

# Malawi - CRS Cleaned residency episodes

**Malawi Epidemiology and Intervention Research Unit**

Report generated on: November 15, 2025

Visit our data catalog at: <http://192.168.0.96/microdataportal/index.php>

## Identification

### SURVEY ID NUMBER

MWI-MEIRU-CRSresidency-2002-v1

### TITLE

CRS Cleaned residency episodes

### COUNTRY

| Name   | Country code |
|--------|--------------|
| Malawi | MWI          |

### STUDY TYPE

Demographic Surveillance

### ABSTRACT

crs\_residency\_episodes\_clean datasets consist of episodes (with a start and a stop date) when people were resident in the CRS. Each individual may have multiple records, and records should be unique on ident, crshse and startdate.

The cleaning process first cleans data errors which result in overlapping episodes in the same house (e.g. when the dates of someone's episode overlap with another of their episodes in the same household) and then cleans the overlapping episodes in different houses: the latter may be data errors or they may be correct due to some men being head of multiple households if they are married polygamously. To distinguish between the errors and the polygamous episodes the dataset polyheads is merged in (a derived table made by house\_head\_list.do). The overlapping head episodes are kept for the \_poly version of the output but not for the \_nopoly version. The data cleaning may have resulted in some of the other variables not making sense, so they are altered as well.

### UNIT OF ANALYSIS

Individual

## Version

### VERSION DATE

2020-06-17

## Producers and sponsors

### PRIMARY INVESTIGATORS

| Name                                               | Affiliation                                   |
|----------------------------------------------------|-----------------------------------------------|
| Malawi Epidemiology and Intervention Research Unit | London School of Hygein and Tropical Medicine |

### OTHER IDENTIFICATIONS/ACKNOWLEDGMENTS

| Name                           |
|--------------------------------|
| MEIRU Core datamanagement team |

## Data collection

### DATA COLLECTION MODE

Face-to-face [f2f]

### DATA COLLECTORS

| Name                                               | Abbreviation |
|----------------------------------------------------|--------------|
| Malawi Epidemiology and Intervention Research Unit | MEIRU        |

## Access policy

### ACCESS CONDITIONS

These data are made available for licensed access under the following conditions:

1. Data and other material provided by MEIRU will not be redistributed or sold to other individuals, institutions or organisations without MEIRU's written agreement.
2. In the case of multi-centre datasets, data originating from a single contributing member centre of the collaboration may not be analysed or reported on in isolation without the express permission of the member centre concerned.
3. No attempt will be made to re-identify respondents, and there will be no use of the identity of any person or establishment discovered inadvertently. Any such discovery will be reported immediately to MEIRU.
4. No attempt will be made to produce links between datasets provided by MEIRU or between MEIRU data and other datasets that could identify individuals.
5. Any books, articles, conference papers, theses, dissertations, reports or other publications employing data obtained from MEIRU will cite the source, in line with the citation requirement provided with the dataset.
6. An electronic copy of all publications based on the requested data will be sent to MEIRU.

### ACCESS AUTHORITY

| Name                                               |
|----------------------------------------------------|
| Malawi Epidemiology and Intervention Research Unit |

### ARCHIVE WHERE STUDY IS ORIGINALLY STORED

Chilumba

## Metadata production

### DDI DOCUMENT ID

DDI-MWI-MEIRU-crsresidency-2002-v1

### PRODUCERS

| Name               | Abbreviation | Affiliation | Role                                                             |
|--------------------|--------------|-------------|------------------------------------------------------------------|
| Chifundo Kanjala   | CK           | MEIRU       | Documentation planning and supervision                           |
| Elizabeth Munthali | EM           | MEIRU       | Metadata entry and editing                                       |
| Estelle McLean     | EM           | LSHTM       | Production of the documentation used to create this DDI document |

