

South Africa - SAPRIN Individual Surveillance Episodes 2020 Datasets

**Prof Mark Collinson, Dr Kobus Herbst, Prof Steve Tollman, Dr Eric Maimela, Prof
Willem Hanekom**

Report generated on: July 1, 2025

Visit our data catalog at: <https://saprindata.samrc.ac.za/index.php>

Identification

SURVEY ID NUMBER
SAPRIN.SISED2020V1

TITLE
SAPRIN Individual Surveillance Episodes 2020 Datasets

COUNTRY

Name	Country code
South Africa	RSA

STUDY TYPE
Demographic Surveillance

SERIES INFORMATION
This dataset contains demographic surveillance data covering the period from 1 Jan 1993 to 31 December 2017.

ABSTRACT
The 'South African Population Research Infrastructure Network' (SAPRIN) is a national research infrastructure funded through the Department of Science, Technology and Innovation and hosted by the South African Medical Research Council. One of SAPRIN's initial goals has been to harmonise and share the longitudinal data from the three current Health and Demographic Surveillance System *(HDSS)* Nodes. These long-standing nodes are the MRC/Wits University Agincourt HDSS in Bushbuckridge District, Mpumalanga, established in 1993, with a population of 113 113 people; the University of Limpopo DIMAMO HDSS in the Capricorn District of Limpopo, established in 1996, with a current population of 38 479; and the Africa Health Research Institute (AHRI) HDSS in uMkhanyakude District, KwaZulu-Natal, established in 2000, with a current population of 139 250.

For an individual to be eligible for inclusion in the surveillance, the individual must be a member of a household resident within the geographic boundaries of a SAPRIN node. For a household to be resident, it must have at least one household member who is resident within the surveillance area. Households and household membership are self-defined by the household informant interviewed by the fieldworker at their place of residence (or during a telephonic interview with the household informant). Household members so identified could be resident, that is sleep the majority of nights at this household's place of residence, or could be resident elsewhere (usually outside the surveillance area, but potentially within the surveillance area, in which case they will be a resident household member of the household resident at that place - In this dataset, individuals are members of a single household at a time, and in this example, the non-resident member of the household who is resident elsewhere in the surveillance area, will be reflected in the dataset as a resident member of the household this individual is co-resident with and not also as a non-resident member of this household). The resident status of household members can change: they can move out of the surveillance area to be resident elsewhere, but still be considered household members (so-called 'temporary migration'), such cases are reflected in the data as episodes of external residence; or temporary migrants can return to take up residence again with the household, initiating a new episode of residence internal to the surveillance area.

In addition to these periods of internal and external residence punctuated by in- and outmigration, surveillance episodes can be started by the birth of an individual, if the child is born to a resident mother, their birth starts a period of internal residency for the child; if the child is born to a mother who is a temporary migrant (externally resident) and the child is considered to be a member of the household, a period of external residency ensues for the child. Residency episodes (whether internal or external) are of course terminated by the death of the individual, if that happens whilst the individual is under surveillance.

All SAPRIN nodes conducted a baseline household census at their beginning and all individuals enrolled at this point start their surveillance episode with enumeration. However, nodes may extend their area of surveillance at certain points after the initial household census, by doing another baseline census in these new areas, and all individuals enrolled then, also start their surveillance episodes with enumeration. For integrity in the longitudinal surveillance of individuals, the identity of newly encountered individuals is checked against the database and merged with prior records if the individual is already in the database. In the case of newly incorporated areas into the surveillance area, it is entirely possible to find individuals that have previously out-migrated from the surveillance area to reside in this new area, in such cases an individual will have more than one surveillance episode that starts with enumeration, their enumeration in the original baseline census as well as their enumeration in the newly extended surveillance area.

This dataset represents a snapshot of the continually evolving data in the underlying longitudinal databases maintained by the SAPRIN nodes. In these databases the rightmost extent of the individual's surveillance episode is indicated by the data collection date of the last time the individual's membership of a household under surveillance has been confirmed. Each dataset has a right censor date (31 December 2017 for the current version of the dataset) and individual surveillance episodes are terminated at that point if the individual is still under surveillance beyond the cut-off date.

Each individual surveillance episode is associated with a physical location, for internal residency episodes it is the actual place of residence of the individual, for external residence episodes (periods of temporary migration) it is the place of residence of the individual's household. If an individual change their place of residency from one location within the surveillance area to another location still within the surveillance area, the episode at the original location is terminated with a location exit event, and a new episode starts with a location entry event at the destination location. It is also possible for the household the individual is a member of, to change their place of residency in the surveillance area, whilst the individual is externally resident (is a temporary migrant), in which case the individual's external resident episode will also be split with a location exit-entry pair of events.

At every household visit written consent is obtained from the household respondent for continued participation in the surveillance and such consent can be withdrawn. When this happens all household members' surveillance episodes are terminated with a refusal event. It is possible for households to again provide consent to participate in the surveillance after some time, in such cases surveillance events are restarted with a permission event.

As mentioned previously, surveillance episodes are continually extended by the last data collection event if the individual remains under surveillance. In certain cases, individuals may be lost to follow-up and surveillance episodes where the date of last data collection is more than one year prior to the right censor data are terminated as lost to follow up at that last data collection date. Individuals with data collection dates within a year of the right censor date is considered still to be under surveillance up to this last data collection date.

Each surveillance episode contains the identifier of the household the individual is a member of during that episode. Under relatively rare circumstances it is possible for an individual to change household membership whilst still resident at the same location, or to change membership whilst externally resident, in these cases the surveillance episode will be split with a pair of membership end and membership start events. More commonly membership start and end events coincide with location exit and entry events or in- and out-migration events. Memberships also obviously start at birth or enumeration and end at death, refusal to participate or lost to follow-up.

In about half of the cases, individuals have a single episode from first enumeration, birth or in-migration, to their eventual death, out-migration or currently still under surveillance. In the remaining cases, individuals transition from internal residency to external residency via out-migration, or from one location to another via internal migration with a location exit and entry event, or some other rarer form of transition involving membership change, refusal or lost to follow-up. Usually these series of surveillance episodes are continuous in time, with no gaps between episodes, but gaps can form, e.g. when an individual out-migrates and end membership with the household and so is no longer under surveillance, only to return via in-migration at some future date and take up membership with same or different household.

