South Africa - SAPRIN Individual Surveillance Episodes 2023 Datasets

Dr Kobus Herbst, Prof Mark Collinson, Tinofa Mutevedzi, Prof Steve Tollman, Prof Eric Maimela, Prof Collins Iwuji

Report generated on: January 25, 2024

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

Overview

Identification

ID NUMBER
SAPRIN.SISED2023V2

Version

VERSION DESCRIPTION
v2: Dataset for public distribution.

PRODUCTION DATE
2024-01-22

NOTES
v2: Dataset for public distribution.

Overview

ABSTRACT
The ‘South African Population Research Infrastructure Network’ (SAPRIN) is a national research infrastructure funded through the Department of Science 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 current population of 105 461 people; the University of Limpopo DIMAMO HDSS in the Capricorn District of Limpopo, established in 1996, with a current population of 99 741; and the Africa Health Research Institute (AHRI) HDSS in uMkhanyakude District, KwaZulu-Natal, established in 2000, with a current population of 153 410.

For an individual to 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 extend 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 2022 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 2023 Datasets consists of three types of the Demographic surveillance datasets:

1.SAPRIN Individual Surveillance Episodes 2023: Basic Dataset. This dataset contains only the internal and external residency episodes for an individual.

2.SAPRIN Individual Surveillance Episodes 2023: Age-Year 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.

3.SAPRIN Individual Surveillance Episodes 2023: 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.

KIND OF DATA
Event history data

UNITS OF ANALYSIS
Exposure Episodes

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	Vocabulary	URI
Fertility, Mortality, Migration		

KEYWORDS
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 1st July 2022, there were 106 394 people under surveillance of whom 25% were not resident within the surveillance area, with a total of about 2.6 m person years of observation. 31% of the population is under 15 years old. The population is almost exclusively Xitsonga 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 1st July 2022, there were 99 087 people under surveillance, of whom 23% were not resident within the surveillance area, with about 865,000 person years of observation. 29% of the population is under 15 years old. The population is predominantly Sepedi 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 form portions where the Kwamsane township stradles 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 1st July 2022, there were 154 815 people under surveillance of whom 32% were not resident within the surveillance area, with about 2.5 m person years of observation. 30 % of the population is under 15 years old. The population is almost exclusively isiZulu 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 INVESTIGATOR(S)

Name	Affiliation
Dr Kobus Herbst	SAPRIN
Prof Mark Collinson	SAPRIN
Tinofa Mutevedzi	SAPRIN
Prof Steve Tollman	Agincourt
Prof Eric Maimela	DIMAMO
Prof Collins Iwuji	AHRI

OTHER PRODUCER(S)

Name	Affiliation	Role
Dr Molulaqhooa Linda Maoyi	SAPRIN	Dataset Creation
Dr Kobus Herbst	SAPRIN	Dataset Creation

FUNDING

Name	Abbreviation	Role
Department of Science and Innovation	DSI	Current Funder

OTHER ACKNOWLEDGEMENTS

Name	Affiliation	Role
Agincourt Data Team	Agincourt	Providing Data
DIMAMO Data Team	DIMAMO	Providing Data
AHRI Data Team	AHRI	Providing Data
Centre for High Performance Computing	Centre for High Performance Computing	Providing IT Infrastucture for Data Processing

Metadata Production

METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
Molulaqhooa Linda Maoyi	MLM	SAPRIN	Documentation of Study and Review of the metadata
Kobus Herbst	KH	SAPRIN	Documentation of Study and Review of the metadata
Daniel Ohene-Kwofie	DOH	Agincourt	Technical Assistance

Joseph Tlouyamma	JT	DIMAMO	Technical Assistance
Dickman Gareta	DG	AHRI	Technical Assistance

DATE OF METADATA PRODUCTION
2024-01-22

DDI DOCUMENT VERSION
Version 2 (January 2024)

DDI DOCUMENT ID
DDI.SAPRIN.SISED2023V2

Sampling

Sampling Procedure

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

Questionnaires

Overview

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 Collection

Data Collection Dates

Start	End	Cycle
1993-01-01	2022-12-31	Agincourt
1996-01-01	2022-12-31	DIMAMO
2000-01-01	2022-12-31	AHRI

