

**South Africa**

**Dr Kobus Herbst, AHRI, SAPRIN**

**Dr Guy Harling, AHRI**

**Dr Mark Siedner, AHRI**

**Dr Willem Hanekom, AHRI**

**AHRI.Surveillance Episodes YrAge\_Labour**

**Study Documentation**

April 24, 2025

# Metadata Production

|                             |   |
|-----------------------------|---|
| <b>Metadata Producer(s)</b> | Africa Health Research Institute (AHRI) |
| <b>Identification</b>       | DDI.AHRI.SEYAL                          |

# Table of Contents

|  |    |
|--|----|
| <a href="#">Overview</a> .....                         | 4  |
| <a href="#">Scope &amp; Coverage</a> .....             | 5  |
| <a href="#">Producers &amp; Sponsors</a> .....         | 5  |
| <a href="#">Sampling</a> .....                         | 5  |
| <a href="#">Data Collection</a> .....                  | 6  |
| <a href="#">Data Processing &amp; Appraisal</a> .....  | 6  |
| <a href="#">Accessibility</a> .....                    | 6  |
| <a href="#">Files Description</a> .....                | 7  |
| <a href="#">SurveillanceEpisodesYrAge_Labour</a> ..... | 7  |
| <a href="#">Variables List</a> .....                   | 8  |
| <a href="#">SurveillanceEpisodesYrAge_Labour</a> ..... | 8  |
| <a href="#">Variables Description</a> .....            | 10 |
| <a href="#">SurveillanceEpisodesYrAge_Labour</a> ..... | 11 |

## AHRI.Surveillance Episodes YrAge\_Labour

### Overview

|                       |            |
|-----------------------|------------|
| <b>Identification</b> | AHRI.SEYAL |
| <b>Version</b>        | v1.0.0     |

#### Abstract

Africa Health Research Institute (AHRI) is a multidisciplinary, independent research institute. AHRI's goal is to become a source of fundamental discoveries into the susceptibility, transmission and cure of HIV, TB and related diseases. We also seek ways to improve diagnosis, prevention and treatment. To achieve this, we bring together leading researchers from different fields, use cutting edge science to improve people's health, and help to train the next generation of African scientists.

AHRI has conducted individual and household demographic surveillance since 2000. Demographic surveillance data are processed for longitudinal analyses by organizing the demographic surveillance data into individual and household residency episodes at a geographical physical location, and individual membership episodes within a household. Start events of residency and memberships include enumeration, birth, in-migration and relocating into a household from within the study population. Exit events of residency and memberships include death (by cause), out-migration and relocating to another location in the study population. AHRI annually produces a suite of Individual Demographic surveillance datasets which contains detailed records of all individuals under surveillance since January 2000. The datasets detail individual residency and membership patterns, household mortality (including cause of death), fertility, migration, HIV status, health care utilization, and household socio-economic profile.

Each record in the dataset represents a period of observation for an individual during which all the recorded characteristics of the individual stay constant. For example, on the birthday of the individual a new episode will start, because the age of the individual has changed. An out-migration will result in a new episode, because the location or residential status has changed. Any change in one of the status values, such as education, marital status, social-economic status and HIV will likewise result in a new episode on the date of the change.

The following are the versions of the Demographic surveillance datasets:

- CoresidencyEpisodes
- EducationCleaned
- EducationStatuses
- HIVStatusEpisodes
- HouseholdAssets
- HouseholdMap
- HouseholdObservations
- HouseholdSocioEconomicStatus
- HouseholdStatusEpisodes
- IndividualHSEEpisodes
- IndividualMap
- LabourEpisodes
- LabourStatuses
- LocationMap
- MaritalStatuses
- PartnerEpisodes
- SurveillanceEpisodesBasic
- SurveillanceEpisodesYTAge
- SurveillanceEpisodes YrAge\_Coresidency
- SurveillanceEpisodesYrAge\_HIV
- SurveillanceEpisodes VrAge\_HSE
- SurveillanceEpisodes YrAge\_Labour
- SurveillanceEpisodesYrAge\_Partners
- SurveillanceEpisodesYrAgeDel