The SAPRIN Individual Surveillance Episodes 2020 Datasets consists of three types of the Demographic surveillance datasets:

- 1.SAPRIN Individual Surveillance Episodes 2020: Basic Dataset. This dataset contains only the internal and external residency episodes for an individual.
- 2.SAPRIN Individual Surveillance Episodes 2020: Age-Year-Delivery Dataset. This dataset splits the basic surveillance episodes at calendar year end and at the date when the age in years (birth-day) of an individual changes. In the case of women who have given births, episodes are split at the time of delivery as well.
- 3.SAPRIN Individual Surveillance Episodes 2020: Detailed Dataset. This dataset adds to the dataset 2 time-varying attributes such as education, employment, marital status and socio-economic status.

KIND OF DATA

Event history data

UNIT OF ANALYSIS

Exposure Episodes

Version

VERSION DESCRIPTION

v1: Dataset for public distribution.

VERSION DATE

2020-06-01

VERSION NOTES

v1: Dataset for public distribution.

Scope

NOTES

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 or marital status, will likewise result in a new episode on the date of the change.

TOPICS

Topic
Fertility, Mortality, Migration

KEYWORDS

Keyword
Fertility, Mortality, Migration

Coverage

GEOGRAPHIC COVERAGE

The South African Population Research Infrastructure Network (SAPRIN) currently represents a network of three Health and Demographic Surveillance System (HDSS) nodes located in rural South Africa, namely: 1) MRC/Wits University Agincourt HDSS in Bushbuckridge District, Mpumalanga, which has collected data since 1993. The nodal website is: <http://www.agincourt.co.za>; 2) the University of Limpopo DIMAMO HDSS in the Capricorn District of Limpopo, which has collected data since 1996. The nodal website is: N/A; 3) and the Africa Health Research Institute (AHRI) HDSS in uMkhanyakude District, KwaZulu-Natal, which has collected data since 2000. The nodal website is: <http://www.ahri.org>. The Agincourt HDSS covers a surveillance area of approximately 420 square kilometres and is located in the Bushbuckridge District, Mpumalanga in the rural northeast of South Africa close to the Mozambique border. At baseline in 1992, 57 600 people were recorded in 8900 households in 20 villages; by 2006, the population had increased to about 70 000 people in 11 700 households. As of December 2017, there were 113 113 people under surveillance of whom 28% were not resident within the surveillance area, with a total of about 2m person years of observation. 33% of the population is under 15 years old. The population is almost exclusively Shangaan-speaking. The Agincourt HDSS has population density of over 200 persons per square kilometre. The Agincourt HDSS extends between latitudes 24° 50' and 24° 56' S and longitudes 31°08' and 31° 25' E. The altitude is about 400-600m above sea level. DIMAMO is located in the Capricorn district, Limpopo Province approximately 40 kilometres from Polokwane, the capital city of Limpopo Province and 15-50 kilometres from the University of Limpopo. The site covers an area of approximately 400 square kilometres. The initial total population observed was about 8 000 but the field site was expanded in 2010. As of December 2017, there were 38 479 people under surveillance, of whom 22% were not resident within the surveillance area, with about 400,000 person years of observation. 30% of the population is under 15 years old. The population is predominantly Sotho speaking. Most households have electricity. Some households have piped water either inside the house or in their yards, but most fetch water from taps situated at strategic points in the villages. Most households have a pit latrine in their yards. The area lies between latitudes and 23°65' and 23°90' S and longitudes 29°65' and 29°85' E. The HDSS is located on a high plateau area (approximately 1250 m above sea level) where communities typically consist of households clustered in villages, with access to local land for small-scale food production. Africa Health Research Institute (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 for portions where the Kwamsane township straddles the highway) and in the north by the Inyalazi river for portions of the boundary. The surveillance area is approximately 850 square kilometres. As of December 2017, there were 139 250 people under surveillance of whom 28% were not resident within the surveillance area, with about 1.7m person years of observation. 32% of the population is under 15 years old. The population is almost exclusively Zulu-speaking. The surveillance area is typical of many rural areas of South Africa in that while predominantly rural, it contains an urban township and informal peri-urban settlements. The area is characterized by

large variations in population densities (20-3000 people per square kilometre). The area lies between latitudes -28°24' and 28°20'N and longitudes 32°10' and 31°58'E.

UNIVERSE

Households resident in dwellings within the study area will be eligible for inclusion in the household component of SAPRIN. All individuals identified by the household proxy informant as a member of the household will be enumerated. A resident household member is an individual that intends to sleep the majority of time at the dwelling occupied by the household over a four-month period. Households will 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 most of the time. They 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.

Producers and sponsors

PRIMARY INVESTIGATORS

Name	Affiliation
Prof Mark Collinson	SAPRIN
Dr Kobus Herbst	SAPRIN
Prof Steve Tollman	Agincourt
Dr Eric Maimela	DIMAMO
Prof Willem Hanekom	AHRI

PRODUCERS

Name	Affiliation	Role
Taurayi Mudzana	SAPRIN	Technical Assistance
Tinofa Mutevedzi	SAPRIN	Technical Assistance
Chodziwadziwa Kabudula	Agincourt	Technical Assistance
Joseph Tlouyamma	DIMAMO	Technical Assistance
Dickman Gareta	AHRI	Technical Assistance

FUNDING AGENCY/SPONSOR

Name	Abbreviation	Role
Department of Science,Technology and Innovation	DSTI	Current Funder

OTHER IDENTIFICATIONS/ACKNOWLEDGMENTS

Name	Affiliation	Role
Agincourt Data Team	Agincourt	Providing Data
DIMAMO Data Team	DIMAMO	Providing Data
AHRI Data Team	AHRI	Providing Data
Steve Tollman	Agincourt	
Eric Maimela	DIMAMO	
Willem Hanekom	AHRI	
Centre for High Performance Computing	Centre for High Performance Computing	Providing IT Infrastructure for Data Processing

Sampling

SAMPLING PROCEDURE

This dataset is not based on a sample but contains information from the complete demographic surveillance areas.

