



**Economic and Social Data Service**

## **Quarterly Labour Force Survey, January - March, 2011: Unrestricted Access Teaching Dataset**

---

### **ESDS Government**

## **User Guide**

**Contact:**

ESDS Government

The Cathie Marsh Centre for Census and Survey Research (CCSR)

University of Manchester, Manchester, M13 9PL

Tel: 0161 275 1980 (ESDS Government helpdesk)

Email: [govsurveys@esds.ac.uk](mailto:govsurveys@esds.ac.uk)

## Contents

<b>About this User Guide .....</b>	<b>3</b>
<b>Introduction to the Labour Force Survey.....</b>	<b>3</b>
Main Topics .....	3
The Datasets.....	4
The Sample Design .....	4
Other Useful Sources of Information .....	4
<b>Introduction to the QLFS January-March 2011: Unrestricted Access Teaching Dataset .....</b>	<b>5</b>
Weighting the Dataset .....	5
List of Variables .....	6
Missing Values.....	7
Downloading the Data .....	7
<b>Appendix I: Data Dictionary.....</b>	<b>8</b>
<b>Appendix II: SPSS Syntax for variables derived for the QLFS January-March 2011: Unrestricted Access Teaching Dataset .....</b>	<b>13</b>

## **About this User Guide**

This is a user guide for the Quarterly Labour Force Survey (QLFS), January-March, 2011: Unrestricted Access Teaching Dataset. This teaching dataset has been created for teaching related purposes only such as for teaching basic statistics or for students to write report for coursework. Those wishing to use the Labour Force Survey (LFS) for research purposes are advised to use the original SN6782 Quarterly Labour Force Survey, January-March, 2011. The QLFS January-March 2011: Unrestricted Access Teaching Dataset contains a reduced number of variables from the original QLFS January-March 2011. It gives students the opportunity to gain experience of using nationally representative UK data.

## **Introduction to the Labour Force Survey**

The Labour Force Survey (LFS) is the largest regular social survey in the United Kingdom. It is a unique source of information using international definitions of employment and unemployment and economic inactivity, together with a wide range of related topics such as occupation, training, hours of work and personal characteristics of household members aged 16 years and over living at private addresses in the United Kingdom.

Its purpose is to provide information on the UK labour market that can then be used to develop, manage, evaluate and report on labour market policies. The survey is managed by the Social Surveys division of the Office for National Statistics (ONS) in Great Britain and by the Central Survey Unit of the Department of Finance and Personnel in Northern Ireland. Data from each Quarterly Labour Force survey are anonymised and deposited with the UK Data Archive at the University of Essex and are distributed through the Economic and Social Data Service (ESDS).

## **Main Topics**

The LFS questionnaire comprises a 'core' of questions which are included in every quarter of the survey, together with 'non-core' questions which vary from quarter to quarter. The core questions cover respondent's household, family structure, basic housing information and

demographic details of household members. The 'non-core' questions cover economic activity, education and health. They might also include a few questions asked on behalf of other government departments such as on childcare or regional mobility and might be included in the survey only every 2 or 3 years.

### **The Datasets**

There are a number of different types of LFS datasets available from ESDS such as quarterly dataset, household data, local area data, Eurostat data, two-quarterly and five-quarterly longitudinal datasets. The [ESDS web site](#)<sup>1</sup> contains fuller explanations of each of the datasets. The teaching dataset is based on the January-March 2011 Quarterly LFS, which contains data from that quarter only.

### **The Sample Design**

Since 1992, the LFS has had a single stage, stratified random sample design drawn from the Postcode Address File (PAF). A household, once initially selected for interview, is retained in the sample for a total of five consecutive quarters. The first interview is face-to-face and subsequent ones are by telephone. Interviewers can accept proxy information for household members who are unavailable when the interview takes place.

### **Other Useful Sources of Information**

Further information about the methodology of the LFS can be found in the [LFS user guide – Volume 1 Background and Methodology](#)<sup>2</sup> on the ESDS web site. This contains detailed explanations of the survey design, sampling errors and weighting, among other things.

There is also a series of [ESDS guides](#)<sup>3</sup>, including introductory guides to using the large scale government surveys and to the analysis packages SPSS and Stata.

---

<sup>1</sup> <http://www.esds.ac.uk/findingData/lfsTitles.asp>

<sup>2</sup> [http://www.esds.ac.uk/doc/6782%5Cmrdoc%5Cpdf%5Clfs\\_user\\_guide\\_vol1\\_background2011.pdf](http://www.esds.ac.uk/doc/6782%5Cmrdoc%5Cpdf%5Clfs_user_guide_vol1_background2011.pdf)

<sup>3</sup> <http://www.esds.ac.uk/government/resources/>

Additional information about LFS, including publications, summaries of labour market statistics, guidance and methodology is available on the [ONS webpage](#)<sup>4</sup>.

