

UK Data Service

---



# Crime Survey for England and Wales 2013- 2014: Unrestricted Access Teaching Dataset

## User guide

Author: Jennifer Buckley and Margherita Ceraolo

Version: 1.0

Date: 16 June 2016

## Contents

Background to the Crime Survey for England and Wales (CSEW).....	3
Crime Survey for England and Wales (CSEW) 2013-2014 Unrestricted Access Teaching Dataset .....	4
List of variables in the CSEW 2013-2014 Unrestricted Access Teaching Dataset .....	6
Codebook .....	7
Code for additional variables .....	17

## **Background to the Crime Survey for England and Wales (CSEW)**

The Crime Survey for England and Wales (CSEW) is a face-to-face victimisation survey in which people resident in households in England and Wales are asked about their experiences of a range of crimes in the 12 months prior to the interview. Respondents to the survey are also asked about their perceptions of crime and attitudes towards crime related issues such as the police and criminal justice system.

Previously known as the *British Crime Survey* (BCS), the survey was first conducted in 1982. In 2001, the then BCS, moved to an annual format with continuous sampling. The survey now only covers England and Wales with separate surveys of Scotland and Northern Ireland. However, the first and third surveys were carried out in England, Wales and Scotland (hence 'British' Crime Survey).

The CSEW uses the Postcode Address File (PAF) to select a representative sample of households in England and Wales. The CSEW does not cover the population living in group residences (such as halls of residence and prisons) or other institutions, nor does it cover crime against commercial or public sector bodies.

The core sample size has increased over the years from around 11,000 in the earlier cycles to 35,371 in the 2013/14 CSEW.

### **Fieldwork**

At each sampled address the interviewer establishes that the address is eligible; ineligible addresses include vacant properties, second homes, non-residential addresses and establishments where people are living in group residences. If one PAF address leads to two households, the interviewer randomly selects which household to approach. Once the household is determined to be eligible, individuals aged 16 or over in the selected household are listed and one adult is randomly selected for interview. No substitutes are permitted.

Face-to-face interviews are carried out using computer-assisted personal interviewing (CAPI) where interviewers record responses to the questionnaire on tablet computers. Self-completion modules are used in the CSEW to collect information on topic areas that respondents could feel uncomfortable talking about to an interviewer.

### **The questionnaire**

The CSEW questionnaire consists of question modules such as victimization, anti-social behavior and demographic characteristics of the respondent and household.

Within modules there can also be further filtering with so that some questions are only asked of smaller sub-samples. Respondents are randomly allocated into one of four sub-samples, A, B, C or D which each represent around a quarter of the overall sample.

Further information about CSEW can be found on the [Office for National Statistics' Crime Survey for England and Wales webpage.](#)

## Crime Survey for England and Wales (CSEW) 2013-2014 Unrestricted Access Teaching Dataset

This teaching dataset is based on the [Crime Survey for England and Wales \(CSEW\) 2013-2014](#). It contains data for 8,843 cases selected at random from the CSEW 2013-14 (adult non-victim form dataset), which is 25 percent of the original 35,371<sup>1</sup>.

### Variables

The dataset contains 32 variables, covering the following topics:

- demographic details
- perceptions of crime module
- experience of crime
- antisocial behavior

All the variables within the dataset are individual level variables and require individual based analysis. There is a mix of discrete and continuous variables. A full list of variable names and labels is on page 6 and variable frequencies are provided from page 7.

Most of the variables come directly from the CSEW 2013-14 dataset deposited at the UK Data archive. The documentation for the CSEW 2013-2014 includes a copy of the [questionnaire](#).

### *Derived variables*

The dataset contains new scalar variables for teaching and learning. They each provide measures of key concepts and have been derived from multiple variables from the CSEW 2013-14 (using principal components analysis). For example, *worryx* measures worry about crime and comes from five variables relating to worries about a range of crimes: *wburgl wmugged wraped wattack wraceatt*. For this variable, a higher score indicate a higher level of worry. The teaching dataset contains all the original variables used to create the new scalar variables (and as a result, users should note that the scalar variables should not be used as if they are independent of the variables used to create them).

The dataset also contains the following changes from the original CSEW dataset.