|   |                      |
|---|----------------------|
| SurveillanceEpisodes VrAgeDel_HSE<br>SurveillanceEpisodesYrAgeDeL_HSE_Labour<br>SurveillanceEpisodesYrAgeDeL_HSE_Labour_Partner_Parents<br>SurveillanceEpisodes VrAgeDel_HSE_Labour_Partner_Parents_HIV<br>SurveillanceEpisodesYrAgeDeL_HSE_Labour_Partners |                      |
| <b>Kind of Data</b>   | Event History Data   |
| <b>Unit of Analysis</b>   | Episodes of exposure |

| <b>Scope &amp; Coverage</b>   |   |
|---|---|
| <b>Keywords</b>   | Fertility, Mortality, Migration, Social-economic status, HIV, ART |
| <b>Topics</b>   | Surveillance Episodes YrAge_Labour                                |
| <b>Time Period(s)</b>   | 2000-2023   |
| <b>Countries</b>  | South Africa  |
| <b>Geographic Coverage</b><br>AHRI is situated in the south-east portion of the uMkhanyakude district of KwaZulu-Natal province near the town of Mtubatuba. It is bounded on the west by the Umfolozi-Hluhluwe nature reserve, on the south by the Umfolozi river, on the east by the N2 highway (except form portions where the KwaMsane township strangles the highway) and in the north by the provincial road R22. The surveillance area is approximately 845km <sup>2</sup> in size.   |   |
| <b>Universe</b><br>Households resident in dwellings within the study area are eligible for inclusion in the household component of AHRI Demographic Surveillance. All individuals identified by the household proxy informant as members of the households are enumerated. A resident household member is an individual that intends to sleep majority of time at the dwelling occupied by the household over a four-month period. Households include resident and non-resident members. An individual is a non-resident member if they have close ties to the household, but do not physically reside with the household at the dwelling most of the time. Non-resident members can also be called temporary migrants and they are enumerated within the household list. Because household membership is not tied to physical residency, an individual may be a member of more than one household. |   |

| <b>Producers &amp; Sponsors</b> |  |
|---------------------------------|--|
| <b>Primary Investigator(s)</b>  | Dr Kobus Herbst, AHRI, SAPRIN<br>Dr Guy Harling, AHRI<br>Dr Mark Siedner, AHRI<br>Dr Willem Hanekom, AHRI  |
| <b>Other Producer(s)</b>        | Africa Health Research Institute (AHRI)  |
| <b>Funding Agency/ies</b>       | Wellcome Trust (WT) , Core funding<br>SAPRIN (SAPRIN)  |
| <b>Other Acknowledgment(s)</b>  | AHRI Research Data Management Team , Research Data Management , AHRI<br>AHRI Data Collection Team , Data Collection , AHRI<br>AHRI Community Engagement Team , Community Engagement , AHRI |

| <b>Sampling</b>   |  |
|---|--|
| <b>Sampling Procedure</b><br>This dataset is not based on a sample but contains information from the complete demographic surveillance areas. |  |

## Data Collection

|                              |                                    |
|------------------------------|------------------------------------|
| <b>Data Collection Dates</b> | start 2000-01-01<br>end 2023-12-31 |
|------------------------------|------------------------------------|

## Data Processing & Appraisal

### Data Editing

Data collected in this study was centrally stored on AHRI's secure database servers. Data was collected through interviews at the call centre on central server using Survey Solutions, and in the field on paper and encrypted tablet computers using REDCap, and then securely uploaded to a central server. Database rules and constraints were enforced on the master database to ensure that only valid data was stored. Personally identifiable data was restricted to separate database tables accessible only to data management staff .