Data collection

DATES OF DATA COLLECTION

Start	End	Cycle
1993-01-01	2017-12-31	Agincourt
1996-01-01	2017-12-31	DIMAMO
2000-01-01	2017-12-31	AHRI

TIME PERIODS

Start date	End date	Cycle
1993-01-01	2017-12-31	Agincourt
1996-01-01	2017-12-31	DIMAMO
2000-01-01	2017-12-31	AHRI

DATA COLLECTION NOTES

In all the HDSS nodes, data are collected from a household proxy respondent, preferably the head of household or any next available senior adult resident household member, after informed consent was obtained by trained fieldworkers. Respondents are informed of the purpose and confidentiality of the interview, their right to refuse participation or withdraw from the study, and that scientists would be given access to anonymised data to analyse and publish information. Informed consent was verbal in all HDSS nodes until 2016. Written informed consent started in 2017 in AHRI, and 2018 in DIMAMO and 2019 in Agincourt. Until 2016 for Agincourt and AHRI, and 2017 for DIMAMO, data collection was field-based 'paper and pen' personal interviews (PAPI), before changing to field-based computer-assisted personal interviews (CAPI). Since 2019, all SAPRIN HDSS nodes collect data in 3 annual rounds over a 45-week data collection schedule; one field-based CAPI round, sandwiched on either side by a Call-Centre-based computer assisted telephonic interview (CATI), to create 3 data points at an interval of approximately 4 months in each calendar year. In the past HDSS nodes had different data collection frequencies. AHRI data collection was 2 PAPI rounds per year from inception to 2011, changing to 3 PAPI rounds per year between 2012 and 2016, before becoming 1 PAPI round and 2 CATI rounds from 2017. Agincourt and DIMAMO have been collecting data once annually in a census-type format, over 4-5-month period until 2018.

Questionnaires

QUESTIONNAIRES

The data on this Repository is not the result of a single questionnaire but is a result of harmonised data from three different sites longitudinally collected over more than twenty years using different questionnaires that varied over time and site.

Data Processing

DATA EDITING

The first step in the data preparation process is quality assurance. The SAPRIN Management hub team assess the data submitted to ensure it is in the correct format and falls within expected value ranges. Other potential issues checked include: missing data, incorrect data types, unexpected duplicate or orphan records. The SAPRIN Management hub assess this conversion by running both original operational database and the SAPRIN database created from the operational database through the iSHARE data quality assessment and indicator process. The data quality checking process is conducted using Pentaho Data Integration (PDI). PDI provides the Extract, Transform, and Load (ETL) capabilities that facilitates the process of capturing, cleansing, and storing data using a uniform and consistent format that is accessible and relevant to

end users. The principle of the data quality checks is that if the data conversion conducted by the nodes was complete and accurate, there should be little or no difference in the data quality and demographic indicators between the base and SAPRIN versions of the nodal data. If the data submitted by the nodes meets the criteria for inclusion into the consolidated dataset the data moves to the second step of the data production process. However, if the data fail the inclusion checks, this could then lead to another iteration of data submission and quality control checks until SAPRIN Management hub is satisfied that they have high quality data. To produce this final standard dataset, the data is processed using PDI on the Centre for High Performance Computing cluster.

Data Appraisal

ESTIMATES OF SAMPLING ERROR

Not Applicable

Access policy

CONTACTS

Name	Affiliation	Email	URL
SAPRIN Data Management	SAPRIN	saprindata@mrc.ac.za	Link

ACCESS CONDITIONS

This data is made available for access under the following conditions:

- 1)The data and other materials provided by SAPRIN will not be redistributed or sold to other individuals, institutions, or organizations without the written agreement of SAPRIN.
- 2)The data will be used for statistical and scientific research purposes only. They will be used solely for reporting of aggregated information, and not for investigation of specific individuals or organisations. The Data User will neither use nor permit others to use the data in any way other than listed in the original application (Analysis Plan) for access to the dataset.
- 3)No attempt will be made to re-identify respondents, and no use will be made of the identity of any person or establishment discovered inadvertently. Any such discovery should immediately be reported to SAPRIN.
- 4)No attempt will be made to produce links among datasets provided by SAPRIN, or among data from SAPRIN and other datasets that could identify individuals or organizations.
- 5)The Data User will ensure that the data are kept in a secured environment and that only authorized users have access to the data.
- 6)Any books, articles, conference papers, theses, dissertations, reports, or other publications that employ data obtained from SAPRIN will cite the source of data in accordance with the Citation Requirement provided with each dataset.
- 7)An electronic copy of all reports and publications based on the requested data will be sent to SAPRIN.
- 8)The original collector of the data, SAPRIN, and relevant funding agencies bear no responsibility for use of the data or for interpretations or inferences based upon such uses.
- 9) Once the data set has served its indicated purpose it must be destroyed. If the dataset needs to be lodged for publication purposes, a reference (a digital object identifier will be maintained by SAPRIN for this purpose) to the original dataset on the SAPRIN data repository should be used. Derived or aggregated datasets produced from the original dataset do not fall within this provision and may be lodged as publication datasets. If the same dataset is needed for a different purpose, the dataset should be re-requested and the new purposes indicated.

CITATION REQUIREMENTS

Any use of this data must cite the digital object identifier (doi) associated with the appropriate dataset. Using the following form:

Mudzana, T; Mutevedzi, T; Herbst, K; Collinson, M (2020): SAPRIN Individual Surveillance Episodes 2020: Basic Dataset. South African Population Research Infrastructure Network.
<http://doi.org/10.23667/SAPRIN.SISEBD2020>

Mudzana, T; Mutevedzi, T; Herbst, K; Collinson, M (2020): SAPRIN Individual Surveillance Episodes 2020: Detailed Dataset. South African Population Research Infrastructure Network.
<http://doi.org/10.23667/SAPRIN.SISEDD2020>

Mudzana, T; Mutevedzi, T; Herbst, K; Collinson, M (2020): SAPRIN Individual Surveillance Episodes 2020: Age-Year-Delivery Dataset. South African Population Research Infrastructure Network.
<http://doi.org/10.23667/SAPRIN.SISEAYDD2020>

ACCESS AUTHORITY

Name	Affiliation	Email	URL
SAPRIN Data Management	SAPRIN	saprindata@mrc.ac.za	Link

Disclaimer and copyrights

DISCLAIMER

The user of the data acknowledges that the original collector of the data and the relevant funding agencies bear no responsibility for the data's use or interpretation or inferences based upon it.