### DATE OF METADATA PRODUCTION

2020-06-17

## Data Dictionary

| Data file                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | Cases  | Variables |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|-----------|
| <b>all_crs_members</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 79201  | 1         |
| <p><b>crs_residency_episodes_clean_nopoly</b></p> <p>crs_residency_episodes_clean_nopoly.dta is a cleaned version of crs_memball which is stset ready for survival analysis. Polygamous heads are now only living in one household and any overlapping episodes have been modified so there is no double counting of person time. This dataset should be used for calculating rates and numbers of residents but not for locating people and households.</p> <p>In the variable labels (S) indicates that the variable is unchanged from the original source and (D) indicates that it is a 'derived' variable, often using more than one source variable. For both types information on where the data come from and how they are processed is found in the 'Recoding and Derivation' section</p>                                                  | 139917 | 19        |
| <p><b>crs_residency_episodes_clean_poly</b></p> <p>crs_residency_episodes_clean_poly.dta is a cleaned version of crs_memball. It differs from crs_residency_episodes_clean_nopoly.dta in that polygamous heads are living in all of their households. Any overlapping episodes of other people have been modified so there is no double counting of their person time. This dataset should NOT be used for calculating rates and numbers of residents but can be used for locating people and households.</p> <p>In the variable labels (S) indicates that the variable is unchanged from the original source and (D) indicates that it is a 'derived' variable, often using more than one source variable. For both types information on where the data come from and how they are processed is found in the 'Recoding and Derivation' section</p> | 141407 | 14        |



**Data file: all\_crs\_members**

Cases: 79201

Variables: 1

**Variables**

| ID | Name  | Label                 | Question |
|----|-------|-----------------------|----------|
| V1 | ident | Unique identifier (S) |          |

Total: 1

**Data file: crs\_residency\_episodes\_clean\_nopoly**

crs\_residency\_episodes\_clean\_nopoly.dta is a cleaned version of crs\_memball which is stset ready for survival analysis. Polygamous heads are now only living in one household and any overlapping episodes have been modified so there is no double counting of person time. This dataset should be used for calculating rates and numbers of residents but not for locating people and households.

In the variable labels (S) indicates that the variable is unchanged from the original source and (D) indicates that it is a 'derived' variable, often using more than one source variable. For both types information on where the data come from and how they are processed is found in the 'Recoding and Derivation' section

Cases: 139917

Variables: 19

**Variables**

| ID  | Name        | Label                                  | Question |
|-----|-------------|----------------------------------------|----------|
| V2  | ident       | Unique identifier (S)                  |          |
| V3  | startdate   | Start date of episode (D)              |          |
| V4  | stopdate    | End date of episode (D)                |          |
| V5  | crshse      | Household identifier (S)               |          |
| V6  | rsnmovecenm | Reason for moving into household (S)   |          |
| V7  | rsnmovedep  | Reason for moving out of household (S) |          |
| V8  | gpsdeg      | GPS coordinate: south degrees          |          |
| V9  | gpssmin     | GPS coordinate: south minutes          |          |
| V10 | gpsdeg      | GPS coordinate: east degrees           |          |
| V11 | gpsemin     | GPS coordinate: east minutes           |          |
| V12 | sex         | Sex of individual (S)                  |          |
| V13 | birth_date  | Date of birth (D)                      |          |
| V14 | starttype   | Reason for start of episode (D)        |          |
| V15 | stoptype    | Reason for end of episode (D)          |          |
| V16 | st          | survival analysis variable (D)         |          |
| V17 | d           | survival analysis variable (D)         |          |
| V18 | origin      | survival analysis variable (D)         |          |
| V19 | t           | survival analysis variable (D)         |          |
| V20 | t0          | survival analysis variable (D)         |          |

Total: 19

**Data file: crs\_residency\_episodes\_clean\_poly**

crs\_residency\_episodes\_clean\_poly.dta is a cleaned version of crs\_memball. It differs from crs\_residency\_episodes\_clean\_nopoly.dta in that polygamous heads are living in all of their households. Any overlapping episodes of other people have been modified so there is no double counting of their person time. This dataset should NOT be used for calculating rates and numbers of residents but can be used for locating people and households.