1. The variables measuring the deprivation of respondents' local area are deprivation quintiles, which rank areas into five groups indicating the most to least deprived. These have been calculated from deprivation deciles in the main CSEW 2013-2014 dataset.
2. The original CSEW weighting variable for individual level analyses (*c11cdivwgt*), which adjusts for unequal probabilities and non-response, has been scaled to have a mean of 1.

All variables created specifically for this dataset are suffixed with an x and the Stata script used to create the extra variables from the original crime survey can be found on page 17.

---

<sup>1</sup> A sub-sample was selected in order to make the data available under an Open Government Licence. The sub-sample was taken a random and without replacement using the `sample` command in Stata.

### **How to obtain the CSEW 2013-2014 Unrestricted Access Teaching Dataset**

The CSEW 2013-2014 Unrestricted Access Teaching Dataset is open access, and can be simply downloaded from the [UK Data Service](#). The Teaching Dataset is available in two formats: SPSS and Stata.

### **Weighting the dataset**

The Teaching Dataset contains an individual level weight called *IndivWgtx*. Weights adjust for unequal selection probabilities and non-response and users of the data should use the appropriate weight.

### **Missing values within the dataset**

#### *Don't know and refusal*

When asked a question, survey respondents may respond 'do not know' or refuse to answer. Such responses are recorded using specific codes; for example, don't know is often coded as 9. It is useful to get variable frequencies first in any analysis to examine the distribution of responses and the proportion of 'don't know' and 'refusal' responses. In most analysis, 'refusal' codes are excluded. 'Don't know' codes are also usually excluded unless there is interest in 'don't know' responses such as in the case of attitudinal questions. In the SPSS version of this teaching dataset do not know responses and refusals are set as missing values.

#### *System missing*

Variables within the dataset can contain empty cells where no data is recorded. In SPSS empty cell are automatically coded as 'system missing' which is denoted by a dot (.). In STATA 'system missing' data are also denoted by a dot (.). For some variables in the dataset, only a sub-set of respondents were asked the question due to the modular design of the survey (see the discussion of the Questionnaire above). Where the question was not asked, a system missing response is recorded and therefore some variables have a large number of system missing responses.

### **Notes for teachers**

This dataset has been designed for teaching purposes only. As the data comes from [SN 7616 Crime Survey for England and Wales 2013-2014](#) weighted analyses can be taken to be representative of England and Wales as a whole. Students may therefore use these results to discuss crime-related statistics and attitudes to the Criminal Justice System at the national level in their reports while they learn about statistics, statistics packages and large-scale national surveys.

Please use the original for all non- teaching purposes. The full dataset is available for download (after registration) from the UK Data Service website: [www.ukdataservice.ac.uk](http://www.ukdataservice.ac.uk).

The Teaching Dataset is available under the [Open Government Licence](#). For more information about making it available to your students, see the pages about using teaching data with your class here: <http://ukdataservice.ac.uk/use-data/teaching/practical-resources.aspx>

## List of variables in the CSEW 2013-2014 Unrestricted Access Teaching Dataset

Variable names and labels correspond to the main CSEW 2013-2014 dataset, where applicable. Variables derived for this teaching dataset are suffixed with an 'x'.

For variables from question modules asked to only specific sub-samples, the sub-samples are indicated in brackets after the variable label.