COPYRIGHT

This dataset documentation is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License. The dataset is shared in terms of the data-use agreement accepted at the time of data download.

Metadata production

PRODUCERS

DATE OF METADATA PRODUCTION

2025-06-30

Data Dictionary

Data file	Cases	Variables
SISEBD2020V1 SAPRIN Individual Surveillance Episodes 2020: Basic Dataset	0	33
SISEDD2020V1 SAPRIN Individual Surveillance Episodes 2020: Detailed Dataset	0	56
SISEAYDD2020V1 SAPRIN Individual Surveillance Episodes 2020: Age Year Delivery Dataset	0	42

Data file: SISEBD2020V1

SAPRIN Individual Surveillance Episodes 2020: Basic Dataset

Cases: 0

Variables: 33

Variables

ID	Name	Label	Question
V1	Nodeld	Nodeld	
V2	IndividualId	IndividualId	
V3	Sex	Sex	
V4	DoB	DoB	
V5	DoD	DoD	
V6	MotherId	MotherId	
V7	FatherId	FatherId	
V8	Episode	Episode	
V9	Episodes	Episodes	
V10	StartDate	StartDate	
V11	EndDate	EndDate	
V12	Days	Days	
V13	LocationId	LocationId	
V14	HouseholdId	HouseholdId	
V15	Resident	Resident	
V16	Born	Born	
V17	Enumeration	Enumeration	
V18	InMigration	InMigration	
V19	LocationEntry	LocationEntry	
V20	ExtResStart	ExtResStart	
V21	Participation	Participation	
V22	Died	Died	
V23	OutMigration	OutMigration	
V24	LocationExit	LocationExit	
V25	ExtResEnd	ExtResEnd	
V26	Refusal	Refusal	
V27	LostToFollowUp	LostToFollowUp	
V28	Current	Current	
V29	MembershipStart	MembershipStart	
V30	MembershipEnd	MembershipEnd	
V31	Memberships	Memberships	
V32	GapStart	GapStart	
V33	GapEnd	GapEnd	

Total: 33

Data file: SISED2020V1

SAPRIN Individual Surveillance Episodes 2020: Detailed Dataset

Cases: 0

Variables: 56

Variables

ID	Name	Label	Question
V34	Nodeld	Nodeld	
V35	IndividualId	IndividualId	
V36	MotherId	MotherId	
V37	FatherId	FatherId	
V38	Sex	Sex	
V39	DoB	DoB	
V40	DoD	DoD	
V41	Age	Age	
V42	Year	Year	
V43	Episodes	Episodes	
V44	Episode	Episode	
V45	StartDate	StartDate	
V46	EndDate	EndDate	
V47	Days	Days	
V48	Resident	Resident	
V49	Gap	Gap	
V50	Enumeration	Enumeration	
V51	Born	Born	
V52	Participation	Participation	
V53	InMigration	InMigration	
V54	LocationEntry	LocationEntry	
V55	ExtResStart	ExtResStart	
V56	Died	Died	
V57	Refusal	Refusal	
V58	LostToFollowUp	LostToFollowUp	
V59	Current	Current	
V60	OutMigration	OutMigration	
V61	LocationExit	LocationExit	
V62	ExtResEnd	ExtResEnd	
V63	MembershipStart	MembershipStart	
V64	MembershipEnd	MembershipEnd	
V65	GapStart	GapStart	
V66	GapEnd	GapEnd	
V67	LocationId	LocationId	
V68	HouseholdId	HouseholdId	
V69	HHRelation	HHRelation	
V70	Memberships	Memberships	

ID	Name	Label	Question
V71	Education	Education	
V72	MaritalStatus	MaritalStatus	
V73	SEIdx	SEIdx	
V74	CurrentEmployment	CurrentEmployment	
V75	EmploymentType	EmploymentType	
V76	MotherCoResident	MotherCoResident	
V77	MotherDead	MotherDead	
V78	FatherCoResident	FatherCoResident	
V79	FatherDead	FatherDead	
V80	DwellingIdx	DwellingIdx	
V81	LifeStockIdx	LifeStockIdx	
V82	ModerntAssetIdx	ModerntAssetIdx	
V83	PowerSupplyIdx	PowerSupplyIdx	
V84	WaterSanitationIdx	WaterSanitationIdx	
V85	EducationObsDate	EducationObsDate	
V86	MaritalStatusObsDate	MaritalStatusObsDate	
V87	SEObsDate	SEObsDate	
V88	EmploymentSector	EmploymentSector	
V89	EmploymentObsDate	EmploymentObsDate	

Total: 56

Data file: SISEAYDD2020V1

SAPRIN Individual Surveillance Episodes 2020: Age Year Delivery Dataset

Cases: 0

Variables: 42

Variables

ID	Name	Label	Question
V90	Nodeld	Nodeld	
V91	IndividualId	IndividualId	
V92	Sex	Sex	
V93	DoB	DoB	
V94	DoD	DoD	
V95	MotherId	MotherId	
V96	FatherId	FatherId	
V97	Age	Age	
V98	CalendarYear	CalendarYear	
V99	Episode	Episode	
V100	Episodes	Episodes	
V101	StartDate	StartDate	
V102	EndDate	EndDate	
V103	Days	Days	
V104	LocationId	LocationId	
V105	HouseholdId	HouseholdId	
V106	Resident	Resident	
V107	Born	Born	
V108	Enumeration	Enumeration	
V109	InMigration	InMigration	
V110	LocationEntry	LocationEntry	
V111	ExtResStart	ExtResStart	
V112	Participation	Participation	
V113	Died	Died	
V114	OutMigration	OutMigration	
V115	LocationExit	LocationExit	
V116	ExtResEnd	ExtResEnd	
V117	Refusal	Refusal	
V118	LostToFollowUp	LostToFollowUp	
V119	Current	Current	
V120	MembershipStart	MembershipStart	
V121	MembershipEnd	MembershipEnd	
V122	Memberships	Memberships	
V123	GapStart	GapStart	
V124	GapEnd	GapEnd	
V125	AgeStart	AgeStart	
V126	YrStart	YrStart	

ID	Name	Label	Question
V127	AgeEnd	AgeEnd	
V128	YrEnd	YrEnd	
V129	Delivery	Delivery	
V130	ChildrenBorn	ChildrenBorn	
V131	ChildrenEverBorn	ChildrenEverBorn	