Time Periods

Start	End	Cycle
1993-01-01		Agincourt
1996-01-01		DIMAMO
2000-01-01		AHRI

Data Collection Mode

Household and individual component The desired frequency of data collection for the household component is a function of the resolution required to accurately detect key vital events, specifically births & still births, deaths (especially neonatal deaths) [20] and migration. Reporting of still births and neonatal deaths can be improved by recording pregnancies and then subsequently enquiring about the pregnancy outcome. This implies that the household informant needs to be aware of the pregnancy status of each female household member. This is more likely to be the case during the second two trimesters of the pregnancy. Thus, an interview frequency that guarantees at least one interview during the last two trimesters of a pregnancy will have an increased chance of recording the pregnancy - this implies an interview frequency of at least three times per year at the household level. The accurate recording of residency episodes (the period between birth or in-migration and outmigration or death) requires at least one interview during the period of residency to record the presence of a resident household member. The frequency of household interviews will therefore determine the shortest residency episode that can be reliably detected, in other words three household interviews per year will reliably detect residency episodes with duration of more than four months. Frequent household visits, however, have a higher burden of participation for household informants and cost more. Given the penetration of mobile phone ownership (>96%) in the study populations, it is possible to reduce the frequency of household visits but retain the benefit of accurately recording pregnancy outcomes and migration at a lower cost both to households and the Institution by augmenting household visits with telephone interviews. Therefore, SAPRIN-nodes will conduct a single household visit each year, and a short, scheduled telephone interview with the household informants twice a year. Such interviews will quickly review the household membership roster with the informant noting the current residential status of each member, and pregnancy status of female household members between the ages of 12 and 50. Where a pregnancy, birth, death or migration event is detected, the relevant questionnaire will then be telephonically administered. Telephone interviews and contacts will be conducted by trained interviewers based at a call centre at the nodal office. Consent to conduct a telephonic interview will be obtained from household respondents in the household visit prior to the start of telephonic interview round. The household and individual components will be combined into a single dwelling visit during which the household informant will be invited to be interviewed (household component) as well as all eligible resident household members aged > 15 years (individual component). Interviews will be done by a team of fieldworkers that will visit dwellings according to a predefined visit schedule. Due to youth being in school at the time of some interviews, visits can be planned for afternoons and weekends. Dwelling visits in an area will be preceded by community entry activities conducted by a community engagement team. Verbal autopsy Verbal autopsy interviews will be conducted independently from the household visits, a minimum of one month after the date of death of the deceased to allow for an appropriate mourning period. Interviews will be conducted by specially trained fieldworkers either telephonically or in person depending on the availability of the care giver to be interviewed. Routine service delivery data Access to the service utilisation data will be obtained in terms of a memorandum of agreement between the Node and the service delivery authority, facilitated by the SAPRIN Management Hub, and complaint to the Protection of Personal Information Act (POPIA or the Act), No. 4 of 2013, and any associated regulations, including the Code of Conduct of Research. Currently the AHRI Node has such agreements with the KwaZulu-Natal Provincial Department of Health and the National Health Laboratory Service and similar agreements will be entered into with other service delivery departments, such as Social Welfare, Home Affairs and Education. Data will be transferred to the institute directly from the service delivery authority

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

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 SAPRINQA data quality assessment and indicator process. The data quality checking process is conducted using the SAPRIN QA Julia Code. The Julia Code 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 the SAPRIN Management hub is satisfied that they have high quality data.To produce this final standard dataset, the data is processed using Julia Code on the Centre for High Performance Computing cluster .