Variable name	Variable label
<b>Background, demographic and weighting variables</b>	
rowlabel	Case identifier (9 digits)
split	Follow-up module split
sex	Adult number 1 (respondent): Sex
ysarea	How long lived in this area
resyrago	Living at this address 12 months ago or not
work2	Any paid work in last week
tenure1	In which way do you occupy this accommodation
livharm1	ONS harmonised marital status
agegrp7	Age group (7 bands)
ethgrp2a	Ethnic Group (5 categories)
educat3	Respondent education (5 categories)
rural2	Type of area 2004: urban/rural
edeprivex	England: Index of multiple deprivation by quintile (1=20% most deprived wards)
wdeprivex	Wales: Index of multiple deprivation by quintile (1=20% most deprived wards)
IndivWgtx	Individual-level weight (mean=1)
<b>Experience and perceptions of crime</b>	
Cause2m	One MAIN cause of crime in Britain today (Module D)
walkdark	How safe do you feel walking alone after dark (Module D)
walkday	How safe do you feel walking alone in this area during the day (Module D)
homealon	How safe do you feel when alone in home at night (Module D)
wburgl	How worried about having your home broken into (Module C)
wmugged	How worried about being mugged and robbed (Module C)
wcarstol	How worried about having car stolen (Module B)
wfromcar	How worried about having things stolen from your car (Module B)
wraped	How worried about being raped (Module C)
wattack	How worried about being physically attacked by strangers (Module C)
wraceatt	How worried about being attacked because of skin colour, ethnic origin or religion (Module C)
worryx	Worry about being a victim of crime (high score = high level of worry) (Module C) [Derived from wburgl wmugged wraped wattack wraceatt]
bcsvictim	Experience of any crime in the previous 12 months
<b>Anti-social behaviour</b>	
rubbcomm	How common is litter or rubbish in immediate area
vandcomm	How common is vandalism or graffiti in immediate area
poorhou	How common are homes in poor condition/run down
antisocx	Anti-social behaviour in their neighbourhood (high score = high levels of anti-social behaviour)

## Codebook

-----  
rowlabel

Case identifier (9 digits)

-----  
type: numeric (double)  
range: [1.352e+08,1.476e+08] units: 10  
unique values: 8,843 missing .: 0/8,843  
mean: 1.4e+08  
std. dev: 4.4e+06  
percentiles: 10% 25% 50% 75% 90%  
1.4e+08 1.4e+08 1.4e+08 1.4e+08 1.5e+08

-----  
split

Follow-up module split

-----  
type: numeric (double)  
label: split  
range: [1,4] units: 1  
unique values: 4 missing .: 0/8,843  
tabulation: Freq. Numeric Label  
2,308 1 A (Experience of the poilice)  
2,267 2 B (Attitudes to the CJS)  
2,194 3 C (Crime preventing)  
2,074 4 D (Online security)

-----  
sex

Adult number 1 (respondent): Sex

-----  
type: numeric (double)  
label: sex  
range: [1,2] units: 1  
unique values: 2 missing .: 0/8,843  
tabulation: Freq. Numeric Label  
4,037 1 Male  
4,806 2 Female

-----  
yrsarea  
How long lived in this area?  
-----

type: numeric (double)  
label: yrsarea

range: [1,9] units: 1  
unique values: 8 missing .: 0/8,843

tabulation: Freq. Numeric Label

502	1	Less than a month
484	2	12 months but less than 2 years
451	3	2 years but less than 3 years
597	4	3 years but less than 5 years
1,225	5	5 years but less than 10 years
1,686	6	10 years but less than 20 years
3,897	7	20 years or longer
1	9	Don't know

-----  
resyrago  
Living at this address 12 months ago or not?  
-----

type: numeric (double)  
label: resyrago

range: [1,2] units: 1  
unique values: 2 missing .: 7,334/8,843

tabulation: Freq. Numeric Label

652	1	Yes
857	2	No
7,334	.	

-----  
work2  
Any paid work in last week?  
-----

type: numeric (double)  
label: work2

range: [1,8] units: 1  
unique values: 3 missing .: 0/8,843

tabulation: Freq. Numeric Label

4,703	1	Yes
4,138	2	No
2	8	Refusal



-----  
tenure1 In  
which way do you occupy this accommodation?  
-----

type: numeric (double)  
label: tenure1  
  
range: [1,9] units: 1  
unique values: 8 missing .: 0/8,843

tabulation:	Freq.	Numeric	Label
	3,059	1	Own it outright
	2,515	2	Buying it with help of mortgage/loan
	50	3	Pay part rent part mortgage
	2,914	4	Rent it
	281	5	Live here rent free
	1	6	Squatting
	11	8	Refusal
	12	9	Don't know

-----  
livharm1  
Respondent marital status  
-----

type: numeric (double)  
label: livharm1  
  
range: [-1,6] units: 1  
unique values: 7 missing .: 0/8,843

tabulation:	Freq.	Numeric	Label
	13	-1	Not classified
	3,931	1	Married
	893	2	Cohabiting
	1,988	3	Single
	314	4	Separate
	800	5	Divorced
	904	6	Widowed

