

South Africa - AHRI.Education Statuses Dataset

Nompumelelo Mkwazi

Report generated on: February 18, 2026

Visit our data catalog at: <https://data.ahri.org/index.php>

Identification

SURVEY ID NUMBER

AHRI.EducationStatuses

TITLE

AHRI.Education Statuses Dataset

COUNTRY

Name	Country code
South Africa	ZA

ABSTRACT

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

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

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

The following are the versions of the Demographic surveillance datasets:

CoresidencyEpisodes
 EducationCleaned
 EducationStatuses
 HIVStatusEpisodes
 HouseholdAssets
 HouseholdMap
 HouseholdObservations
 HouseholdSocioEconomicStatus
 HouseholdStatusEpisodes
 IndividualHSEEpisodes
 IndividualMap
 LabourEpisodes
 LabourStatuses
 LocationMap
 MaritalStatuses
 PartnerEpisodes
 SurveillanceEpisodesBasic
 SurveillanceEpisodesYTAge
 SurveillanceEpisodesYrAge_Coresidency
 SurveillanceEpisodesYrAge_HIV
 SurveillanceEpisodesYrAge_HSE
 SurveillanceEpisodesYrAge_Labour

SurveillanceEpisodesYrAge_Partners
 SurveillanceEpisodesYrAgeDel
 SurveillanceEpisodes VrAgeDel_HSE
 SurveillanceEpisodesYrAgeDeL_HSE_Labour
 SurveillanceEpisodesYrAgeDeL_HSE_Labour_Partner_Parents
 SurveillanceEpisodes VrAgeDel_HSE_Labour_Partner_Parents_HIV
 SurveillanceEpisodesYrAgeDeL_HSE_Labour_Partners

KIND OF DATA

Event History Data

UNIT OF ANALYSIS

Episodes of exposure

Version

VERSION DESCRIPTION

v2.0.0

Scope

TOPICS

Topic	Vocabulary	URI
Education Statuses	Africa Health Research Institute	www.ahri.org

KEYWORDS

Keyword	Vocabulary	URI
Fertility, Mortality, Migration, Social-economic status, HIV, ART	Africa Health Research Institute	www.ahri.org

Coverage

GEOGRAPHIC COVERAGE

AHRI is situated in the south-east portion of the uMkhanyakude district of KwaZulu-Natal province near the town of Mtubatuba. It is bounded on the west by the Umfolozi-Hluhluwe nature reserve, on the south by the Umfolozi river, on the east by the N2 highway (except form portions where the KwaMsane township strangles the highway) and in the north by the provincial road R22. The surveillance area is approximately 845km² in size.

UNIVERSE

Households resident in dwellings within the study area are eligible for inclusion in the household component of AHRI Demographic Surveillance. All individuals identified by the household proxy informant as members of the households are enumerated. A resident household member is an individual that intends to sleep majority of time at the dwelling occupied by the household over a four-month period. Households include resident and non-resident members. An individual is a non-resident member if they have close ties to the household, but do not physically reside with the household at the dwelling most of the time. Non-resident members can also be called temporary migrants and they are enumerated within the household list. Because household membership is not tied to physical residency, an individual may be a member of more than one household.

Producers and sponsors

PRIMARY INVESTIGATORS

Name	Affiliation
Dr. Kobus Herbst	AHRI, SAPRIN

Dr. Guy Harling	AHRI
Dr. Mark Siedner	AHRI
Prof Willem Hanekom	AHRI

PRODUCERS

Name	Abbreviation
Africa Health Research Institute	AHRI

FUNDING AGENCY/SPONSOR

Name	Abbreviation	Role
Wellcome Trust	WT	Core funding
SAPRIN	SAPRIN	

OTHER IDENTIFICATIONS/ACKNOWLEDGMENTS

Name	Affiliation	Role
AHRI Research Data Management Team	AHRI	Research Data Management
AHRI Data Collection Team	AHRI	Data Collection
AHRI Community Engagement Team	AHRI	Community Engagement

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
2000-02-22	2024-12-17

Data Processing

DATA EDITING

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

Access policy

ACCESS CONDITIONS

Access to the data requires accurate completion of the online data access application form accessible on the AHRI Data repository(<<https://data.ahri.org/>>). Data users are required to abide by the data use conditions stipulated on the application for access to the data. Failure to do so may result in their data access privileges being revoked by the Data Custodian. In order to recognise the effort and intellectual contributions of AHRI investigators in producing and curating the

data, users of AHRI data must acknowledge the source of the data and abide by the terms and conditions under which the data is accessed and must cite the dataset in publication using the citation provided as part of this documentation. All analytical datasets published on the AHRI Data Repository are assigned digital object identifier (DOIs) and the DOIs can be found on the Data Repository under Study Description tab - Access policy. AHRI data users are required to always cite the dataset using the relevant DOI.