## **Introduction to the QLFS January-March 2011: Unrestricted Access Teaching**

### **Dataset**

The QLFS January-March 2011: Unrestricted Access Teaching Dataset contains 13 variables. The variables have been chosen to enable users to explore the data and apply some basic statistical techniques such as regression analysis, ANOVA, comparing means and cross-tabs.

The total sample size of the Unrestricted Access Teaching Dataset is 25,162, which is a sub-sample of the original QLFS dataset. It is based on a random selection of 30 per cent of the main QLFS sample, excluding children aged 0-15. The syntax for this is in Appendix II.

Details of the variables selected for the Unrestricted Access Teaching Dataset are given below including a data dictionary which lists information about variable names, values, labels, missing values and frequency. The name of the variables and their labels remain the same or very close to those in the original LFS dataset. However, due to concerns about statistical disclosure, some variables have been recoded and their level of detail reduced for this teaching dataset. Variables that differ from the original LFS have been suffixed with “r”.<sup>5</sup> The syntax for how these variables were recoded is in Appendix II. In addition, a new ID variable was created in order to avoid linking the Unrestricted Access Teaching Dataset with the original dataset.

### **Weighting the Dataset**

Respondents to the survey did not have an equal chance of selection. There were also differences in response rates between groups. Weights were calculated for LFS to correct for this. To enable users to make inferences from this sample to the entire eligible population the person-weight variable ‘pwt10’ should be used.

---

<sup>4</sup> <http://www.ons.gov.uk/ons/taxonomy/index.html?nscl=Labour+Market>

<sup>5</sup> With the exception of “ilodefr” (economic activity –reported) which is the original variable name.

## **List of Variables**

Observations: 25,162

Variables: 13

---

Variable name	Variable label	Variable type
casenew <sup>6</sup>	New random ID number	Scalar
pwt10	Integer weight 2010	Scalar
sex	Sex of respondent	Nominal
ageeulr <sup>7</sup>	Age bands in 5 year intervals – 12 categories	Ordinal
marsta3r <sup>8</sup>	Marital status – 3 categories (recoded)	Nominal
hiqul11d	Highest qualification – 6 categories	Nominal
ethuk6r <sup>9</sup>	Ethnicity –6 categories	Nominal
ilodefr	Economic activity (reported) – 3 categories	Nominal
stat3r <sup>10</sup>	Employment status – 3 categories	Nominal
ftptwk	Full-time or part-time in main job	Nominal
tothrs	Total hours worked in reference week	Scalar
nsecmj3r <sup>11</sup>	NS-SEC 3 (main job) (SOC2010)	Ordinal
govtof2	Government Office Region 2 and 3 combined	Nominal

---

<sup>6</sup> New computed variable

<sup>7</sup> Derived from “ageeul” (28 categories)

<sup>8</sup> Derived from “marsta” and “mardy6”

<sup>9</sup> Derived from “ethuk8” (8 categories)

<sup>10</sup> Derived from “stat” (4 categories)

<sup>11</sup> Derived from “nsecmj10” (8 categories)

### **Missing Values**

Missing values occur when a respondent does not answer a particular question. This may happen for a number of reasons, including refusal or inability to answer a question, or where it does not apply to them. It is common practice to exclude missing values from analyses, although they may occasionally be of interest. Both SPSS and Stata have different conventions for missing values:

	<b>SPSS dataset</b>	<b>Stata dataset</b>
No answer	-8	.a
Does not apply	-9	.b
System missing	.	.

### **Downloading the Data**

Unlike other government surveys the QLFS January-March 2011: Unrestricted Access Teaching Dataset does not require any registration and can be downloaded via the ESDS website:

<http://www.ccsr.ac.uk/esds/data/>.

The dataset is available in two formats: SPSS and STATA.

## Appendix I: Data Dictionary

---

casenew New random ID number

---

type: numeric (double)

range: [3,30000]                      units: 1  
unique values: 17055                      missing .: 0/25162

mean: 15125  
std. dev: 8659.63

percentiles:	10%	25%	50%	75%	90%
	3008	7646	15138.5	22636	27045

---

pwt10 Integer weight 2010

---

type: numeric (double)  
label: pwt10, but 1297 nonmissing values are not labeled

range: [0,11787]                      units: 1  
unique values: 1297                      missing .: 0/25162

examples: 462  
          533  
          595  
          691

---

sex Sex

---

type: numeric (double)  
label: sex

range: [1,2]                              units: 1  
unique values: 2                              missing .: 0/25162

tabulation:	Freq.	Numeric	Label
	12181	1	Male
	12981	2	Female



---

ageeulr

Age bands in 5 year intervals (recoded)