Data Appraisal

Estimates of Sampling Error

Not Applicable

File Description

Variable List

SISEAYD2023V1

Content	
Cases	12919716
Variable(s)	38
Structure	Type: Keys: ()
Version	
Producer	
Missing Data	

Variables

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V112	NodeId	SAPRIN Node Identifier	discrete	numeric	
V113	IndividualId	Unique individual identifier	contin	numeric	
V114	Sex	Sex of individual	discrete	numeric	
V115	DoB	Date of birth	discrete	character	
V116	DoD	Date of death	discrete	character	
V117	MotherId	Mother's IndividualId	contin	numeric	
V118	FatherId	Father's IndividualId	contin	numeric	
V119	CalendarYear	Calendar year during episode	contin	numeric	
V120	Age	Age in completed years during episode	contin	numeric	
V121	StartDate	Start date of episode (inclusive)	discrete	character	
V122	EndDate	End date of episode (inclusive)	discrete	character	
V123	Episodes	Total number of episodes for individual	contin	numeric	
V124	Episode	This episode number (first=1, last=Episodes)	contin	numeric	
V125	LocationId	Where individual was resident, household residence if non-resident	contin	numeric	
V126	HouseholdId	Unique household identifier of the household the individual is a member of	contin	numeric	
V127	Resident	Whether individual is resident for duration of episode	discrete	numeric	
V128	Enumeration	Episode starts with an enumeration	discrete	numeric	
V129	Born	Episode starts with the birth of the individual	discrete	numeric	
V130	InMigration	Episode starts with an in-migration	discrete	numeric	
V131	LocationEntry	Episode starts with an internal migration	discrete	numeric	
V132	ExtResStart	Flag to indicate start of external residence	discrete	numeric	
V133	Participation	Resume participation after refusal	discrete	numeric	
V134	YrStart	Flag episode start due to calendar year change	discrete	numeric	
V135	AgeStart	Flag episode start due to age change	discrete	numeric	
V136	Died	Episode ends with the death of the individual	discrete	numeric	
V137	OutMigration	Episode ends with the out-migration	discrete	numeric	
V138	LocationExit	Episode ends with an internal migration	discrete	numeric	

V139	ExtResEnd	Flag to indicate end of external residence	discrete	numeric
V140	LostToFollowUp	Individual was lost to follow-up at the end of the episode	discrete	numeric
V141	Refusal	Individual refused follow-up	discrete	numeric
V142	YrEnd	Flag episode end due to calendar year change	discrete	numeric
V143	AgeEnd	Flag episode end due to age change	discrete	numeric
V144	MembershipStart	Flag to indicate start of household membership	discrete	numeric
V145	MembershipEnd	Flag to indicate end of household membership	discrete	numeric
V146	Memberships	Number of concurrent household memberships	discrete	numeric
V147	Gap	Gap in individual exposure record	discrete	numeric
V148	Days	Duration in days of episode	contin	numeric
V149	Current	Individual still under surveillance	discrete	numeric

SISEAYDD2023V1

Content	
Cases	13028770
Variable(s)	41
Structure	Type: Keys: ()
Version	
Producer	
Missing Data	

Variables

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V150	NodeId	SAPRIN Node Identifier	discrete	numeric	
V151	IndividualId	Unique individual identifier	contin	numeric	
V152	Sex	Sex of individual	discrete	numeric	
V153	DoB	Date of birth	discrete	character	
V154	DoD	Date of death	discrete	character	
V155	MotherId	Mother's IndividualId	contin	numeric	
V156	FatherId	Father's IndividualId	contin	numeric	
V157	CalendarYear	The calendar year in which the episode false	contin	numeric	
V158	Age	The age in completed years of the individual	contin	numeric	
V159	StartDate	Start date of episode (inclusive)	discrete	character	
V160	EndDate	End date of episode (inclusive)	discrete	character	
V161	Episodes	Total number of episodes for individual	contin	numeric	
V162	Episode	This episode number (first=1, last=Episodes)	contin	numeric	
V163	LocationId	Where individual was resident, household residence if non-resident	contin	numeric	
V164	HouseholdId	Unique household identifier of the household the individual is a member of	contin	numeric	
V165	Resident	Whether individual is resident for duration of episode	discrete	numeric	
V166	Enumeration	Episode starts with an enumeration	discrete	numeric	
V167	Born	Episode starts with the birth of the individual	discrete	numeric	
V168	InMigration	Episode starts with an in-migration	discrete	numeric	
V169	LocationEntry	Episode starts with an internal migration	discrete	numeric	
V170	ExtResStart	Flag to indicate start of external residence	discrete	numeric	
V171	Participation	Resume participation after refusal	discrete	numeric	
V172	YrStart	Flag episode start due to calendar year change	discrete	numeric	
V173	AgeStart	Flag episode start due to age change	discrete	numeric	
V174	Died	Episode ends with the death of the individual	discrete	numeric	
V175	OutMigration	Episode ends with the out-migration	discrete	numeric	
V176	LocationExit	Episode ends with an internal migration	discrete	numeric	