-----  
agegrp7  
Age group (7 bands)  
-----

type: numeric (double)  
label: agegrp7  
  
range: [1,7] units: 1  
unique values: 7 missing .: 0/8,843

tabulation:	Freq.	Numeric	Label
	685	1	16-24
	1,359	2	25-34
	1,442	3	35-44
	1,515	4	45-54
	1,428	5	55-64
	1,315	6	65-74
	1,099	7	75+

-----  
ethgrp2a  
Ethnic Group (5 categories)  
-----

type: numeric (double)  
label: ethgrp2a

range: [1,5] units: 1  
unique values: 5 missing .: 10/8,843

tabulation: Freq. Numeric Label  
7,954 1 White  
88 2 Mixed  
403 3 Asian or Asian British  
288 4 Black or Black British  
100 5 Chinese  
10 .

-----  
educat3  
Respondent education (5 categories)  
-----

type: numeric (double)  
label: educat3

range: [1,5] units: 1  
unique values: 5 missing .: 21/8,843

tabulation: Freq. Numeric Label  
1,818 1 None  
1,751 2 O level/GCSE  
1,592 3 Apprenticeship or A/AS level  
3,287 4 Degree or diploma  
374 5 Other  
21 .

-----  
rural2  
Type of area 2004: urban/rural  
-----

type: numeric (double)  
label: rural2

range: [1,2] units: 1  
unique values: 2 missing .: 0/8,843

tabulation: Freq. Numeric Label  
6,755 1 Urban  
2,088 2 Rural

-----  
edeprivex  
England: Index of multiple deprivation by quintile (1=20% most deprived wards)  
-----

type: numeric (float)

range: [1,5] units: 1  
unique values: 5 missing .: 703/8,843

tabulation: Freq. Value  
1,521 1  
1,595 2  
1,689 3  
1,665 4  
1,670 5  
703 .

wdeprivex  
Wales: Index of multiple deprivation by quintile (1=20% most deprived wards)

-----  
type: numeric (float)  
range: [1,5] units: 1  
unique values: 5 missing .: 8,140/8,843  
tabulation: Freq. Value  
119 1  
130 2  
164 3  
172 4  
118 5  
8,140 .

-----  
IndivWgtx  
Individual-level weight (mean=1)

-----  
type: numeric (float)  
range: [.21917011,5.1739893] units: 1.000e-08  
unique values: 8,822 missing .: 0/8,843  
mean: .99572  
std. dev: .624757  
percentiles: 10% 25% 50% 75% 90%  
.41893 .572498 .821092 1.23573 1.70401

-----  
cause2m  
One MAIN cause of crime in Britain today?

-----  
type: numeric (double)  
label: cause2m  
range: [1,99] units: 1  
unique values: 14 missing .: 6,769/8,843  
examples: 9 I. Too few police  
.

-----  
walkdark  
How safe do you feel walking alone after dark?

-----  
type: numeric (double)  
label: walkdark  
range: [1,9] units: 1  
unique values: 5 missing .: 6,769/8,843  
tabulation: Freq. Numeric Label  
624 1 Very safe  
824 2 Fairly safe  
411 3 A bit unsafe  
198 4 Very unsafe  
17 9 Don't know  
6,769 .

-----  
walkday

How safe do you feel walking alone in this area during the day?

---

```
      type: numeric (double)
      label: walkday

      range: [1,9]                units: 1
unique values: 5                  missing .: 6,769/8,843

      tabulation: Freq.  Numeric  Label
                  1,573      1    Very safe
                   438      2    Fairly safe
                    50      3    A bit unsafe
                    10      4    Very unsafe
                     3      9    Don't know
                   6,769      .
```

---

homealon

How safe do you feel when alone in home at night?

---

```
      type: numeric (double)
      label: homealon

      range: [1,9]                units: 1
unique values: 5                  missing .: 6,769/8,843

      tabulation: Freq.  Numeric  Label
                  1,358      1    Very safe
                   578      2    Fairly safe
                   109      3    A bit unsafe
                    27      4    Very unsafe
                     2      9    Don't know
                   6,769      .
```

---

wburgl

How worried about having your home broken into?