<https://github.com/SAPRIN-MRC/SaprinEpisodes.jl.git>

CITATION REQUIREMENTS

Herbst, K., Harling, G., Siedner, M., & Hanekom, W. (2026). AHRI.Education Statuses Dataset [Data set]. Africa Health Research Institute.

DOI: <https://doi.org/10.23664/AHRI.EDUCATIONSTATUSES>

Metadata production

DDI DOCUMENT ID

DDI.AHRI.EducationStatuses

PRODUCERS

Name	Abbreviation
Africa Health Research Institute	AHRI

Data Dictionary

Data file	Cases	Variables
EducationStatuses	1795797	5

Data file: EducationStatuses

Cases: 1795797

Variables: 5

Variables

ID	Name	Label	Question
V301	IndividualId	Individual ID	
V302	ObservationDate	Observation Date	
V303	CurrentEducation	Current Education	
V304	HighestSchoolLevel	Highest School Level	
V305	HighestNonSchoolLevel	Highest Non-School Level	

Total: 5

INDIVIDUALID: Individual ID**Data file:** EducationStatuses**Overview**

Valid: 1795797 Invalid: 0 Minimum: 1 Maximum: 263153 Mean: 120933.687 Standard deviation: 77410.781

Type: Continuous Decimal: 0 Width: 9 Range: 1 - 263153 Format: Numeric

OBSERVATIONDATE: Observation Date**Data file:** EducationStatuses**Overview**

Valid: 1795797 Minimum: 2000-02-22 Maximum: 2024-12-17

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

CURRENTEDUCATION: Current Education**Data file:** EducationStatuses**Overview**

Valid: 1795797 Invalid: 0

Type: Discrete Decimal: 0 Width: 9 Range: -1 - 6 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Undefined	200030	11.1%
1	Part time	6358	0.4%
2	Full time	640523	35.7%
3	No	936903	52.2%
5	Refused	1648	0.1%
6	Unknown	10335	0.6%

HIGHESTSCHOOLEVEL: Highest School Level**Data file:** EducationStatuses**Overview**

Valid: 1795797 Invalid: 0

Type: Discrete Decimal: 0 Width: 18 Range: -1 - 99 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown	82245	4.6%
0	Grade 0	3209	0.2%
1	Grade 1	61771	3.4%
2	Grade 2	68944	3.8%
3	Grade 3	74405	4.1%
4	Grade 4	83518	4.7%
5	Grade 5	71527	4%
6	Grade 6	76376	4.3%
7	Grade 7	99300	5.5%
8	Grade 8	96067	5.3%
9	Grade 9	106291	5.9%
10	Grade 10	152697	8.5%
11	Grade 11	179950	10%
12	Grade 12	499203	27.8%
13	Special School	18	0%
21	Higher Certificate	0	0%
22	National Diploma	0	0%
23	Degree	0	0%
24	Honours	0	0%
25	Masters	0	0%
26	Doctors	0	0%
31	ABET 1	0	0%
32	ABET 2	0	0%
33	ABET 3	0	0%
34	ABET 4	0	0%
41	FET 1	0	0%
42	FET 2	0	0%
43	FET 3	0	0%
50	LT 1 year	29344	1.6%
51	Creche	261	0%
96	Refused	1982	0.1%
97	NAD	5288	0.3%
98	No education	103401	5.8%
99	Not attending	0	0%

HIGHESTNONSCHOOLLEVEL: Highest Non-School Level**Data file: EducationStatuses****Overview**

Valid: 1795797 Invalid: 0
 Type: Discrete Decimal: 0 Width: 18 Range: -1 - 99 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown	41848	2.3%
21	Higher Certificate	41121	2.3%
22	National Diploma	19696	1.1%
23	Degree	32203	1.8%
24	Honours	1115	0.1%
25	Masters	1771	0.1%
26	Doctors	134	0%
31	ABET 1	1334	0.1%
32	ABET 2	171	0%
33	ABET 3	108	0%
34	ABET 4	324	0%
41	FET 1	3036	0.2%
42	FET 2	3663	0.2%
43	FET 3	7819	0.4%
45		3338	0.2%
96	Refused	1806	0.1%
97	NAD	103462	5.8%
98	No education	1532848	85.4%
99	Not attending	0	0%

Download related resources