---

type: numeric (double)  
label: ageeulr

range: [1,12]                      units: 1  
unique values: 12                    missing .: 0/25162

tabulation:	Freq.	Numeric	Label
	1604	1	16-19
	1931	2	20-24
	1980	3	25-29
	2002	4	30-34
	2132	5	35-39
	2355	6	40-44
	2371	7	45-49
	2227	8	50-54
	2032	9	55-59
	2145	10	60-64
	1761	11	65-69
	2622	12	70 and over

---

marsta3r

Marital status (recoded)

---

type: numeric (double)  
label: marsta3r

range: [1,3]                      units: 1  
unique values: 3                    missing .: 0/25162

tabulation:	Freq.	Numeric	Label
	8240	1	Single, never married
	13643	2	Married/ Cohabiting/ Civil Partner
	3279	3	Divorced/ Widowed/ Previously in Civil Partnership

---

---

hiqul1d

Highest qualification (detailed grouping)

---

---

type: numeric (double)  
label: hiqul1d

range: [1,7]                      units: 1  
unique values: 7                    missing .: 0/25162  
unique mv codes: 2                  missing .\*: 4856/25162

tabulation:	Freq.	Numeric	Label
	4488	1	Degree or equivalent
	1769	2	Higher education
	4660	3	GCE A Level or equiv
	4686	4	GCSE grades A-C or equiv
	2218	5	Other qualifications
	2332	6	No qualification
	153	7	D/K
	2411	.a	
	2445	.b	

---

---

ethuk6r

Ethnicity UK 6 cat. (recoded)

---

---

type: numeric (double)  
label: ethuk6r

range: [1,6]                      units: 1  
unique values: 6                    missing .: 1240/25162  
unique mv codes: 2                  missing .\*: 128/25162

tabulation:	Freq.	Numeric	Label
	21609	1	White
	173	2	Mixed/Multiple ethnic groups
	470	3	Indian
	414	4	Pakistani and Bangladeshi
	529	5	Black/African/Caribbean/Black British
	599	6	Other ethnic group (includes Northern Irish Gypsy, Traveller or Irish Traveller)
	1240	.	
	128	.a	

---

ilodefr Economic activity (reported)

---

type: numeric (double)  
label: ilodefr

range: [1,3]                      units: 1  
unique values: 3                    missing .: 0/25162  
unique mv codes: 1                missing .\*: 1240/25162

tabulation:	Freq.	Numeric	Label
	14182	1	In employment
	1153	2	ILO unemployed
	8587	3	Inactive
	1240	.a	

---

stat3r Employment status (recoded)

---

type: numeric (double)  
label: stat3r

range: [1,3]                      units: 1  
unique values: 3                    missing .: 0/25162  
unique mv codes: 1                missing .\*: 7021/25162

tabulation:	Freq.	Numeric	Label
	15659	1	Employee
	2411	2	Self-employed
	71	3	Government scheme or unpaid family worker
	7021	.b	

---

ftptwk Full-time or part-time in main job

---

type: numeric (double)  
label: ftptwk

range: [1,2]                      units: 1  
unique values: 2                    missing .: 0/25162  
unique mv codes: 2                missing .\*: 7038/25162

tabulation:	Freq.	Numeric	Label
	12606	1	Full-time
	5518	2	Part-time
	12	.a	
	7026	.b	

---

tothrs Total hours worked in reference week

---

type: numeric (double)  
label: tothrs, but 92 nonmissing values are not labeled

range: [0,97] units: 1  
unique values: 93 missing .: 0/25162  
unique mv codes: 2 missing .\*: 11247/25162

examples: 30  
40  
.b

---

nsecmj3r NS-SEC 3 class (main job, SOC2010, recoded)

---

type: numeric (double)  
label: nsecmj3r

range: [1,4] units: 1  
unique values: 4 missing .: 0/25162  
unique mv codes: 1 missing .\*: 7045/25162

tabulation:	Freq.	Numeric	Label
	4200	1	Higher managerial, administrative and professional occupations
	2939	2	Intermediate occupations
	4412	3	Routine and manual occupations
	6566	4	Never worked, unemployed, and nec
	7045	.b	

---

govtof2 Government Office Region 2 and 3 combined

---

type: numeric (double)  
label: govtof2

range: [1,13] units: 1  
unique values: 12 missing .: 0/25162

tabulation:	Freq.	Numeric	Label
	1170	1	North East
	2823	2	North West (inc Merseyside)
	2240	4	Yorkshire and the Humber
	1901	5	East Midlands
	2181	6	West Midlands
	2363	7	Eastern
	2704	8	London
	3288	9	South East
	2148	10	South West
	1211	11	Wales
	2122	12	Scotland
	1011	13	Northern Ireland