Total: 42

NODEID: NodeId**Data file: SISEBD2020V1****Overview**

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: 1 - 3 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
1	Agincourt
2	DIMAMO
3	AHRI

Description

DEFINITION

SAPRIN Node Identifier 1=Agincourt; 2=DIMAMO; 3=AHRI

STARTDATE: StartDate**Data file: SISEBD2020V1****Overview**

Valid: 0 Invalid: 0

Type: Discrete Width: 11 Range: - Format: character

Description

DEFINITION

Start date of episode (inclusive) Format yyyy-MM-dd in csv file

INDIVIDUALID: IndividualId**Data file: SISEBD2020V1****Overview**

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Width: 12 Range: - Format: Numeric

Description

DEFINITION

Unique anonymised individual identifier. Only unique within nodal dataset

SEX: Sex**Data file: SISEBD2020V1****Overview**

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: 1 - 3 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
1	Male
2	Female
3	Unknown

Description

DEFINITION

Sex of the individual. 0= Missing 1 = Male 2 = Female 3 = Unknown

DOB: DoB**Data file: SISEBD2020V1****Overview**

Valid: 0 Invalid: 0

Type: Discrete Width: 11 Range: - Format: character

Description

DEFINITION

Date of birth. Format yyyy-MM-dd in csv file Date of Birth 1800-01-01 is unknown

DOD: DoD**Data file: SISEBD2020V1****Overview**

Valid: 0 Invalid: 0

Type: Discrete Width: 11 Range: - Format: character

Description

DEFINITION

Date of death Format yyyy-MM-dd in csv file

MOTHERID: MotherId

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Width: 12 Range: - Format: Numeric

Description

DEFINITION

Mother's IndividualId 0 = Unknown

FATHERID: FatherId

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Width: 12 Range: - Format: Numeric

Description

DEFINITION

Father's IndividualId 0 = Unknown

EPISODE: Episode

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Width: 12 Range: - Format: Numeric

Description

DEFINITION

This episode number (first=1, last=Episodes)

EPISODES: Episodes

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0
 Type: Continuous Decimal: 0 Width: 12 Range: - Format: Numeric

Description

DEFINITION

Total number of episodes for individual

ENDDATE: EndDate

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0
 Type: Discrete Width: 11 Range: - Format: character

Description

DEFINITION

End date of episode (inclusive) Format yyyy-MM-dd in csv file

DAYS: Days

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0
 Type: Continuous Decimal: 0 Width: 12 Range: - Format: Numeric

Description

DEFINITION

Duration in days of episode $EndDate - StartDate + 1$

LOCATIONID: LocationId

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0
 Type: Continuous Decimal: 0 Width: 12 Range: - Format: Numeric

Description

DEFINITION

Where individual was resident, household residence if non-resident. For internal residency episodes it is the actual place of residence, for external residency episodes, it is the place of residence of the individual's household.

HOUSEHOLDID: HouseholdId

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Width: 12 Range: - Format: Numeric

Description

DEFINITION

Unique household identifier of the household the individual is a member of

RESIDENT: Resident

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: 1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
1	Internal residence;

Description

DEFINITION

Whether individual is resident for duration of episode 1 = Internal residence; 0 = External residence

BORN: Born

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode starts with the birth of the individual. Can start an internal or external residency episode depending on the residency status of the mother, if the mother is an internal resident, then the child is as well, if the mother is externally resident then child is as well. If the child becomes resident shortly after birth, then the child will be in-migrated and start a residency episode at that point.

ENUMERATION: Enumeration

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode starts with an enumeration. Can start an internal or external residency episode depending on the usual place of residency of the individual.

INMIGRATION: InMigration

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-------	----------

-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode starts with an in-migration. Either as a new first episode for the individual; or following the end of an external residency; or after a gap in residency.

LOCATIONENTRY: LocationEntry

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode starts with an internal migration. Individual changed residency from one location to another for internal residency episodes; for external residency episodes, the household the individual is a member of, changed residency from one location to another

EXTRESSTART: ExtResStart

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

External residency start. Individual out-migrated, but retained membership in the household, in other words the start of a temporary migration episode. If this is the start of the first episode for the individual then the household reported a new household member that is currently externally resident.

PARTICIPATION: Participation

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Resume participation after refusal. The individual (or more specifically the individual's household) re-consented to participate after a previous refusal

DIED: Died

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode ends with the death of the individual

OUTMIGRATION: OutMigration

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode ends with the out-migration. If the individual retains membership in a household under surveillance this will be followed by an episode of external residence (temporary migration), otherwise this may be the start of a gap in the surveillance episodes of the individual; or it may be the end of the last surveillance episode for the individual.

LOCATIONEXIT: LocationExit

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode ends with an internal migration. Individual changed residency from one location to another for internal residency episodes; for external residency episodes, the household the individual is a member of, changed residency from one location to another. LocationExit and LocationEntry always occurs in pairs.

EXTRESEND: ExtResEnd

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

External residency end. Individual in-migrated after a period of external residency; or the individual is no longer a member of any household under surveillance.

REFUSAL: Refusal

Data file: SISEBD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Individual refused follow-up. The household the individual is a member of refused further participation in the surveillance.

LOSTTOFOLLOWUP: LostToFollowUp**Data file: SISEBD2020V1****Overview**

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Individual was lost to follow-up at the end of the episode. Individual was last reported on by the household respondent more than a year before the current right censor date.

CURRENT: Current**Data file: SISEBD2020V1****Overview**

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Individual still under surveillance. Either last reported on by the household respondent after the right censor date, or within a year prior to the right censor date.

MEMBERSHIPSTART: MembershipStart**Data file: SISEBD2020V1****Overview**

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Flag to indicate start of household membership. This flag can be set in conjunction with other start flags, in which case it means that a new membership started at the same time. If none of the other entry flags are set it means that the individual changed household membership without changing residency status or place of residence.

MEMBERSHIPEND: MembershipEnd**Data file: SISEBD2020V1****Overview**

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Flag to indicate end of household membership. This flag can be set in conjunction with other end flags, in which case it means that the membership ended at the same time. If none of the other exit flags are set it means that the individual changed household membership without changing residency status or place of residence

MEMBERSHIPS: Memberships**Data file: SISEBD2020V1****Overview**

Valid: 0 Invalid: 0
 Type: Discrete Decimal: 0 Width: 12 Range: - Format: Numeric

Description

DEFINITION

Number of concurrent household memberships at the start of the episode.