---

```
      type: numeric (double)
      label: wburgl

      range: [1,9]                units: 1
unique values: 6                  missing .: 6,649/8,843

      tabulation: Freq.  Numeric  Label
                  225      1    Very worried
                   595      2    Fairly worried
                  1,040      3    Not very worried
                   332      4    Not at all worried
                     1      5    (Not applicable)
                     1      9    Don't know
                   6,649      .
```

-----  
wmugged  
How worried about being mugged and robbed?  
-----

```
      type: numeric (double)
      label: wmugged

      range: [1,9]                units: 1
unique values: 6                missing .: 6,649/8,843

      tabulation: Freq.  Numeric  Label
                  167      1      Very worried
                  412      2      Fairly worried
                 1,088      3      Not very worried
                  514      4      Not at all worried
                   4      5      (Not applicable)
                   9      9      Don't know
                 6,649      .
```

-----  
wcarstol  
How worried about having car stolen?  
-----

```
      type: numeric (double)
      label: wcarstol

      range: [1,5]                units: 1
unique values: 5                missing .: 7,080/8,843

      tabulation: Freq.  Numeric  Label
                  98      1      Very worried
                 284      2      Fairly worried
                 896      3      Not very worried
                 455      4      Not at all worried
                  30      5      (Not applicable)
                 7,080      .
```

-----  
wfromcar  
How worried about having things stolen from your car?  
-----

```
      type: numeric (double)
      label: wfromcar

      range: [1,9]                units: 1
unique values: 5                missing .: 7,110/8,843

      tabulation: Freq.  Numeric  Label
                  93      1      Very worried
                 347      2      Fairly worried
                 874      3      Not very worried
                 418      4      Not at all worried
                   1      9      Don't know
                 7,110      .
```

-----  
wrape

How worried about being raped?  
-----

```
      type: numeric (double)
      label: wrape

      range: [1,9]                units: 1
unique values: 6                missing .: 6,649/8,843

      tabulation: Freq.  Numeric  Label
                  152      1      Very worried
                  179      2      Fairly worried
                  691      3      Not very worried
                  1,083    4      Not at all worried
                   78      5      (Not applicable)
                   11      9      Don't know
                 6,649      .
```

-----  
wata

How worried about being physically attacked by strangers?  
-----

```
      type: numeric (double)
      label: wata

      range: [1,9]                units: 1
unique values: 6                missing .: 6,649/8,843

      tabulation: Freq.  Numeric  Label
                  175      1      Very worried
                  394      2      Fairly worried
                 1,033    3      Not very worried
                   581    4      Not at all worried
                    2      5      (Not applicable)
                    9      9      Don't know
                 6,649      .
```

-----  
wra

How worried about being attacked because of skin colour, ethnic origin or religion?  
-----

```
      type: numeric (double)
      label: wra

      range: [1,9]                units: 1
unique values: 6                missing .: 6,649/8,843

      tabulation: Freq.  Numeric  Label
                   78      1      Very worried
                   132     2      Fairly worried
                   611     3      Not very worried
                 1,296    4      Not at all worried
                    67     5      (Not applicable)
                    10     9      Don't know
                 6,649      .
```

-----  
worryx

Worry about being a victim of crime (high score = high level of worry)

-----  
type: numeric (float)  
range: [-2.9023592,1.3885418] units: 1.000e-11  
unique values: 294 missing .: 6,796/8,843  
mean: .023542  
std. dev: .966015  
percentiles: 10% 25% 50% 75% 90%  
-1.2253 -.360114 .226307 .770854 1.1319  
-----

bcsvictim

Experience of any crime in the previous 12 months?

-----  
type: numeric (double)  
label: bcsvicti  
range: [0,1] units: 1  
unique values: 2 missing .: 0/8,843  
tabulation: Freq. Numeric Label  
7,460 0 Not a victim of crime  
1,383 1 Victim of crime  
-----

rubbcomm

How common is litter or rubbish in immediate area?