In the variable labels (S) indicates that the variable is unchanged from the original source and (D) indicates that it is a 'derived' variable, often using more than one source variable. For both types information on where the data come from and how they are processed is found in the 'Recoding and Derivation' section

Cases: 141407

Variables: 14

**Variables**

| ID  | Name        | Label                                  | Question |
|-----|-------------|----------------------------------------|----------|
| V21 | ident       | Unique identifier (S)                  |          |
| V22 | startdate   | Start date of episode (D)              |          |
| V23 | stopdate    | End date of episode (D)                |          |
| V24 | crshse      | Household identifier (S)               |          |
| V25 | rsnmovecenm | Reason for moving into household (S)   |          |
| V26 | rsnmovedep  | Reason for moving out of household (S) |          |
| V27 | gpsdeg      | GPS coordinate: south degrees          |          |
| V28 | gpssmin     | GPS coordinate: south minutes          |          |
| V29 | gpsdeg      | GPS coordinate: east degrees           |          |
| V30 | gpsemin     | GPS coordinate: east minutes           |          |
| V31 | sex         | Sex of individual (S)                  |          |
| V32 | birth_date  | Date of birth (D)                      |          |
| V33 | starttype   | Reason for start of episode (D)        |          |
| V34 | stoptype    | Reason for end of episode (D)          |          |

Total: 14





 **IDENT: Unique identifier (S)**

**Data file:** all\_crs\_members

**Overview**

Valid: 79201    Invalid: 0  
Type: Discrete    Width: 7    Range: -    Format: character

**Description**

---

DEFINITION  
Unique identifier

---

**GPSEDEG: GPS coordinate: east degrees****Data file:** crs\_residency\_episodes\_clean\_nopoly**Overview**

Valid: 139917    Invalid: 0  
 Type: Discrete    Decimal: 0    Width: 2    Range: 34 - 34    Format: Numeric

**Description**

## DEFINITION

GPS coordinate: east degrees

**GPSEMIN: GPS coordinate: east minutes****Data file:** crs\_residency\_episodes\_clean\_nopoly**Overview**

Valid: 139917    Invalid: 0    Minimum: 4.71    Maximum: 16.127    Mean: 11.739    Standard deviation: 2.474  
 Type: Continuous    Decimal: 0    Width: 6    Range: 4.71 - 16.127    Format: Numeric

**Description**

## DEFINITION

GPS coordinate: east minutes

**IDENT: Unique identifier (S)****Data file:** crs\_residency\_episodes\_clean\_nopoly**Overview**

Valid: 139917    Invalid: 0  
 Type: Discrete    Width: 7    Range: -    Format: character

**Description**

## DEFINITION

Unique identifier

**Imputation and derivation**

## DERIVATION

variable: ident(table[s]: crs\_memball)

**STARTDATE: Start date of episode (D)****Data file:** crs\_residency\_episodes\_clean\_nopoly

## Overview

Valid: 139917    Minimum: 1995-10-10    Maximum: 2018-12-31  
 Type: Discrete    Width: 10    Range: -    Format: character

## Description

---

### DEFINITION

Start date of episode

## Imputation and derivation

---

### DERIVATION

variable: startdate(table[s]: crs\_memball)

startdate of some episodes may be changed during cleaning to remove polygamous episodes or data errors causing overlapping episodes

---

## ■ STOPDATE: End date of episode (D)

**Data file:** crs\_residency\_episodes\_clean\_nopoly

## Overview

Valid: 139917    Minimum: 2002-09-06    Maximum: 2018-12-31  
 Type: Discrete    Width: 10    Range: -    Format: character

## Description

---

### DEFINITION

End date of episode

## Imputation and derivation

---

### DERIVATION

variable: stopdate(table[s]: crs\_memball)

stopdate of some episodes may be changed during cleaning to remove polygamous episodes or data errors causing overlapping episodes

---

## ■ CRSHSE: Household identifier (S)

**Data file:** crs\_residency\_episodes\_clean\_nopoly

## Overview

Valid: 139917    Invalid: 0  
 Type: Discrete    Width: 7    Range: -    Format: character

## Description

---

### DEFINITION

Household identifier

## Imputation and derivation

### DERIVATION

variable: crshse(table[s]: crs\_memball)

## RSNMOVECENM: Reason for moving into household (S)

Data file: crs\_residency\_episodes\_clean\_nopoly

### Overview

Valid: 89934 Invalid: 49983

Type: Discrete Decimal: 0 Width: 1 Range: 1 - 5 Format: Numeric

### Questions and instructions

#### CATEGORIES

| Value   | Category            | Cases |       |
|---------|---------------------|-------|-------|
| 1       | Marriage            | 20896 | 23.2% |
| 2       | Divorce/sep         | 10729 | 11.9% |
| 3       | education start/end | 10197 | 11.3% |
| 4       | work start/end      | 17574 | 19.5% |
| 5       | Other               | 30538 | 34%   |
| Sysmiss |                     | 49983 |       |

