

## **South Africa**

**Prof. Frank Tanser(Principal Investigator), Africa  
Health Research Institute/ Stellenbosch University**

**Prof Adrian Dobra, University of Washington**

**Prof Till Barnighausen, Africa Health Research Institute/  
Medical Faculty and University Hospital, Heidelberg University,  
Heidelberg Institute of Global Health, Heidelberg, Germany**

**AHRI:Sesikhona! (we are here) Activity Space**

## **Study Documentation**

December 4, 2025

# Metadata Production

<b>Metadata Producer(s)</b>	Africa Health Research Institute (AHRI)
<b>Identification</b>	DDI.AHRI.Sesikhona.ActivitySpace.Survey

# Table of Contents

<a href="#">Overview</a>	4
<a href="#">Scope &amp; Coverage</a>	4
<a href="#">Producers &amp; Sponsors</a>	4
<a href="#">Sampling</a>	4
<a href="#">Data Collection</a>	5
<a href="#">Data Processing &amp; Appraisal</a>	5
<a href="#">Accessibility</a>	5
<a href="#">Files Description</a>	6
<a href="#">AHRI.Sesikhona.ActivitySpace.Survey.2025.v1.0</a>	6
<a href="#">Variables List</a>	7
<a href="#">AHRI.Sesikhona.ActivitySpace.Survey.2025.v1.0</a>	7
<a href="#">Variables Description</a>	8
<a href="#">AHRI.Sesikhona.ActivitySpace.Survey.2025.v1.0</a>	9

## AHRI:Sesikhona! (we are here) Activity Space

### Overview

<b>Identification</b>	AHRI.Sesikhona.ActivitySpace.Survey
<b>Version</b>	v1.0.0
<b><u>Abstract</u></b>	<p>The Sesikhona study is the first of its kind study using smartphone-based GPS technology to record ultra-high resolution mobility patterns of over 200 young adults aged 20-30 years in rural KwaZulu Natal. The study seeks to identify distinct movement typologies within this highly mobile, HIV hyper-endemic population and to quantify travel in and out of known high-risk locations, with the long-term goal of developing a location-intelligent smartphone intervention that delivers real-time, precision messaging to individuals at heightened risk of HIV acquisition, transmission or treatment interruption. Conducted in three phases between 2021 and 2025, the study used the Avicenna Research Platform to collect GPS and survey data. Between 2023 and 2025, consenting individuals self-completed surveys bi-daily and monthly based on their travel, mobility and activity space questions to validate the collected GPS location data, understand reasons for travel and planning thereof. 25978, 25237, 25236, 25235, 22768, 22766</p>
<b>Kind of Data</b>	Survey data
<b>Unit of Analysis</b>	Young adults aged 20-30 years old

### Scope & Coverage

<b>Keywords</b>	Human mobility, HIV, GPS technology, digital technology, travel, mobility and activity space
<b>Topics</b>	GPS technology, smartphone-based GPS, human mobility, young adults, HIV, mobile health, digital technology, mental health
<b><u>Time Period(s)</u></b>	2024-2025
<b><u>Countries</u></b>	South Africa
<b><u>Geographic Coverage</u></b>	Residents from uMkhanyakude in KwaZulu Natal but move anywhere in South Africa
<b><u>Universe</u></b>	A random sample of adults aged 20-30 years old who were residents in Southern PIPSA, participated in the 2019 HIV surveillance and consented to participate in Sesikhona study and opted in for the sexual behavior surveys

### Producers & Sponsors

<b>Primary Investigator(s)</b>	<p>Prof. Frank Tanser(Principal Investigator), Africa Health Research Institute/ Stellenbosch University          Prof Adrian Dobra, University of Washington          Prof Till Barnighausen, Africa Health Research Institute/Medical Faculty and University Hospital, Heidelberg University, Heidelberg Institute of Global Health, Heidelberg, Germany</p>
<b>Other Producer(s)</b>	Africa Health Research Institute (AHRI)
<b>Funding Agency/ies</b>	<p>National Institute of Health (NIH) , Funder          German Science Foundation (DFG) , Funder          Academy of Medical Sciences Newton Fund (NAF) , Funder</p>

### Sampling

#### Sampling Procedure

Participants were randomly selected from the 2019 AHRI HIV surveillance program if they were aged 20-30 years old and had participated in the HIV survey in 2019 and were resident members in the southern PIP and owned a smartphone that supported the study app.