-----  
type: numeric (double)  
label: rubbcomm  
range: [1,5] units: 1  
unique values: 5 missing .: 0/8,843  
tabulation: Freq. Numeric Label  
103 1 Very common  
786 2 Fairly common  
3,258 3 Not very common  
4,682 4 Not at all common  
14 5 Not coded  
-----

vandcomm

How common is vandalism or graffiti in immediate area?

-----  
type: numeric (double)  
label: vandcomm  
range: [1,5] units: 1  
unique values: 5 missing .: 0/8,843  
tabulation: Freq. Numeric Label  
36 1 Very common  
202 2 Fairly common  
2,434 3 Not very common  
6,156 4 Not at all common  
15 5 Not coded  
-----

-----  
poorhou  
How common are homes in poor co conditions/run down?  
-----

type: numeric (double)  
label: poorhou  
  
range: [1,5] units: 1  
unique values: 5 missing .: 0/8,843  
  
tabulation: Freq. Numeric Label  
48 1 Very common  
438 2 Fairly common  
3,000 3 Not very common  
5,331 4 Not at all common  
26 5 Not coded

-----  
antisocx  
Anti-social behaviour in their neighbourhood (high score = high levels of anti-  
social behavior)  
-----

type: numeric (float)  
  
range: [-4.0145574,1.2152667] units: 1.000e-12  
unique values: 738 missing .: 6,694/8,843  
  
mean: .007498  
std. dev: .991067  
  
percentiles: 10% 25% 50% 75% 90%  
-1.42235 -.528008 .184597 .788219 1.21527



## Code for additional variables

Extract from STATA do. file used to create additional variables

```
*** Revise individual and household weights, mean=1 and drop the original variables
gen IndivWgtx=c11Indiv/1280.213
label variable IndivWgtx "Individual-level weight (mean=1)"
drop c11Indiv
gen HhdWgtx= c11HhdWg/674.91
label variable HhdWgtx "Household-level weight (mean=1)"
drop c11HhdWg

** Create indices of multiple deprivation in England and Wales by quintile and drop
the original variables
gen edeprivex=emdidec3
recode edeprivex 1/2=1 3/4=2 5/6=3 7/8=4 9/10=5
label variable edeprivex "England: Index of multiple deprivation by quintile (1=20%
most deprived wards)"
gen wdeprivex=wmdidec4
recode wdeprivex 1/2=1 3/4=2 5/6=3 7/8=4 9/10=5
label variable wdeprivex "Wales: Index of multiple deprivation by quintile (1=20%
most deprived wards)"
drop emdidec3
drop wmdidec4

*** Create new scalar variables using pca
**Worries about crime - use variables from Module C
codebook wburgl wmugged wraped wattack wraceatt
mvdecode wburgl wmugged wraped wattack wraceatt, mv(5=.c)
factor wburgl wmugged wraped wattack wraceatt, pcf
predict worryx
label var worryx "Worry about being a victim of crime (high score = high level of
worry)"

**Anti-social behaviour in their neighbourhood
codebook noisneig teenhang rubbish vandals druguse drunk abancar
factor noisneig teenhang rubbish vandals druguse drunk abancar, pcf
rotate
predict antisocx
label var antisocx "Anti-social behaviour in their neighbourhood (high score = high
levels of anti- social behaviour)"

**Effectiveness of Criminal Justice System
codebook cjspolb cjscpsb cjscrt2a cjscrt2b cjspslb cjsp2b
factor cjspolb cjscpsb cjscrt2a cjscrt2b cjspslb cjsp2b, pcf
predict effectx
label var effectx "Effectiveness of Criminal Justice System (high score= high
opinion)"

**Fairness of Criminal Justice System
codebook fairatt1-fairatt7
factor fairatt1-fairatt7, pcf
predict fairx
label var fairx "Fairness of Criminal Justice System (high score=high opinion)"
```

```
**Confidence in police in their neighbourhood
codebook polatt1-polatt7
factor polatt1-polatt7, pcf
predict confx
label var confx "Confidence in police in their neighbourhood (high score=high level
of confidence)"
```

```
** Reverse and label scalar variables so that high scores=high levels of
worry/confidence/opinions replace worryx=worryx*(-1)
replace antisocx=antisocx*(-1)
replace effectx=effectx*(-1)
replace fairx=fairx*(-1)
replace confx=confx*(-1)
```