### Description

#### DEFINITION

Reason for moving into household

## Imputation and derivation

### DERIVATION

variable: rsnmovecenm(table[s]: crs\_memball)

## RSNMOVEDEP: Reason for moving out of household (S)

Data file: crs\_residency\_episodes\_clean\_nopoly

### Overview

Valid: 90790 Invalid: 49127

Type: Discrete Decimal: 0 Width: 1 Range: 1 - 5 Format: Numeric

## Questions and instructions

### CATEGORIES

| Value   | Category            | Cases |       |
|---------|---------------------|-------|-------|
| 1       | Marriage            | 15412 | 17%   |
| 2       | Divorce/sep         | 11461 | 12.6% |
| 3       | education start/end | 10472 | 11.5% |
| 4       | work start/end      | 24255 | 26.7% |
| 5       | Other               | 29190 | 32.2% |
| Sysmiss |                     | 49127 |       |

### Description

#### DEFINITION

Reason for moving out of household

### Imputation and derivation

#### DERIVATION

variable: rsnmovedep(table[s]: crs\_memball)

### GPSSDEG: GPS coordinate: south degrees

Data file: crs\_residency\_episodes\_clean\_nopoly

#### Overview

Valid: 139917 Invalid: 0  
Type: Discrete Decimal: 0 Width: 2 Range: 10 - 10 Format: Numeric

#### Description

#### DEFINITION

GPS coordinate: south degrees

### GPSSMIN: GPS coordinate: south minutes

Data file: crs\_residency\_episodes\_clean\_nopoly

#### Overview

Valid: 139917 Invalid: 0 Minimum: 23.53 Maximum: 30.765 Mean: 27.156 Standard deviation: 1.458  
Type: Continuous Decimal: 0 Width: 6 Range: 23.53 - 30.765 Format: Numeric

## Description

### DEFINITION

GPS coordinate: south degrees

## SEX: Sex of individual (S)

Data file: crs\_residency\_episodes\_clean\_nopoly

### Overview

Valid: 139917    Invalid: 0  
 Type: Discrete    Decimal: 0    Width: 1    Range: 0 - 9    Format: Numeric

### Questions and instructions

#### CATEGORIES

| Value | Category | Cases |       |
|-------|----------|-------|-------|
| 0     | Male     | 63570 | 45.4% |
| 1     | Female   | 76347 | 54.6% |
| 8     | Unknown  | 0     | 0%    |
| 9     | Missing  | 0     |       |

## Description

### DEFINITION

Sex of individual

### Imputation and derivation

#### DERIVATION

variable: sex (table[s]: gen\_identity)

## BIRTH\_DATE: Date of birth (D)

Data file: crs\_residency\_episodes\_clean\_nopoly

### Overview

Valid: 139917    Minimum: 1905-06-15    Maximum: 2018-12-31  
 Type: Discrete    Width: 10    Range: -    Format: character

## Description

### DEFINITION

Date of birth

## Imputation and derivation

### DERIVATION

variable: birthdd (table[s]: gen\_identity)

variable: birthmm (table[s]: gen\_identity)

variable: birthyear (table[s]: gen\_identity)

variable: birthest (table[s]: gen\_identity)

Birth date is created from individual date elements on gen\_identity: if day missing 15 is used, if month missing June is used, if only year estimate the mid-point of the estimate range is used

## STARTTYPE: Reason for start of episode (D)

Data file: crs\_residency\_episodes\_clean\_nopoly

### Overview

Valid: 139917 Invalid: 0

Type: Discrete Decimal: 0 Width: 1 Range: 1 - 4 Format: Numeric

### Questions and instructions

### CATEGORIES

| Value | Category              | Cases |       |
|-------|-----------------------|-------|-------|
| 1     | Baseline recruit      | 28764 | 20.6% |
| 2     | Birth                 | 20426 | 14.6% |
| 3     | from CRS in-migration | 47139 | 33.7% |
| 4     | external in-migration | 43588 | 31.2% |

### Description

### DEFINITION

Reason for start of episode

## Imputation and derivation

### DERIVATION

variable: starttype(table[s]: crs\_memball)

Some start events maybe updated if they do not make sense in the context of other episodes.