GAPSTART: GapStart**Data file: SISEBD2020V1****Overview**

Valid: 0 Invalid: 0
 Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Start of gap in individual exposure. E.g. between out-migration without retaining any household membership and a subsequent in-migration taking up household membership and residency in the surveillance area again

GAPEND: GapEnd**Data file: SISEBD2020V1****Overview**

Valid: 0 Invalid: 0
 Type: Discrete Decimal: 0 Width: 12 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-------	----------

-1	Unknown/Missing
1	Yes

Description

DEFINITION

End of gap in individual exposure. GapStart and GapEnd flags always occur in conjunction with the other entry or exit flags.

NODEID: Nodeld**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Range: 1 - 3 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
1	Agincourt
2	DIMAMO
3	AHRI

Description

DEFINITION

SAPRIN Node Identifier 1=Agincourt; 2=DIMAMO; 3=AHRI

INDIVIDUALID: IndividualId**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Unique anonymised individual identifier. Only unique within nodal dataset

MOTHERID: MotherId**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Mother's IndividualId 0 = Unknown

FATHERID: FatherId**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0
 Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Father's IndividualId 0 = Unknown

SEX: Sex**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0
 Type: Discrete Decimal: 0 Range: 1 - 3 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
1	Male
2	Female
3	Unknown

Description

DEFINITION

Sex of the individual 0= Missing 1 = Male 2 = Female 3 = Unknown

DOB: DoB**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0
 Type: Discrete Width: 9 Range: - Format: character

Description

DEFINITION

Date of birth. Format yyyy-MM-dd in csv file Date of Birth 1800-01-01 is unknown

DOD: DoD

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Width: 9 Range: - Format: character

Description

DEFINITION

Date of Death. Format yyyy-MM-dd in csv file

AGE: Age

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Age in completed years (valid for the duration of the episode)

YEAR: Year

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Calendar year during episode

EPISODES: Episodes

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0
 Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Total number of episodes for individual

EPISODE: Episode

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0
 Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

This episode number (first=1, last=Episodes)

STARTDATE: StartDate

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0
 Type: Discrete Width: 9 Range: - Format: character

Description

DEFINITION

Start date of episode (inclusive). Format yyyy-MM-dd in csv file

ENDDATE: EndDate

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0
 Type: Discrete Width: 9 Range: - Format: character

Description

DEFINITION

End date of episode (inclusive). Format yyyy-MM-dd in csv file

DAYS: Days

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0
 Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Duration in days of the episode. EndDate - StartDate + 1

RESIDENT: Resident

Data file: SISED2020V1

Overview

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

Questions and instructions

CATEGORIES

Value	Category
1	Internal residence;

Description

DEFINITION

Whether individual is resident for duration of episode 1 = Internal residence; 0 = External residence

GAP: Gap

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0
 Type: Discrete Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Gap in individual exposure

ENUMERATION: Enumeration**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode starts with an enumeration

BORN: Born**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Whether this episode starts with the birth of the individual. Can start an internal or external residency episode depending on the residency status of the mother, if the mother is an internal resident, then the child is as well, if the mother is externally resident then child is as well. If the child becomes resident shortly after birth, then the child will be in-migrated and start a residency episode at that point.

PARTICIPATION: Participation**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0
 Type: Discrete Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Resume participation after refusal. The individual (or more specifically the individual's household) re-consented to participate after a previous refusal

INMIGRATION: InMigration**Data file: SISED2020V1****Overview**

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Whether this episode starts with an in-migration. Either as a new first episode for the individual; or following the end of an external residency; or after a gap in residency

LOCATIONENTRY: LocationEntry**Data file: SISED2020V1****Overview**

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

Questions and instructions

CATEGORIES

Value	Category
-------	----------

-1	Unknown/Missing
1	Yes

Description

DEFINITION

Whether this episode starts with an internal migration into the location. Individual changed residency from one location to another for internal residency episodes; for external residency episodes, the household the individual is a member of, changed residency from one location to another

EXTRESSTART: ExtResStart

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

External residency starts Individual out-migrated, but retained membership in the household, in other words the start of a temporary migration episode. If this is the start of the first episode for the individual then the household reported a new household member that is currently externally resident.

DIED: Died

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing

1	Yes
---	-----

Description

DEFINITION

Whether the episode ends with the death of the individual.

REFUSAL: Refusal

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Surveillance end resulting from household refusal. The household the individual is a member of refused further participation in the surveillance

LOSTTOFOLLOWUP: LostToFollowUp

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Individual was lost to follow-up at the end of the episode. Individual was last reported on by the household respondent more than a year before the current right censor date.

CURRENT: Current

Data file: SISEDD2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Individual still under surveillance. Either last reported on by the household respondent after the right censor date, or within a year prior to the right censor date.

OUTMIGRATION: OutMigration

Data file: SISEDD2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Whether the episode ends with the out-migration of the individual. If the individual retains membership in a household under surveillance this will be followed by an episode of external residence (temporary migration), otherwise this may be the start of a gap in the surveillance episodes of the individual; or it may be the end of the last surveillance episode for the individual.

LOCATIONEXIT: LocationExit

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode ends with an internal migration. Individual changed residency from one location to another for internal residency episodes; for external residency episodes, the household the individual is a member of, changed residency from one location to another. LocationExit and LocationEntry always occurs in pairs.

EXTRESEND: ExtResEnd

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

External residency end. Individual in-migrated after a period of external residency; or the individual is no longer a member of

any household under surveillance.

MEMBERSHIPSTART: MembershipStart

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Flag to indicate start of household membership. This flag can be set in conjunction with other start flags, in which case it means that a new membership started at the same time. If none of the other entry flags are set it means that the individual changed household membership without changing residency status or place of residence.