## Accessibility

### Access Conditions

Access to the data requires accurate completion of the online data access application form accessible on the AHRI Data repository(<<https://data.ahri.org/>>). Data users are required to abide by the data use conditions stipulated on the application for access to the data. Failure to do so may result in their data access privileges being revoked by the Data Custodian. In order to recognise the effort and intellectual contributions of AHRI investigators in producing and curating the data, users of AHRI data must acknowledge the source of the data and abide by the terms and conditions under which the data is accessed and must cite the dataset in publication using the citation provided as part of this documentation. All analytical datasets published on the AHRI Data Repository are assigned digital object identifier (DOIs) and the DOIs can be found on the Data Repository under Study Description tab - Access policy. AHRI data users are required to always cite the dataset using the relevant DOI.

### Citation Requirements

Herbst, K., Harling, G., Siedner, M., & Hanekom, W. (2025). AHRI.Surveillance Episodes YrAge\_Labour [Data set]. Africa Health Research Institute. DOI:<https://doi.org/10.23664/AHRI.SEYAL>

# Files Description

Dataset contains 1 file(s)

| <b>SurveillanceEpisodesYrAge_Labour</b> |         |
|---|---------|
| <b># Cases</b>                          | 6216557 |
| <b># Variable(s)</b>                    | 42      |

# Variables List

Dataset contains 42 variable(s)

| File SurveillanceEpisodesYrAge_Labour |                            |  |            |              |         |         |          |
|---------------------------------------|----------------------------|--|------------|--------------|---------|---------|----------|
| #                                     | Name                       | Label  | Type       | Format       | Valid   | Invalid | Question |
| 1                                     | <a href="#">Individu..</a> | Unique individual identifier   | continuous | numeric-9.0  | 6216557 | 0       | -        |
| 2                                     | <a href="#">Sex</a>        | Sex of individual  | discrete   | numeric-9.0  | 6216557 | 0       | -        |
| 3                                     | <a href="#">DoB</a>        | Date of birth  | discrete   | character-11 | 6216557 | -       | -        |
| 4                                     | <a href="#">DoD</a>        | Date of death  | discrete   | character-11 | 516421  | -       | -        |
| 5                                     | <a href="#">MotherId</a>   | Mother's IndividualId  | continuous | numeric-9.0  | 4270925 | 1945632 | -        |
| 6                                     | <a href="#">FatherId</a>   | Father's IndividualId  | continuous | numeric-9.0  | 2264986 | 3951571 | -        |
| 7                                     | <a href="#">LocationId</a> | Where individual was resident, household residence if non-resident         | continuous | numeric-9.0  | 6216557 | 0       | -        |
| 8                                     | <a href="#">Househol..</a> | Unique household identifier of the household the individual is a member of | continuous | numeric-9.0  | 6216557 | 0       | -        |
| 9                                     | <a href="#">HHRelati..</a> | Relationship to head of household  | discrete   | numeric-22.0 | 6216557 | 0       | -        |
| 10                                    | <a href="#">Resident</a>   | Whether individual is resident for duration of episode                     | discrete   | numeric-15.0 | 6216557 | 0       | -        |
| 11                                    | <a href="#">Episode</a>    | This episode number (first=1, last=Episodes)                               | continuous | numeric-9.0  | 6216557 | 0       | -        |
| 12                                    | <a href="#">Episodes</a>   | Total number of episodes for individual                                    | continuous | numeric-9.0  | 6216557 | 0       | -        |
| 13                                    | <a href="#">StartDate</a>  | Start date of episode (inclusive)  | discrete   | character-11 | 6216557 | -       | -        |
| 14                                    | <a href="#">EndDate</a>    | End date of episode (inclusive)  | discrete   | character-11 | 6216557 | -       | -        |
| 15                                    | <a href="#">Days</a>       | Duration in days of episode  | continuous | numeric-9.0  | 6216557 | 0       | -        |
| 16                                    | <a href="#">Calendar..</a> | Calendar year during episode   | continuous | numeric-9.0  | 6216557 | 0       | -        |
| 17                                    | <a href="#">Age</a>        | Age in completed years during episode                                      | continuous | numeric-9.0  | 6216557 | 0       | -        |
| 18                                    | <a href="#">AgeGroup</a>   | Age group during episode   | discrete   | numeric-9.0  | 6216557 | 0       | -        |
| 19                                    | <a href="#">Born</a>       | Episode starts with the birth of the individual                            | discrete   | numeric-15.0 | 6216557 | 0       | -        |
| 20                                    | <a href="#">Enumerat..</a> | Episode starts with an enumeration   | discrete   | numeric-15.0 | 6216557 | 0       | -        |
| 21                                    | <a href="#">InMigrat..</a> | Episode starts with an in-migration  | discrete   | numeric-15.0 | 6216557 | 0       | -        |
| 22                                    | <a href="#">Location..</a> | Episode starts with an internal migration                                  | discrete   | numeric-15.0 | 6216557 | 0       | -        |
| 23                                    | <a href="#">ExtResSt..</a> | Flag to indicate start of external residence                               | discrete   | numeric-15.0 | 6216557 | 0       | -        |
| 24                                    | <a href="#">Particip..</a> | Resumed participation after refusal  | discrete   | numeric-15.0 | 6216557 | 0       | -        |