## Data Collection

<b>Data Collection Dates</b>	start 2024-10-01 end 2025-06-30
------------------------------	------------------------------------

## Data Processing & Appraisal

### Data Editing

Travel, Mobility and Activity space survey responses from Sesikhona participants via the Avicenna app

## Accessibility

### Access Conditions

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

### Citation Requirements

Tanse, F., Dobra, A., & Barnighausen, T. (2025). AHRI:Sesikhona! (we are here) Activity Space [Data set]. Africa Health Research Institute.

DOI: <https://doi.org/10.23664/AHRI.SESIKHONA.ACTIVITYSPACE.SURVEY>

# Files Description

Dataset contains 1 file(s)

<b>AHRI.Sesikhona.ActivitySpace.Survey.2025.v1.0</b>	
<b># Cases</b>	2241
<b># Variable(s)</b>	16

# Variables List

Dataset contains 16 variable(s)

File AHRI.Sesikhona.ActivitySpace.Survey.2025.v1.0							
#	Name	Label	Type	Format	Valid	Invalid	Question
1	<a href="#">pid</a>	Participant ID	continuous	numeric-12.0	2241	0	-
2	<a href="#">maq_1</a>	Work	discrete	numeric-12.0	353	1888	-
3	<a href="#">maq_2</a>	School	discrete	numeric-12.0	160	2081	-
4	<a href="#">maq_3</a>	Market/Shops	discrete	numeric-12.0	436	1805	-
5	<a href="#">maq_4</a>	Town	discrete	numeric-12.0	510	1731	-
6	<a href="#">maq_5</a>	Church	discrete	numeric-12.0	276	1965	-
7	<a href="#">maq_6</a>	Bar/Drinking Venue	discrete	numeric-12.0	60	2181	-
8	<a href="#">maq_7</a>	Social Event(Funeral Wedding Sporting)	discrete	numeric-12.0	142	2099	-
9	<a href="#">maq_8</a>	Clinic/Hospital/Doctor	discrete	numeric-12.0	252	1989	-
10	<a href="#">maq_9</a>	Visit your Partner	discrete	numeric-12.0	275	1966	-
11	<a href="#">maq_10</a>	Visit another family member	discrete	numeric-12.0	240	2001	-
12	<a href="#">maq_11</a>	Visit a Friend	discrete	numeric-12.0	356	1885	-
13	<a href="#">maq_12</a>	Travel Out of Town	discrete	numeric-12.0	125	2116	-
14	<a href="#">maq_13</a>	Other	discrete	numeric-12.0	383	1858	-
15	<a href="#">record_t..</a>	Record Time	discrete	character-32	2241	0	-
16	<a href="#">sur_status</a>	Status	discrete	numeric-12.0	2241	0	-

# Variables Description

Dataset contains 16 variable(s)