## STOPTYPE: Reason for end of episode (D)

Data file: crs\_residency\_episodes\_clean\_nopoly

### Overview

Valid: 139917 Invalid: 0

Type: Discrete Decimal: 0 Width: 1 Range: 1 - 4 Format: Numeric



## Questions and instructions

### CATEGORIES

| Value | Category               | Cases |       |
|-------|------------------------|-------|-------|
| 1     | Present                | 45092 | 32.2% |
| 2     | Death                  | 3987  | 2.8%  |
| 3     | to CRS out-migration   | 47139 | 33.7% |
| 4     | external out-migration | 43699 | 31.2% |

### Description

#### DEFINITION

Reason for end of episode

### Imputation and derivation

#### DERIVATION

variable: stoptype(table[s]: crs\_memball)

The 'present' code is added for people who don't have an end event (last known still in the area), some end events maybe updated if they do not make sense in the context of other episodes.

## ST: survival analysis variable (D)

**Data file:** crs\_residency\_episodes\_clean\_nopoly

### Overview

Valid: 139917    Invalid: 0  
 Type: Discrete    Decimal: 0    Width: 1    Range: 1 - 1    Format: Numeric

### Description

#### DEFINITION

survival analysis variable

### Imputation and derivation

#### DERIVATION

The data are stset using the following code

```
stset stopdate, id(ident) failure(stoptype==2) origin(birth_date) scale(365.25) enter(startdate) time0(startdate)
```

## D: survival analysis variable (D)

**Data file:** crs\_residency\_episodes\_clean\_nopoly

## Overview

Valid: 139917    Invalid: 0  
 Type: Discrete    Decimal: 0    Width: 1    Range: 0 - 1    Format: Numeric

## Description

DEFINITION  
 survival analysis variable

## Imputation and derivation

DERIVATION  
 The data are stset using the following code  
 stset stopdate, id(ident) failure(stoptype==2) origin(birth\_date) scale(365.25) enter(startdate) time0(startdate)

## ORIGIN: survival analysis variable (D)

Data file: crs\_residency\_episodes\_clean\_nopoly

## Overview

Valid: 139917    Invalid: 0    Minimum: -19923    Maximum: 21549    Mean: 12298.218    Standard deviation: 6173.211

Type: Continuous    Decimal: 0    Width: 6    Range: -19923 - 21549    Format: Numeric

## Description

DEFINITION  
 survival analysis variable

## Imputation and derivation

DERIVATION  
 The data are stset using the following code  
 stset stopdate, id(ident) failure(stoptype==2) origin(birth\_date) scale(365.25) enter(startdate) time0(startdate)

## T: survival analysis variable (D)

Data file: crs\_residency\_episodes\_clean\_nopoly

## Overview

Valid: 139917    Invalid: 0    Minimum: 0.00137    Maximum: 110.281    Mean: 20.139    Standard deviation: 16.538  
 Type: Continuous    Decimal: 0    Width: 19    Range: 0.00136892539356605 - 110.280629705681    Format: Numeric

## Description

DEFINITION

survival analysis variable

## Imputation and derivation

---

### DERIVATION

The data are stset using the following code

```
stset stopdate, id(ident) failure(stoptype==2) origin(birth_date) scale(365.25) enter(startdate) time0(startdate)
```

---

## T0: survival analysis variable (D)

**Data file:** crs\_residency\_episodes\_clean\_nopoly

### Overview

Valid: 139917   Invalid: 0   Minimum: 0   Maximum: 101.495   Mean: 16.127   Standard deviation: 15.648  
 Type: Continuous   Decimal: 0   Width: 16   Range: 0 - 101.494866529774   Format: Numeric

### Description

---

### DEFINITION

survival analysis variable

## Imputation and derivation

---

### DERIVATION

The data are stset using the following code

```
stset stopdate, id(ident) failure(stoptype==2) origin(birth_date) scale(365.25) enter(startdate) time0(startdate)
```