V177	ExtResEnd	Flag to indicate end of external residence	discrete	numeric
V178	LostToFollowUp	Individual was lost to follow-up at the end of the episode	discrete	numeric
V179	Refusal	Individual refused follow-up	discrete	numeric
V180	YrEnd	Flag episode end due to calendar year change	discrete	numeric
V181	AgeEnd	Flag episode end due to age change	discrete	numeric
V182	MembershipStart	Flag to indicate start of household membership	discrete	numeric
V183	MembershipEnd	Flag to indicate end of household membership	discrete	numeric
V184	Memberships	Number of concurrent household memberships	discrete	numeric
V185	Gap	Gap in individual exposure record	discrete	numeric
V186	ChildrenEverBorn	Number of children ever born to individual	discrete	numeric
V187	ChildrenBorn	Number of babies delivered at start of episode	discrete	numeric
V188	Delivery	Delivery flag	discrete	numeric
V189	Days	Duration in days of episode	contin	numeric
V190	Current	Individual still under surveillance	discrete	numeric

SISEDB2023V1

Content	
Cases	1262967
Variable(s)	32
Structure	Type: Keys: ()
Version	
Producer	
Missing Data	

Variables

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V191	NodeId	SAPRIN Node Identifier	discrete	numeric	
V192	IndividualId	Unique individual identifier	contin	numeric	
V193	Sex	Sex of individual	discrete	numeric	
V194	DoB	Date of birth	discrete	character	
V195	DoD	Date of death	discrete	character	
V196	MotherId	Mother's IndividualId	contin	numeric	
V197	FatherId	Father's IndividualId	contin	numeric	
V198	StartDate	Start date of episode (inclusive)	discrete	character	
V199	EndDate	End date of episode (inclusive)	discrete	character	
V200	Episodes	Total number of episodes for individual	contin	numeric	
V201	Episode	This episode number (first=1, last=Episodes)	contin	numeric	
V202	LocationId	Where individual was resident, household residence if non-resident	contin	numeric	
V203	HouseholdId	Unique household identifier of the household the individual is a member of	contin	numeric	
V204	Resident	Whether individual is resident for duration of episode	discrete	numeric	
V205	Enumeration	Episode starts with an enumeration	discrete	numeric	
V206	Born	Episode starts with the birth of the individual	discrete	numeric	
V207	InMigration	Episode starts with an in-migration	discrete	numeric	
V208	LocationEntry	Episode starts with an internal migration	discrete	numeric	
V209	ExtResStart	Flag to indicate start of external residence	discrete	numeric	
V210	Participation	Resume participation after refusal	discrete	numeric	
V211	Died	Episode ends with the death of the individual	discrete	numeric	
V212	OutMigration	Episode ends with the out-migration	discrete	numeric	
V213	LocationExit	Episode ends with an internal migration	discrete	numeric	
V214	ExtResEnd	Flag to indicate end of external residence	discrete	numeric	
V215	LostToFollowUp	Individual was lost to follow-up at the end of the episode	discrete	numeric	
V216	Refusal	Individual refused follow-up	discrete	numeric	
V217	MembershipStart	Flag to indicate start of household membership	discrete	numeric	

V218	MembershipEnd	Flag to indicate end of household membership	discrete	numeric
V219	Memberships	Number of concurrent household memberships	discrete	numeric
V220	Gap	Gap in individual exposure record	discrete	numeric
V221	Days	Duration in days of episode	contin	numeric
V222	Current	Individual still under surveillance	discrete	numeric

SAPRIN Node Identifier (NodeId)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 9
Decimals: 0
Range: 1-3