# File : AHRI.Sesikhona.ActivitySpace.Survey.2025.v1.0

## # pid: Participant ID

**Information** [Type= continuous] [Format=numeric] [Range= 39636-95065] [Missing=\*]

**Statistics [NW/ W]** [Valid=2241 /-] [Invalid=0 /-] [Mean=52173.469 /-] [StdDev=11827.971 /-]

## # maq\_1: Work

**Information** [Type= discrete] [Format=numeric] [Range= 1-2] [Missing=\*]

**Statistics [NW/ W]** [Valid=353 /-] [Invalid=1888 /-]

Value	Label	Cases	Percentage
1	Yes	353	100.0%
2	No	0	
Sysmiss		1888	

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

## # maq\_2: School

**Information** [Type= discrete] [Format=numeric] [Range= 1-2] [Missing=\*]

**Statistics [NW/ W]** [Valid=160 /-] [Invalid=2081 /-]

Value	Label	Cases	Percentage
1	Yes	160	100.0%
2	No	0	
Sysmiss		2081	

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

## # maq\_3: Market/Shops

**Information** [Type= discrete] [Format=numeric] [Range= 1-2] [Missing=\*]

**Statistics [NW/ W]** [Valid=436 /-] [Invalid=1805 /-]

Value	Label	Cases	Percentage
1	Yes	436	100.0%
2	No	0	
Sysmiss		1805	

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

## # maq\_4: Town

**Information** [Type= discrete] [Format=numeric] [Range= 1-2] [Missing=\*]

**Statistics [NW/ W]** [Valid=510 /-] [Invalid=1731 /-]

Value	Label	Cases	Percentage
1	Yes	510	100.0%
2	No	0	
Sysmiss		1731	

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

## # maq\_5: Church

**Information** [Type= discrete] [Format=numeric] [Range= 1-2] [Missing=\*]

**Statistics [NW/ W]** [Valid=276 /-] [Invalid=1965 /-]

Value	Label	Cases	Percentage
1	Yes	276	100.0%
2	No	0	

## File : AHRI.Sesikhona.ActivitySpace.Survey.2025.v1.0

### # maq\_5: Church

Value	Label	Cases	Percentage
Sysmiss		1965	

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

### # maq\_6: Bar/Drinking Venue

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
--------------------	--

<b>Statistics [NW/ W]</b>	[Valid=60 /-] [Invalid=2181 /-]
---------------------------	---------------------------------

Value	Label	Cases	Percentage
1	Yes	60	100.0%
2	No	0	
Sysmiss		2181	

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

### # maq\_7: Social Event(Funeral Wedding Sporting)

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
--------------------	--

<b>Statistics [NW/ W]</b>	[Valid=142 /-] [Invalid=2099 /-]
---------------------------	----------------------------------

Value	Label	Cases	Percentage
1	Yes	142	100.0%
2	No	0	
Sysmiss		2099	

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

### # maq\_8: Clinic/Hospital/Doctor

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
--------------------	--

<b>Statistics [NW/ W]</b>	[Valid=252 /-] [Invalid=1989 /-]
---------------------------	----------------------------------

Value	Label	Cases	Percentage
1	Yes	252	100.0%
2	No	0	
Sysmiss		1989	

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

### # maq\_9: Visit your Partner

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
--------------------	--

<b>Statistics [NW/ W]</b>	[Valid=275 /-] [Invalid=1966 /-]
---------------------------	----------------------------------

Value	Label	Cases	Percentage
1	Yes	275	100.0%
2	No	0	
Sysmiss		1966	

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

### # maq\_10: Visit another family member

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
--------------------	--

<b>Statistics [NW/ W]</b>	[Valid=240 /-] [Invalid=2001 /-]
---------------------------	----------------------------------

Value	Label	Cases	Percentage
1	Yes	240	100.0%
2	No	0	

## File : AHRI.Sesikhona.ActivitySpace.Survey.2025.v1.0

### # maq\_10: Visit another family member

Value	Label	Cases	Percentage
Sysmiss		2001	

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

### # maq\_11: Visit a Friend

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
--------------------	--

<b>Statistics [NW/ W]</b>	[Valid=356 /-] [Invalid=1885 /-]
---------------------------	----------------------------------

Value	Label	Cases	Percentage
1	Yes	356	100.0%
2	No	0	
Sysmiss		1885	

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

### # maq\_12: Travel Out of Town

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
--------------------	--

<b>Statistics [NW/ W]</b>	[Valid=125 /-] [Invalid=2116 /-]
---------------------------	----------------------------------

Value	Label	Cases	Percentage
1	Yes	125	100.0%
2	No	0	
Sysmiss		2116	

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

### # maq\_13: Other

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
--------------------	--

<b>Statistics [NW/ W]</b>	[Valid=383 /-] [Invalid=1858 /-]
---------------------------	----------------------------------

Value	Label	Cases	Percentage
1	Yes	383	100.0%
2	No	0	
Sysmiss		1858	

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

### # record\_time: Record Time

<b>Information</b>	[Type= discrete] [Format=character] [Missing=*]
--------------------	---

<b>Statistics [NW/ W]</b>	[Valid=2241 /-] [Invalid=0 /-]
---------------------------	--------------------------------

### # sur\_status: Status

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= 0-6] [Missing=*]
--------------------	--

<b>Statistics [NW/ W]</b>	[Valid=2241 /-] [Invalid=0 /-]
---------------------------	--------------------------------

Value	Label	Cases	Percentage
0	Unknown	0	
1	Completed	1942	86.7%
2	Canceled	5	0.2%
3	Expired	289	12.9%
4	Blocked	5	0.2%
5	Invalid Criteria	0	
6	In progress	0	

**File : AHRI.Sesikhona.ActivitySpace.Survey.2025.v1.0**

**# sur\_status: Status**

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