| <b>File SurveillanceEpisodesYrAge_Labour</b> |                            |   |          |              |         |         |          |
|--|----------------------------|---|----------|--------------|---------|---------|----------|
| #  | Name                       | Label   | Type     | Format       | Valid   | Invalid | Question |
| 25   | <a href="#">Died</a>       | Episode ends with the death of the individual                 | discrete | numeric-15.0 | 6216557 | 0       | -        |
| 26   | <a href="#">OutMigra..</a> | Episode ends with the out-migration                           | discrete | numeric-15.0 | 6216557 | 0       | -        |
| 27   | <a href="#">Location..</a> | Episode ends with an internal migration                       | discrete | numeric-15.0 | 6216557 | 0       | -        |
| 28   | <a href="#">ExtResEnd</a>  | Flag to indicate end of external residence                    | discrete | numeric-15.0 | 6216557 | 0       | -        |
| 29   | <a href="#">Refusal</a>    | Individual refused follow-up                                  | discrete | numeric-15.0 | 6216557 | 0       | -        |
| 30   | <a href="#">LostToFo..</a> | Individual was lost to follow-up at the end of the episode    | discrete | numeric-15.0 | 6216557 | 0       | -        |
| 31   | <a href="#">Current</a>    | Individual still under surveillance                           | discrete | numeric-15.0 | 6216557 | 0       | -        |
| 32   | <a href="#">HHRelati..</a> | Relationship to head of household changed during this episode | discrete | numeric-15.0 | 6216557 | 0       | -        |
| 33   | <a href="#">Membersh..</a> | Flag to indicate start of household membership                | discrete | numeric-15.0 | 6216557 | 0       | -        |
| 34   | <a href="#">Membersh..</a> | Flag to indicate end of household membership                  | discrete | numeric-15.0 | 6216557 | 0       | -        |
| 35   | <a href="#">GapStart</a>   | Start of a gap in individual exposure record                  | discrete | numeric-15.0 | 6216557 | 0       | -        |
| 36   | <a href="#">GapEnd</a>     | End of gap in individual exposure record                      | discrete | numeric-15.0 | 6216557 | 0       | -        |
| 37   | <a href="#">AgeChange</a>  | Flag age changed  | discrete | numeric-15.0 | 6216557 | 0       | -        |
| 38   | <a href="#">YrChange</a>   | Flag calendar year changed                                    | discrete | numeric-15.0 | 6216557 | 0       | -        |
| 39   | <a href="#">Membersh..</a> | Number of concurrent household memberships                    | discrete | numeric-9.0  | 6216557 | 0       | -        |
| 40   | <a href="#">Educatio..</a> | Education level achieved during this calendar year            | discrete | numeric-18.0 | 4642043 | 1574514 | -        |
| 41   | <a href="#">Employme..</a> | Employment Status   | discrete | numeric-15.0 | 2845395 | 3371162 | -        |
| 42   | <a href="#">Employme..</a> | Flag to indicate change in employment status                  | discrete | numeric-15.0 | 2845395 | 3371162 | -        |