Valid cases: 12919716
Invalid: 0

Unique individual identifier (IndividualId)
File: SISEAYD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-300795

Valid cases: 12919716
Invalid: 0
Minimum: 1
Maximum: 300795
Mean: 125419.5
Standard deviation: 81468.9

Sex of individual (Sex)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 12
Decimals: 0
Range: 0-3

Valid cases: 12919716
Invalid: 0

Date of birth (DoB)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 12919716
Minimum: NaN
Maximum: NaN

Date of death (DoD)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 988762
Minimum: NaN
Maximum: NaN

Mother's IndividualId (MotherId)
File: SISEAYD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 9-300793

Valid cases: 8299800
Invalid: 4619916
Minimum: 9
Maximum: 300793
Mean: 152017
Standard deviation: 72016.4

Father's IndividualId (FatherId)
File: SISEAYD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 49-300792

Valid cases: 4256134
Invalid: 8663582
Minimum: 49
Maximum: 300792
Mean: 162307.8
Standard deviation: 67371.2

Calendar year during episode (CalendarYear)
File: SISEAYD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1992-2022

Valid cases: 12919716
Invalid: 0
Minimum: 1992
Maximum: 2022
Mean: 2011.3
Standard deviation: 8

Age in completed years during episode (Age)
File: SISEAYD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 0-122

Valid cases: 12919716
Invalid: 0
Minimum: 0
Maximum: 122
Mean: 25.2
Standard deviation: 18.9

Start date of episode (inclusive) (StartDate)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 12919716
Minimum: NaN
Maximum: NaN

End date of episode (inclusive) (EndDate)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 12919716
Minimum: NaN
Maximum: NaN

Total number of episodes for individual (Episodes)
File: SISEAYD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-74

Valid cases: 12919716
Invalid: 0
Minimum: 1
Maximum: 74
Mean: 31.6
Standard deviation: 17.4

This episode number (first=1, last=Episodes) (Episode)
File: SISEAYD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-74

Valid cases: 12919716
Invalid: 0
Minimum: 1
Maximum: 74
Mean: 16.3
Standard deviation: 13.6

Where individual was resident, household residence if non-resident (LocationId)
File: SISEAYD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-34959

Valid cases: 12919716
Invalid: 0
Minimum: 1
Maximum: 34959
Mean: 15552.9
Standard deviation: 9241.7

Unique household identifier of the household the individual is a member of (HouseholdId)
File: SISEAYD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-39437

Valid cases: 12919716
Invalid: 0
Minimum: 1
Maximum: 39437
Mean: 17404.2
Standard deviation: 10452.2

Whether individual is resident for duration of episode (Resident)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Episode starts with an enumeration (Enumeration)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Episode starts with the birth of the individual (Born)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Episode starts with an in-migration (InMigration)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Episode starts with an internal migration (LocationEntry)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Flag to indicate start of external residence (ExtResStart)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Resume participation after refusal (Participation)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Flag episode start due to calendar year change (YrStart)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Flag episode start due to age change (AgeStart)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Episode ends with the death of the individual (Died)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Episode ends with the out-migration (OutMigration)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Episode ends with an internal migration (LocationExit)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Flag to indicate end of external residence (ExtResEnd)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Individual was lost to follow-up at the end of the episode (LostToFollowUp)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Individual refused follow-up (Refusal)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Flag episode end due to calendar year change (YrEnd)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Flag episode end due to age change (AgeEnd)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Flag to indicate start of household membership (MembershipStart)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Flag to indicate end of household membership (MembershipEnd)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Number of concurrent household memberships (Memberships)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 12
Decimals: 0
Range: 1-10

Valid cases: 12919716
Invalid: 0

Gap in individual exposure record (Gap)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

Duration in days of episode (Days)
File: SISEAYD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-366

Valid cases: 12919716
Invalid: 0
Minimum: 1
Maximum: 366
Mean: 169.9
Standard deviation: 102.9

Individual still under surveillance (Current)
File: SISEAYD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 12919716
Invalid: 0