---

**IDENT: Unique identifier (S)****Data file:** crs\_residency\_episodes\_clean\_poly**Overview**

Valid: 141407    Invalid: 0  
 Type: Discrete    Width: 7    Range: -    Format: character

**Description**

DEFINITION  
 Unique identifier

**Imputation and derivation**

DERIVATION  
 variable: ident(table[s]: crs\_memball)

**STARTDATE: Start date of episode (D)****Data file:** crs\_residency\_episodes\_clean\_poly**Overview**

Valid: 141407    Minimum: 1995-10-10    Maximum: 2018-12-31  
 Type: Discrete    Width: 10    Range: -    Format: character

**Description**

DEFINITION  
 Start date of episode

**Imputation and derivation**

DERIVATION  
 variable: startdate(table[s]: crs\_memball)  
 startdate of some episodes may be changed during cleaning to remove data errors causing overlapping episodes

**STOPDATE: End date of episode (D)****Data file:** crs\_residency\_episodes\_clean\_poly**Overview**

Valid: 141407    Minimum: 2002-08-21    Maximum: 2018-12-31  
 Type: Discrete    Width: 10    Range: -    Format: character

**Description**

DEFINITION

End date of episode

## Imputation and derivation

### DERIVATION

variable: stopdate(table[s]: crs\_memball)

stopdate of some episodes may be changed during cleaning to remove data errors causing overlapping episodes

## CRSHSE: Household identifier (S)

Data file: crs\_residency\_episodes\_clean\_poly

### Overview

Valid: 141407 Invalid: 0

Type: Discrete Width: 7 Range: - Format: character

### Description

### DEFINITION

Household identifier

## Imputation and derivation

### DERIVATION

variable: crshse(table[s]: crs\_memball)

## RSNMOVECENM: Reason for moving into household (S)

Data file: crs\_residency\_episodes\_clean\_poly

### Overview

Valid: 91164 Invalid: 50243

Type: Discrete Decimal: 0 Width: 1 Range: 1 - 5 Format: Numeric

### Questions and instructions

### CATEGORIES

| Value   | Category            | Cases |       |
|---------|---------------------|-------|-------|
| 1       | Marriage            | 21812 | 23.9% |
| 2       | Divorce/sep         | 10731 | 11.8% |
| 3       | education start/end | 10200 | 11.2% |
| 4       | work start/end      | 17668 | 19.4% |
| 5       | Other               | 30753 | 33.7% |
| Sysmiss |                     | 50243 |       |

## Description

### DEFINITION

Reason for moving into household

## Imputation and derivation

### DERIVATION

variable: rsnmovecenm(table[s]: crs\_memball)

## RSNMOVEDEP: Reason for moving out of household (S)

Data file: crs\_residency\_episodes\_clean\_poly

### Overview

Valid: 91828 Invalid: 49579

Type: Discrete Decimal: 0 Width: 1 Range: 1 - 5 Format: Numeric

## Questions and instructions

### CATEGORIES

| Value   | Category            | Cases |       |
|---------|---------------------|-------|-------|
| 1       | Marriage            | 15430 | 16.8% |
| 2       | Divorce/sep         | 12019 | 13.1% |
| 3       | education start/end | 10479 | 11.4% |
| 4       | work start/end      | 24452 | 26.6% |
| 5       | Other               | 29448 | 32.1% |
| Sysmiss |                     | 49579 |       |

## Description

### DEFINITION

Reason for moving out of household

## Imputation and derivation

### DERIVATION

variable: rsnmovedep(table[s]: crs\_memball)

## GPSSDEG: GPS coordinate: south degrees

Data file: crs\_residency\_episodes\_clean\_poly

## Overview

Valid: 141407   Invalid: 0  
 Type: Discrete   Decimal: 0   Width: 2   Range: 10 - 10   Format: Numeric

## Description

---

### DEFINITION

GPS coordinate: south degrees

---

## GPSSMIN: GPS coordinate: south minutes

Data file: crs\_residency\_episodes\_clean\_poly

## Overview

Valid: 141407   Invalid: 0   Minimum: 23.53   Maximum: 30.765   Mean: 27.158   Standard deviation: 1.459  
 Type: Continuous   Decimal: 0   Width: 6   Range: 23.53 - 30.765   Format: Numeric