**Appendix II: SPSS Syntax for variables derived for the QLFS January-March 2011:  
Unrestricted Access Teaching Dataset**

**Compute new ethnicity variable: recode ethuk8 into ethuk6**

```
RECODE ethuk8 (-8=-8) (-9=-9) (1=1) (2=2) (3=3) (6=6) (7=5) (8=6) (4  
thru 5=4) INTO ethuk6r.  
VARIABLE LABELS ethuk6r 'Ethnicity UK 6 cat. (recoded)'.  
VALUE LABELS ethuk6r -8 'No answer' -9'Does not apply' 1'White'  
2'Mixed/Multiple ethnic groups' 3'Indian' 4'Pakistani and Bangladeshi'  
5'Black/African/Caribbean/Black British' 6 'Other ethnic group  
(includes Northern Irish Gypsy, Traveller or Irish Traveller)'.  
EXECUTE.
```

**Compute new marital status variable: compute “marsta3r” from “marsta”  
(Marital status) and “mardy” (Married/co-habiting/civil partners)**

```
COMPUTE marsta3r=0.  
IF (marsta=1) marsta3r=1.  
IF (marsta=2) marsta3r=2.  
IF (marsta=3) marsta3r=2.  
IF (marsta=4) marsta3r=3.  
IF (marsta=5) marsta3r=3.  
IF (marsta=6 & mardy6=2) marsta3r=3.  
IF (marsta=6 & mardy6=1) marsta3r=2.  
VARIABLE LABELS marsta3r 'Marital status (recoded)'.  
VALUE LABELS marsta3r 1'Single, never married' 2 'Married/ Cohabiting/  
Civil Partner' 3 'Divorced/ Widowed/ Previously in Civil Partnership'.  
EXECUTE.
```

**Compute new NS-SEC 3 variable: recode secmj10 into nsecmj3r**

```
RECODE nsecmj10 (-9=-9) (-8=-8) (8=4) (1 thru 2=1) (3 thru 4=2) (5 thru  
7=3) INTO nsecmj3r.  
VARIABLE LABELS nsecmj3r'NS-SEC 3 class (main job, SOC2010, recoded)'.  
EXECUTE.
```

```
VALUE LABELS nsecmj3r -8 'No answer' -9'Does not apply' 1'Higher
managerial, administrative and professional occupations' 2'Intermediate
occupations' 3'Routine and manual occupations' 4'Never worked,
unemployed, and nec'.
```

```
EXECUTE.
```

**Compute new employment status variable: recode stat into stat3r**

```
RECODE stat (-9=-9) (-8=-8) (1=1) (2=2) (3 thru 4=3) INTO stat3r.
```

```
VARIABLE LABELS stat3r 'Employment status recoded'.
```

```
VALUE LABELS stat3r -8 'No answer' -9'Does not apply' 1'Employee'
2'Self-employed' 3'Government scheme or unpaid family worker'.
```

```
EXECUTE.
```

**Select only cases age 16+ and delete unselected cases i.e. ages 0-15.**

```
FILTER OFF.
```

```
USE ALL.
```

```
SELECT IF (ageeul>=5).
```

```
EXECUTE.
```

**Compute new age variable: recode ageeul (Age band 28 cat.) into ageeulr (Age band 12 cat.)**

```
RECODE ageeul (14=3) (15=4) (16=5) (17=6) (18=7) (19=8) (20=9) (21=10)
(22=11) (5 thru 8=1) (9 thru 13=2) (23 thru 28=12) INTO ageeulr.
```

```
VARIABLE LABELS ageeulr ' Age bands in 5 year intervals (recoded)'.
```

```
VALUE LABELS ageeulr 1 '16-19' 2 '20-24' 3 '25-29' 4 '30-34' 5 '35-39'
6 '40-44' 7 '45-49' 8 '50-54' 9 '55-59' 10 '60-64' 11 '65-69' 12 '70
and over'.
```

```
EXECUTE.
```

**Reduce the sample size: select only 30% of the cases and delete unselected cases**

```
FILTER OFF.
```

```
USE ALL.
```

```
SAMPLE .30.
```

```
EXECUTE.
```

**Compute a new ID variable with random numbers (from 0-15000) and then sort dataset by the new ID variable ascending.**

```
COMPUTE casenew=TRUNC (UNIFORM(30000)) +1.  
VARIABLE LABELS casenew 'New random ID number'.  
EXECUTE.  
SORT CASES BY casenew (A).  
EXECUTE.
```