SAPRIN Node Identifier (NodeId)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 9
Decimals: 0
Range: 1-3

Valid cases: 13028770
Invalid: 0

Unique individual identifier (IndividualId)
File: SISEAYDD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-300795

Valid cases: 13028770
Invalid: 0
Minimum: 1
Maximum: 300795
Mean: 125543.1
Standard deviation: 81467

Sex of individual (Sex)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 12
Decimals: 0
Range: 0-3

Valid cases: 13028770
Invalid: 0

Date of birth (DoB)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 13028770
Minimum: NaN
Maximum: NaN

Date of death (DoD)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 994621
Minimum: NaN
Maximum: NaN

Mother's IndividualId (MotherId)
File: SISEAYDD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 9-300793

Valid cases: 8360239
Invalid: 4668531
Minimum: 9
Maximum: 300793
Mean: 152058.6
Standard deviation: 71985.5

Father's IndividualId (FatherId)
File: SISEAYDD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 49-300792

Valid cases: 4289585
Invalid: 8739185
Minimum: 49
Maximum: 300792
Mean: 162325.2
Standard deviation: 67346.7

The calendar year in which the episode false (CalendarYear)
File: SISEAYDD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1992-2022

Valid cases: 13028770
Invalid: 0
Minimum: 1992
Maximum: 2022
Mean: 2011.3
Standard deviation: 8

The age in completed years of the individual (Age)
File: SISEAYDD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 0-122

Valid cases: 13028770
Invalid: 0
Minimum: 0
Maximum: 122
Mean: 25.2
Standard deviation: 18.8

Start date of episode (inclusive) (StartDate)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 13028770
Minimum: NaN
Maximum: NaN

End date of episode (inclusive) (EndDate)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 13028770
Minimum: NaN
Maximum: NaN

Total number of episodes for individual (Episodes)
File: SISEAYDD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-77

Valid cases: 13028770
Invalid: 0
Minimum: 1
Maximum: 77
Mean: 31.9
Standard deviation: 17.6

This episode number (first=1, last=Episodes) (Episode)
File: SISEAYDD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-77

Valid cases: 13028770
Invalid: 0
Minimum: 1
Maximum: 77
Mean: 16.5
Standard deviation: 13.7

Where individual was resident, household residence if non-resident (LocationId)
File: SISEAYDD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-34959

Valid cases: 13028770
Invalid: 0
Minimum: 1
Maximum: 34959
Mean: 15553.4
Standard deviation: 9241.5

Unique household identifier of the household the individual is a member of (HouseholdId)
File: SISEAYDD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-39437

Valid cases: 13028770
Invalid: 0
Minimum: 1
Maximum: 39437
Mean: 17404.8
Standard deviation: 10452.2

Whether individual is resident for duration of episode (Resident)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Episode starts with an enumeration (Enumeration)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Episode starts with the birth of the individual (Born)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Episode starts with an in-migration (InMigration)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Episode starts with an internal migration (LocationEntry)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Flag to indicate start of external residence (ExtResStart)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Resume participation after refusal (Participation)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Flag episode start due to calendar year change (YrStart)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Flag episode start due to age change (AgeStart)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Episode ends with the death of the individual (Died)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Episode ends with the out-migration (OutMigration)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Episode ends with an internal migration (LocationExit)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Flag to indicate end of external residence (ExtResEnd)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Individual was lost to follow-up at the end of the episode (LostToFollowUp)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Individual refused follow-up (Refusal)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Flag episode end due to calendar year change (YrEnd)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Flag episode end due to age change (AgeEnd)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Flag to indicate start of household membership (MembershipStart)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Flag to indicate end of household membership (MembershipEnd)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Number of concurrent household memberships (Memberships)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 12
Decimals: 0
Range: 1-10

Valid cases: 13028770
Invalid: 0

Gap in individual exposure record (Gap)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Number of children ever born to individual (ChildrenEverBorn)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 12
Decimals: 0
Range: 0-16

Valid cases: 13028770
Invalid: 0

Number of babies delivered at start of episode (ChildrenBorn)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 12
Decimals: 0
Range: 0-9