## Description

---

### DEFINITION

GPS coordinate: south minutes

---

## GPSEDEG: GPS coordinate: east degrees

Data file: crs\_residency\_episodes\_clean\_poly

## Overview

Valid: 141407   Invalid: 0  
 Type: Discrete   Decimal: 0   Width: 2   Range: 34 - 34   Format: Numeric

## Description

---

### DEFINITION

GPS coordinate: east degrees

---

## GPSEMIN: GPS coordinate: east minutes

Data file: crs\_residency\_episodes\_clean\_poly

## Overview

Valid: 141407   Invalid: 0   Minimum: 4.71   Maximum: 16.127   Mean: 11.731   Standard deviation: 2.475  
 Type: Continuous   Decimal: 0   Width: 6   Range: 4.71 - 16.127   Format: Numeric

## Description

---

### DEFINITION

GPS coordinate: east minutes

**SEX: Sex of individual (S)****Data file:** crs\_residency\_episodes\_clean\_poly**Overview**

Valid: 141407    Invalid: 0  
 Type: Discrete    Decimal: 0    Width: 1    Range: 0 - 9    Format: Numeric

**Questions and instructions**

## CATEGORIES

| Value | Category | Cases |     |
|-------|----------|-------|-----|
| 0     | Male     | 65059 | 46% |
| 1     | Female   | 76348 | 54% |
| 8     | Unknown  | 0     | 0%  |
| 9     | Missing  | 0     |     |

**Description**

## DEFINITION

Sex of individual

**Imputation and derivation**

## DERIVATION

variable: sex (table[s]: gen\_identity)

**BIRTH\_DATE: Date of birth (D)****Data file:** crs\_residency\_episodes\_clean\_poly**Overview**

Valid: 141407    Minimum: 1905-06-15    Maximum: 2018-12-31  
 Type: Discrete    Width: 10    Range: -    Format: character

**Description**

## DEFINITION

Date of birth

**Imputation and derivation**

## DERIVATION



variable: birthdd (table[s]: gen\_identity)

variable: birthmm (table[s]: gen\_identity)

variable: birthyear (table[s]: gen\_identity)

variable: birthest (table[s]: gen\_identity)

Birth date is created from individual date elements on gen\_identity: if day missing 15 is used, if month missing June is used, if only year estimate the mid-point of the estimate range is used

## STARTTYPE: Reason for start of episode (D)

Data file: crs\_residency\_episodes\_clean\_poly

### Overview

Valid: 141407 Invalid: 0

Type: Discrete Decimal: 0 Width: 1 Range: 1 - 4 Format: Numeric

### Questions and instructions

#### CATEGORIES

| Value | Category              | Cases |       |
|-------|-----------------------|-------|-------|
| 1     | Baseline recruit      | 28764 | 20.3% |
| 2     | Birth                 | 20426 | 14.4% |
| 3     | from CRS in-migration | 48431 | 34.2% |
| 4     | external in-migration | 43786 | 31%   |

### Description

#### DEFINITION

Reason for start of episode

### Imputation and derivation

#### DERIVATION

variable: starttype(table[s]: crs\_memball)

Some start events maybe updated if they do not make sense in the context of other episodes.

## STOCTYPE: Reason for end of episode (D)

Data file: crs\_residency\_episodes\_clean\_poly

### Overview

Valid: 141407 Invalid: 0

Type: Discrete Decimal: 0 Width: 1 Range: 1 - 4 Format: Numeric

### Questions and instructions

#### CATEGORIES

| Value | Category               | Cases |       |
|-------|------------------------|-------|-------|
| 1     | Present                | 44943 | 31.8% |
| 2     | Death                  | 3971  | 2.8%  |
| 3     | to CRS out-migration   | 48431 | 34.2% |
| 4     | external out-migration | 44062 | 31.2% |

## Description

---

### DEFINITION

Reason for end of episode

## Imputation and derivation

---

### DERIVATION

variable: stoptype(table[s]: crs\_memball)

The 'present' code is added for people who don't have an end event (last known still in the area), some end events maybe updated if they do not make sense in the context of other episodes.

---