# Variables Description

Dataset contains 42 variable(s)

## File : SurveillanceEpisodesYrAge\_Labour

### # IndividualId: Unique individual identifier

**Information** [Type= continuous] [Format=numeric] [Range= 1-258503] [Missing=\*]

**Statistics [NW/ W]** [Valid=6216557 /-] [Invalid=0 /-] [Mean=117445.522 /-] [StdDev=76056.02 /-]

### # Sex: Sex of individual

**Information** [Type= discrete] [Format=numeric] [Range= 0-3] [Missing=\*]

**Statistics [NW/ W]** [Valid=6216557 /-] [Invalid=0 /-]

| Value | Label   | Cases   | Percentage |
|-------|---------|---------|------------|
| 0     |         | 55      | 0.0%       |
| 1     | Male    | 2938861 | 47.3%      |
| 2     | Female  | 3277641 | 52.7%      |
| 3     | Unknown | 0       |            |

*Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.*

### # DoB: Date of birth

**Information** [Type= discrete] [Format=character] [Missing=\*]

**Statistics [NW/ W]** [Valid=6216557 /-]

### # DoD: Date of death

**Information** [Type= discrete] [Format=character] [Missing=\*]

**Statistics [NW/ W]** [Valid=516421 /-]

### # MotherId: Mother's IndividualId

**Information** [Type= continuous] [Format=numeric] [Range= 17-258500] [Missing=\*]

**Statistics [NW/ W]** [Valid=4270925 /-] [Invalid=1945632 /-] [Mean=154235.243 /-] [StdDev=64550.355 /-]

### # FatherId: Father's IndividualId

**Information** [Type= continuous] [Format=numeric] [Range= 49-258496] [Missing=\*]

**Statistics [NW/ W]** [Valid=2264986 /-] [Invalid=3951571 /-] [Mean=162812.027 /-] [StdDev=57795.497 /-]

### # LocationId: Where individual was resident, household residence if non-resident

**Information** [Type= continuous] [Format=numeric] [Range= 0-35891] [Missing=\*]

**Statistics [NW/ W]** [Valid=6216557 /-] [Invalid=0 /-] [Mean=17777.168 /-] [StdDev=10346.801 /-]

### # HouseholdId: Unique household identifier of the household the individual is a member of

**Information** [Type= continuous] [Format=numeric] [Range= 1-40054] [Missing=\*]

**Statistics [NW/ W]** [Valid=6216557 /-] [Invalid=0 /-] [Mean=19650.391 /-] [StdDev=11285.733 /-]

### # HHRelationshipTypeId: Relationship to head of household

**Information** [Type= discrete] [Format=numeric] [Range= 1-12] [Missing=\*]

**Statistics [NW/ W]** [Valid=6216557 /-] [Invalid=0 /-]

| Value | Label              | Cases   | Percentage |
|-------|--------------------|---------|------------|
| 1     | Head               | 826043  | 13.3%      |
| 2     | Spouse             | 324725  | 5.2%       |
| 3     | Child              | 2106768 | 33.9%      |
| 4     | Son/Daughter-inlaw | 28752   | 0.5%       |
| 5     | Grandchild         | 1346775 | 21.7%      |
| 6     | Parent             | 46359   | 0.7%       |

## File : SurveillanceEpisodesYrAge\_Labour

### # HHRelationshipTypeId: Relationship to head of household

| Value | Label                  | Cases   | Percentage |
|-------|------------------------|---------|------------|
| 7     | Parent-inlaw           | 1658    | 0.0%       |
| 8     | Grandparent            | 6909    | 0.1%       |
| 9     | Sibling                | 1045159 | 16.8%      |
| 10    | Other relative         | 14300   | 0.2%       |
| 11    | Domestic worker/tenant | 326970  | 5.3%       |
| 12    | Unrelated/Other        | 142139  | 2.3%       |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # Resident: Whether individual is resident for duration of episode