Valid cases: 13028770
Invalid: 0

Delivery flag (Delivery)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

Duration in days of episode (Days)
File: SISEAYDD2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-366

Valid cases: 13028770
Invalid: 0
Minimum: 1
Maximum: 366
Mean: 168.5
Standard deviation: 102.8

Individual still under surveillance (Current)
File: SISEAYDD2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 13028770
Invalid: 0

SAPRIN Node Identifier (NodeId)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 9
Decimals: 0
Range: 1-3

Valid cases: 1262967
Invalid: 0

Unique individual identifier (IndividualId)
File: SISEDB2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-300795

Valid cases: 1262967
Invalid: 0
Minimum: 1
Maximum: 300795
Mean: 124231.3
Standard deviation: 80513.5

Sex of individual (Sex)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 12
Decimals: 0
Range: 0-3

Valid cases: 1262967
Invalid: 0

Date of birth (DoB)
File: SISEDB2023V1

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 1262967
Minimum: NaN
Maximum: NaN

Date of death (DoD)
File: SISEDB2023V1

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 90844
Minimum: NaN
Maximum: NaN

Mother's IndividualId (MotherId)
File: SISEDB2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 9-300793

Valid cases: 811225
Invalid: 451742
Minimum: 9
Maximum: 300793
Mean: 149577.1
Standard deviation: 72019.3

Father's IndividualId (FatherId)
File: SISEDB2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 49-300792

Valid cases: 388662
Invalid: 874305
Minimum: 49
Maximum: 300792
Mean: 159634.6
Standard deviation: 67612.5

Start date of episode (inclusive) (StartDate)
File: SISEDB2023V1

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 1262967
Minimum: NaN
Maximum: NaN

End date of episode (inclusive) (EndDate)
File: SISEDB2023V1

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 1262967
Minimum: NaN
Maximum: NaN

Total number of episodes for individual (Episodes)
File: SISEDB2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-22

Valid cases: 1262967
Invalid: 0
Minimum: 1
Maximum: 22
Mean: 2.9
Standard deviation: 2.2

This episode number (first=1, last=Episodes) (Episode)
File: SISEDB2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-22

Valid cases: 1262967
Invalid: 0
Minimum: 1
Maximum: 22
Mean: 2
Standard deviation: 1.5

Where individual was resident, household residence if non-resident (LocationId)
File: SISEDB2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-34959

Valid cases: 1262967
Invalid: 0
Minimum: 1
Maximum: 34959
Mean: 15577.2
Standard deviation: 9275.2

Unique household identifier of the household the individual is a member of (HouseholdId)
File: SISEDB2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-39437

Valid cases: 1262967
Invalid: 0
Minimum: 1
Maximum: 39437
Mean: 17402.5
Standard deviation: 10501

Whether individual is resident for duration of episode (Resident)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Episode starts with an enumeration (Enumeration)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Episode starts with the birth of the individual (Born)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Episode starts with an in-migration (InMigration)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Episode starts with an internal migration (LocationEntry)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Flag to indicate start of external residence (ExtResStart)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Resume participation after refusal (Participation)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Episode ends with the death of the individual (Died)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Episode ends with the out-migration (OutMigration)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Episode ends with an internal migration (LocationExit)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Flag to indicate end of external residence (ExtResEnd)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Individual was lost to follow-up at the end of the episode (LostToFollowUp)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Individual refused follow-up (Refusal)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Flag to indicate start of household membership (MembershipStart)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Flag to indicate end of household membership (MembershipEnd)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Number of concurrent household memberships (Memberships)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 12
Decimals: 0
Range: 1-10

Valid cases: 1262967
Invalid: 0

Gap in individual exposure record (Gap)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0

Duration in days of episode (Days)
File: SISEDB2023V1

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 1-11263

Valid cases: 1262967
Invalid: 0
Minimum: 1
Maximum: 11263
Mean: 1738.1
Standard deviation: 1857.2

Individual still under surveillance (Current)
File: SISEDB2023V1

Overview

Type: Discrete
Format: numeric
Width: 15
Decimals: 0
Range: -1-1

Valid cases: 1262967
Invalid: 0