MEMBERSHIPEND: MembershipEnd

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Flag to indicate end of household membership. This flag can be set in conjunction with other end flags, in which case it means that the membership ended at the same time. If none of the other exit flags are set it means that the individual changed household membership without changing residency status or place of residence

GAPSTART: GapStart**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Start of gap in individual exposure. E.g. between out-migration without retaining any household membership and a subsequent in-migration taking up household membership and residency in the surveillance area again

GAPEND: GapEnd**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

End of gap in individual exposure. GapStart and GapEnd flags always occur in conjunction with the other entry or exit flags

LOCATIONID: LocationId**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Where individual is resident for the duration of this episode. For internal residency episodes it is the actual place of residence, for external residency episodes, it is the place of residence of the individual's household

HOUSEHOLDID: HouseholdId**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Unique household identifier of the household the individual is a member of during this episode

HHRELATION: HHRelation**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Range: 1 - 12 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
1	Head
2	Spouse
3	Child
4	Son/Daughter in law
5	Grandchild
6	Parent
7	Parant in law

8	Grand parent
9	Sibling
10	Other Relative
11	Domestic Worker/tenant
12	Unrelated/other

Description

DEFINITION

Relationship to the head of the household.

MEMBERSHIPS: Memberships

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Number of concurrent household memberships. At the start of the episode

EDUCATION: Education

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Education status of individual during this episode. 0. None - could be missing or unknown too 1. Grade 1 2. Grade 2 3. Grade 3 4. Grade 4 5. Grade 5 6. Grade 6 7. Grade 7 8. Grade 8 9. Grade 9 10. Grade 10 11. Grade 11 12. Grade 12 13. Some tertiary qualification 14. Some tertiary qualification

MARITALSTATUS: MaritalStatus

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
1	Not Married
2	Married
3	Widowed
4	Separated/divorced
5	Informal union

Description

DEFINITION

Marital status of the individual during this episode. 0 = Unknown 1 = Not Married 2 = Married 3 = Widowed 4 = Separated/divorced 5 = Informal Union

SEIDX: SEIdx

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0
Type: Discrete Width: 17 Range: - Format: character

Description

DEFINITION

Socio-economic index of the household the individual belongs to. Calculated using the approach described in the paper by Kabudula, et al.

CURRENEMPLOYMENT: CurrentEmployment

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0
Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Employment status of the individual during this episode.

EMPLOYMENTTYPE: EmploymentType**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Range: 200 - 203 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
200	Unknown
201	Employee
202	Self Employed
203	Informally Employed

Description

DEFINITION

Employment type of the individual during this episode

MOTHERCORESIDENT: MotherCoResident**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Whether mother is co-resident during episode, yes, no or unknown (if MotherId = 0)

MOTHERDEAD: MotherDead**Data file: SISED2020V1**

Overview

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Whether mother is dead for the duration of this episode

FATHERCORESIDENT: FatherCoResident

Data file: SISED2020V1

Overview

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Whether father is co-resident during episode, yes, no or unknown (if FatherId = 0)

FATHERDEAD: FatherDead

Data file: SISED2020V1

Overview

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Whether father is dead for the duration of this episode

DWELLINGIDX: DwellingIdx

Data file: SISEDD2020V1

Overview

Valid: 0 Invalid: 0
 Type: Discrete Width: 17 Range: - Format: character

Description

DEFINITION

The dwelling SE Index. Not available in the AHRI data. Based on the construction material and size of the dwelling

LIFESTOCKIDX: LifeStockIdx

Data file: SISEDD2020V1

Overview

Valid: 0 Invalid: 0
 Type: Discrete Width: 3 Range: - Format: character

Description

DEFINITION

The livestock SE Index. Not available in the AHRI data. Based on the construction material and size of the dwelling

MODERN TASSETIDX: ModerntAssetIdx

Data file: SISEDD2020V1

Overview

Valid: 0 Invalid: 0
 Type: Discrete Width: 17 Range: - Format: character

Description

DEFINITION

The Asset SE Index. Based on the ownership of a list of various assets.

POWERSUPPLYIDX: PowerSupplyIdx

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Width: 2 Range: - Format: character

Description

DEFINITION

The power supply SE Index. Based on the energy source used for lighting and cooking.

WATERSANITATIONIDX: WaterSanitationIdx

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Width: 4 Range: - Format: character

Description

DEFINITION

The water and sanitation SE Index. Based on the energy source used for lighting and cooking.

EDUCATIONOBSDATE: EducationObsDate

Data file: SISED2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Width: 9 Range: - Format: character

Description

DEFINITION

The date on which the education status was recorded. Format yyyy-MM-dd in csv file

MARITALSTATUSOBSDATE: MaritalStatusObsDate**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0
 Type: Discrete Width: 9 Range: - Format: character

Description

DEFINITION

The date on which the marital status was recorded Format yyyy-MM-dd in csv file

SEOBSDATE: SEObsDate**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0
 Type: Discrete Width: 9 Range: - Format: character

Description

DEFINITION

The date on which the socio-economic status was recorded Format yyyy-MM-dd in csv file

EMPLOYMENTSECTOR: EmploymentSector**Data file: SISED2020V1****Overview**

Valid: 0 Invalid: 0
 Type: Continuous Decimal: 0 Range: 300 - 308 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
300	Unknown
301	National Government
302	Provincial Government
303	Local Government
304	Public Corporation
305	Private Sector Employment
306	Non profit
307	Self Employment

308

Household

Description

DEFINITION

Employment sector of the individual during this episode

EMPLOYMENTOBSDATE: EmploymentObsDate

Data file: SISEDD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Width: 9 Range: - Format: character

Description

DEFINITION

Date on which the employment status was recorded Format yyyy-MM-dd in csv file

NODEID: Nodeld**Data file: SISEAYDD2020V1****Overview**

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Range: 1 - 3 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
1	Agincourt
2	DIMAMO
3	AHRI

Description

DEFINITION

SAPRIN Node Identifier. 1=Agincourt;2=DIMAMO;3=AHRI

INDIVIDUALID: IndividualId**Data file: SISEAYDD2020V1****Overview**

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Unique anonymised individual identifier Only unique within nodal dataset

SEX: Sex**Data file: SISEAYDD2020V1****Overview**

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Range: 1 - 3 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category
-------	----------

1	Male
2	Female
3	Unknown

Description

DEFINITION

Sex of the individual. 0=Missing 1 = Male 2 = Female 3 = Unknown

DOB: DoB

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Width: 9 Range: - Format: character