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*] |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-]                           |

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 1843738 | 29.7%      |
| 1     | Yes             | 4372819 | 70.3%      |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # Episode: This episode number (first=1, last=Episodes)

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= continuous] [Format=numeric] [Range= 1-77] [Missing=*]         |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-] [Mean=19.231 /-] [StdDev=14.932 /-] |

### # Episodes: Total number of episodes for individual

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= continuous] [Format=numeric] [Range= 1-77] [Missing=*]         |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-] [Mean=37.461 /-] [StdDev=17.841 /-] |

### # StartDate: Start date of episode (inclusive)

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= discrete] [Format=character] [Missing=*] |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-]                              |

### # EndDate: End date of episode (inclusive)

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= discrete] [Format=character] [Missing=*] |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-]                              |

### # Days: Duration in days of episode

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= continuous] [Format=numeric] [Range= 0-366] [Missing=*]          |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-] [Mean=156.611 /-] [StdDev=103.714 /-] |

### # CalendarYear: Calendar year during episode

|                           |  |
|---------------------------|--|
| <b>Information</b>        | [Type= continuous] [Format=numeric] [Range= 2000-2023] [Missing=*]     |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-] [Mean=2013.044 /-] [StdDev=7.062 /-] |

### # Age: Age in completed years during episode

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= continuous] [Format=numeric] [Range= 0-228] [Missing=*]        |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-] [Mean=25.615 /-] [StdDev=18.233 /-] |

### # AgeGroup: Age group during episode

|                    |   |
|--------------------|---|
| <b>Information</b> | [Type= discrete] [Format=numeric] [Range= 0-18] [Missing=*] |
|--------------------|---|

## File : SurveillanceEpisodesYrAge\_Labour

### # AgeGroup: Age group during episode

Statistics [NW/ W] [Valid=6216557 /-] [Invalid=0 /-]

| Value | Label | Cases  | Percentage |
|-------|-------|--------|------------|
| 0     | <1    | 111693 | 1.8%       |
| 1     | 01-04 | 498824 | 8.0%       |
| 2     | 05-09 | 657539 | 10.6%      |
| 3     | 10-14 | 660196 | 10.6%      |
| 4     | 15-19 | 721658 | 11.6%      |
| 5     | 20-24 | 759551 | 12.2%      |
| 6     | 25-29 | 655365 | 10.5%      |
| 7     | 30-34 | 522163 | 8.4%       |
| 8     | 35-39 | 397405 | 6.4%       |
| 9     | 40-44 | 295171 | 4.7%       |
| 10    | 45-49 | 230061 | 3.7%       |
| 11    | 50-54 | 184439 | 3.0%       |
| 12    | 55-59 | 149956 | 2.4%       |
| 13    | 60-64 | 120267 | 1.9%       |
| 14    | 65-69 | 87027  | 1.4%       |
| 15    | 70-74 | 64217  | 1.0%       |
| 16    | 75-79 | 43852  | 0.7%       |
| 17    | 80-84 | 29103  | 0.5%       |
| 18    | 85+   | 28070  | 0.5%       |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # Born: Episode starts with the birth of the individual

Information [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=\*]

Statistics [NW/ W] [Valid=6216557 /-] [Invalid=0 /-]

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 6168751 | 99.2%      |
| 1     | Yes             | 47806   | 0.8%       |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # Enumeration: Episode starts with an enumeration

Information [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=\*]

Statistics [NW/ W] [Valid=6216557 /-] [Invalid=0 /-]

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 6077790 | 97.8%      |
| 1     | Yes             | 138767  | 2.2%       |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # InMigration: Episode starts with an in-migration

Information [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=\*]

Statistics [NW/ W] [Valid=6216557 /-] [Invalid=0 /-]

## File : SurveillanceEpisodesYrAge\_Labour

### # InMigration: Episode starts with an in-migration

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 6063658 | 97.5%      |
| 1     | Yes             | 152899  | 2.5%       |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # LocationEntry: Episode starts with an internal migration

