05 CY Cyprus	–	10.135	10.301	10.596
06 CZ Czech Republic	70.721	68.824	65.464	63.964
07 DE Germany	–	–	–	–
08 DK Denmark	18.233	17.895	17.993	15.986
09 EE Estonia	16.861	16.361	4.676	5.179
10 ES Spain	190.911	196.532	180.853	173.643
11 FI Finland	17.063	35.071	42.146	42.712
12 FR France	183.072	182.155	182.066	178.143
13 GR Greece	84.007	82.921	81.264	80.282
14 HU Hungary	84.036	87.541	85.080	84.671
15 IE Ireland	111.342	109.768	106.306	105.405
16 IS Iceland	3.755	3.656	3.697	3.685
17 IT Italy	201.835	200.625	199.367	196.236
18 LT Lithuania	7.543	7.558	7.583	7.788
19 LU Luxembourg	17.326	16.095	15.257	14.814
20 LV Latvia	18.756	18.701	19.193	18.834
21 MT Malta	–	–	–	–
22 NL Netherlands	55.959	51.929	66.992	90.264
23 NO Norway	20.664	20.717	21.095	20.213
24 PL Poland	54.545	53.300	46.295	57.575
25 PT Portugal	50.067	47.315	45.626	45.681
26 RO Romania	47.718	46.261	44.805	43.742
27 SE Sweden	15.810	17.909	17.321	50.512
28 SI Slovenia	17.976	19.627	18.751	19.607
29 SK Slovak Republic	31.301	30.442	30.846	30.554
30 UK United Kingdom	144.979	143.058	140.067	135.887

[a] see also "datafileinfo_1998-2001.csv" on the DVD you received from Eurostat