Description

DEFINITION

Date of birth. Format yyyy-MM-dd in csv file Date of Birth 1800-01-01 is unknown

DOD: DoD

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Width: 9 Range: - Format: character

Description

DEFINITION

Date of death. Format yyyy-MM-dd in csv file

MOTHERID: MotherId

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Mother's IndividualId. 0 = Unknown

FATHERID: FatherId

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Father's IndividualId. 0 = Unknown

AGE: Age

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

The age in completed years of the individual

CALENDARYEAR: CalendarYear

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

The calendar year in which the episode falls

EPISODE: Episode

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0
 Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

This episode number (first=1, last=Episodes)

EPISODES: Episodes

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0
 Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Total number of episodes for individual

STARTDATE: StartDate

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0
 Type: Discrete Width: 9 Range: - Format: character

Description

DEFINITION

Start date of episode (inclusive). Format yyyy-MM-dd in csv file

ENDDATE: EndDate

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0
 Type: Discrete Width: 9 Range: - Format: character

Description

DEFINITION

End date of episode (inclusive). Format yyyy-MM-dd in csv file

DAYS: Days

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0
 Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Duration in days of episode $\text{EndDate} - \text{StartDate} + 1$

LOCATIONID: LocationId

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0
 Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Where individual was resident, household residence if non-resident. For internal residency episodes it is the actual place of residence, for external residency episodes, it is the place of residence of the individual's household

HOUSEHOLDID: HouseholdId

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0
 Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Unique household identifier of the household the individual is a member of

RESIDENT: Resident

Data file: SISEAYDD2020V1

Overview

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

Questions and instructions

CATEGORIES

Value	Category
1	Internal residence;

Description

DEFINITION

Whether individual is resident for duration of episode. 1 = Internal residence; 0 = External residence

BORN: Born

Data file: SISEAYDD2020V1

Overview

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode starts with the birth of the individual. Can start an internal or external residency episode depending on the residency status of the mother, if the mother is an internal resident, then the child is as well, if the mother is externally resident then child is as well. If the child becomes resident shortly after birth, then the child will be in-migrated and start a residency episode at that point

ENUMERATION: Enumeration

Data file: SISEAYDD2020V1

Overview

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode starts with an enumeration, Can start an internal or external residency episode depending on the usual place of residency of the individual

INMIGRATION: InMigration

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode starts with an in-migration. Either as a new first episode for the individual; or following the end of an external residency; or after a gap in residency

LOCATIONENTRY: LocationEntry

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode starts with an internal migration. Individual changed residency from one location to another for internal residency episodes; for external residency episodes, the household the individual is a member of, changed residency from one location to another

EXTRESSTART: ExtResStart

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

External residency start. Individual out-migrated, but retained membership in the household, in other words the start of a temporary migration episode. If this is the start of the first episode for the individual then the household reported a new household member that is currently externally resident.

PARTICIPATION: Participation

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Resume participation after refusal. The individual (or more specifically the individual's household) re-consented to

participate after a previous refusal

DIED: Died

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode ends with the death of the individual

OUTMIGRATION: OutMigration

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode ends with the out-migration. If the individual retains membership in a household under surveillance this will be followed by an episode of external residence (temporary migration), otherwise this may be the start of a gap in the surveillance episodes of the individual; or it may be the end of the last surveillance episode for the individual.

LOCATIONEXIT: LocationExit**Data file: SISEAYDD2020V1****Overview**

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Episode ends with an internal migration. Individual changed residency from one location to another for internal residency episodes; for external residency episodes, the household the individual is a member of, changed residency from one location to another. LocationExit and LocationEntry always occurs in pairs

EXTRESEND: ExtResEnd**Data file: SISEAYDD2020V1****Overview**

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

External residency end. Individual in-migrated after a period of external residency; or the individual is no longer a member of any household under surveillance

REFUSAL: Refusal**Data file: SISEAYDD2020V1**

Overview

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Individual refused follow-up. The household the individual is a member of refused further participation in the surveillance

LOSTTOFOLLOWUP: LostToFollowUp

Data file: SISEAYDD2020V1

Overview

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Individual was lost to follow-up at the end of the episode. Individual was last reported on by the household respondent more than a year before the current right censor date.

CURRENT: Current

Data file: SISEAYDD2020V1

Overview

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Individual still under surveillance. Either last reported on by the household respondent after the right censor date, or within a year prior to the right censor date.

MEMBERSHIPSTART: MembershipStart

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Flag to indicate start of household membership. This flag can be set in conjunction with other start flags, in which case it means that a new membership started at the same time. If none of the other entry flags are set it means that the individual changed household membership without changing residency status or place of residence.

MEMBERSHIPEND: MembershipEnd

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Flag to indicate end of household membership. This flag can be set in conjunction with other end flags, in which case it means that the membership ended at the same time. If none of the other exit flags are set it means that the individual changed household membership without changing residency status or place of residence.

MEMBERSHIPS: Memberships

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0
 Type: Discrete Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Number of concurrent household memberships. At the start of the episode

GAPSTART: GapStart

Data file: SISEAYDD2020V1

Overview

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Start of gap in individual exposure. E.g. between out-migration without retaining any household membership and a subsequent in-migration taking up household membership and residency in the surveillance area again

GAPEND: GapEnd

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

End of gap in individual exposure. GapStart and GapEnd flags always occur in conjunction with the other entry or exit flags

AGESTART: AgeStart

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Flag episode start due to age change

YRSTART: YrStart**Data file: SISEAYDD2020V1****Overview**

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Flag episode start due to calendar year change

AGEEND: AgeEnd**Data file: SISEAYDD2020V1****Overview**

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Flag episode end due to age change

YREND: YrEnd**Data file: SISEAYDD2020V1**

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Flag episode end due to calendar year change

DELIVERY: Delivery

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

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

Questions and instructions

CATEGORIES

Value	Category
-1	Unknown/Missing
1	Yes

Description

DEFINITION

Flag a delivery to this woman at start of episode

CHILDRENBORN: ChildrenBorn

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

Type: Discrete Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Number of live born from delivery

CHILDRENEVERBORN: ChildrenEverBorn

Data file: SISEAYDD2020V1

Overview

Valid: 0 Invalid: 0

Type: Continuous Decimal: 0 Range: - Format: Numeric

Description

DEFINITION

Number of children ever born