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*] |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-]                           |

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 6163962 | 99.2%      |
| 1     | Yes             | 52595   | 0.8%       |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # ExtResStart: Flag to indicate start of external residence

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*] |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-]                           |

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 6044697 | 97.2%      |
| 1     | Yes             | 171860  | 2.8%       |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # Participation: Resumed participation after refusal

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*] |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-]                           |

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 6216557 | 100.0%     |
| 1     | Yes             | 0       |            |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # Died: Episode ends with the death of the individual

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*] |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-]                           |

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 6189843 | 99.6%      |
| 1     | Yes             | 26714   | 0.4%       |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # OutMigration: Episode ends with the out-migration

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*] |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-]                           |

## File : SurveillanceEpisodesYrAge\_Labour

### # OutMigration: Episode ends with the out-migration

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 5999011 | 96.5%      |
| 1     | Yes             | 217546  | 3.5%       |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # LocationExit: Episode ends with an internal migration

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*] |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-]                           |

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 6163962 | 99.2%      |
| 1     | Yes             | 52595   | 0.8%       |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # ExtResEnd: Flag to indicate end of external residence

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*] |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-]                           |

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 6125607 | 98.5%      |
| 1     | Yes             | 90950   | 1.5%       |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # Refusal: Individual refused follow-up

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*] |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-]                           |

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 6214285 | 100.0%     |
| 1     | Yes             | 2272    | 0.0%       |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # LostToFollowUp: Individual was lost to follow-up at the end of the episode

|                           |  |
|---------------------------|--|
| <b>Information</b>        | [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*] |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-]                          |

### # Current: Individual still under surveillance

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*] |
| <b>Statistics [NW/ W]</b> | [Valid=6216557 /-] [Invalid=0 /-]                           |

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 6059806 | 97.5%      |
| 1     | Yes             | 156751  | 2.5%       |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

## File : SurveillanceEpisodesYrAge\_Labour

### # HHRelationship: Relationship to head of household changed during this episode

**Information** [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=\*]

**Statistics [NW/ W]** [Valid=6216557 /-] [Invalid=0 /-]

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 6148838 | 98.9%      |
| 1     | Yes             | 67719   | 1.1%       |

*Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.*

### # MembershipStart: Flag to indicate start of household membership

**Information** [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=\*]

**Statistics [NW/ W]** [Valid=6216557 /-] [Invalid=0 /-]

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 5876550 | 94.5%      |
| 1     | Yes             | 340007  | 5.5%       |

*Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.*

### # MembershipEnd: Flag to indicate end of household membership

**Information** [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=\*]

**Statistics [NW/ W]** [Valid=6216557 /-] [Invalid=0 /-]

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 5876550 | 94.5%      |
| 1     | Yes             | 340007  | 5.5%       |

*Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.*

### # GapStart: Start of a gap in individual exposure record

**Information** [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=\*]

**Statistics [NW/ W]** [Valid=6216557 /-] [Invalid=0 /-]

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 6201675 | 99.8%      |
| 1     | Yes             | 14882   | 0.2%       |

*Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.*

### # GapEnd: End of gap in individual exposure record

**Information** [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=\*]

**Statistics [NW/ W]** [Valid=6216557 /-] [Invalid=0 /-]

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 6201675 | 99.8%      |
| 1     | Yes             | 14882   | 0.2%       |

*Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.*

### # AgeChange: Flag age changed

**Information** [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=\*]

## File : SurveillanceEpisodesYrAge\_Labour

### # AgeChange: Flag age changed

Statistics [NW/ W] [Valid=6216557 /-] [Invalid=0 /-]

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 3565792 | 57.4%      |
| 1     | Yes             | 2650765 | 42.6%      |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # YrChange: Flag calendar year changed

Information [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=\*]

Statistics [NW/ W] [Valid=6216557 /-] [Invalid=0 /-]

| Value | Label           | Cases   | Percentage |
|-------|-----------------|---------|------------|
| -1    | Unknown/Missing | 0       |            |
| 0     | No              | 3691426 | 59.4%      |
| 1     | Yes             | 2525131 | 40.6%      |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # Memberships: Number of concurrent household memberships

Information [Type= discrete] [Format=numeric] [Range= 1-10] [Missing=\*]

Statistics [NW/ W] [Valid=6216557 /-] [Invalid=0 /-]

### # EducationStatus: Education level achieved during this calendar year

Information [Type= discrete] [Format=numeric] [Range= -1-99] [Missing=\*]

Statistics [NW/ W] [Valid=4642043 /-] [Invalid=1574514 /-]

| Value | Label              | Cases   | Percentage |
|-------|--------------------|---------|------------|
| -1    | Unknown            | 87345   | 1.9%       |
| 0     | Grade 0            | 62520   | 1.3%       |
| 1     | Grade 1            | 131073  | 2.8%       |
| 2     | Grade 2            | 171682  | 3.7%       |
| 3     | Grade 3            | 177341  | 3.8%       |
| 4     | Grade 4            | 229060  | 4.9%       |
| 5     | Grade 5            | 169520  | 3.7%       |
| 6     | Grade 6            | 183706  | 4.0%       |
| 7     | Grade 7            | 250233  | 5.4%       |
| 8     | Grade 8            | 247244  | 5.3%       |
| 9     | Grade 9            | 240960  | 5.2%       |
| 10    | Grade 10           | 393699  | 8.5%       |
| 11    | Grade 11           | 410260  | 8.8%       |
| 12    | Grade 12           | 1529346 | 32.9%      |
| 13    | Special School     | 0       |            |
| 21    | Higher Certificate | 0       |            |
| 22    | National Diploma   | 0       |            |
| 23    | Degree             | 0       |            |
| 24    | Honours            | 0       |            |
| 25    | Masters            | 0       |            |
| 26    | Doctors            | 0       |            |

## File : SurveillanceEpisodesYrAge\_Labour

### # EducationStatus: Education level achieved during this calendar year

| Value   | Label         | Cases   | Percentage |
|---------|---------------|---------|------------|
| 31      | ABET 1        | 0       |            |
| 32      | ABET 2        | 0       |            |
| 33      | ABET 3        | 0       |            |
| 34      | ABET 4        | 0       |            |
| 41      | FET 1         | 0       |            |
| 42      | FET 2         | 0       |            |
| 43      | FET 3         | 0       |            |
| 50      | LT 1 year     | 0       |            |
| 51      | Creche        | 0       |            |
| 96      | Refused       | 0       |            |
| 97      | NAD           | 0       |            |
| 98      | No education  | 358054  | 7.7%       |
| 99      | Not attending | 0       |            |
| Sysmiss |               | 1574514 |            |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # EmploymentStatus: Employment Status

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= discrete] [Format=numeric] [Range= 0-15] [Missing=*] |
| <b>Statistics [NW/ W]</b> | [Valid=2845395 /-] [Invalid=3371162 /-]                     |

| Value   | Label           | Cases   | Percentage |
|---------|-----------------|---------|------------|
| 0       | Missing/Refused | 0       |            |
| 1       | Employed        | 924484  | 32.5%      |
| 2       | Part-time       | 154976  | 5.4%       |
| 13      | Education       | 281837  | 9.9%       |
| 15      | Unemployed      | 1484098 | 52.2%      |
| Sysmiss |                 | 3371162 |            |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

### # EmploymentStatusChange: Flag to indicate change in employment status

|                           |   |
|---------------------------|---|
| <b>Information</b>        | [Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*] |
| <b>Statistics [NW/ W]</b> | [Valid=2845395 /-] [Invalid=3371162 /-]                     |

| Value   | Label           | Cases   | Percentage |
|---------|-----------------|---------|------------|
| -1      | Unknown/Missing | 0       |            |
| 0       | No              | 2450176 | 86.1%      |
| 1       | Yes             | 395219  | 13.9%      |
| Sysmiss |                 | 3371162 |            |

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.