

South Africa - Data Everywhere Datasets

Report generated on: February 10, 2016

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

Overview

Identification

ID NUMBER

DDI.DataEveryWhere-TasP

Version

VERSION DESCRIPTION

1.1.0

PRODUCTION DATE

2016-02-02

NOTES

1.0.0 Initial Release

-- Releasing Locations dataset

1.1.0 Release with HexagonIds

-- Locations dataset contains HexagonIds

-- Item HexId removed from all datasets

-- Houshold dataset contains LocationId

-- First release of Linkage to Care Dataset

-- First release of DataCollectionCycle dataset

-- First release of Migration dataset

TOPICS

Topic	Vocabulary	URI
Data Everywhere		

Producers and Sponsors

OTHER PRODUCER(S)

Name	Affiliation	Role
Kobus Herbst	Africa Centre	
Jaco Dreyer	Africa Centre	

Metadata Production

METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
Kobus Herbst		Africa Centre	Dataset production/documentation
Jaco Dreyer		Africa Centre	Dataset production/documentation

DATE OF METADATA PRODUCTION

2016-02-02

DDI DOCUMENT VERSION

1.1.0

DDI DOCUMENT ID

DDI.DataEveryWhere-TasP

Sampling

No content available

Questionnaires

No content available

Data Collection

Data Collection Dates

Start	End	Cycle
2012-03-09	2015-12-04	N/A

Time Periods

Start	End	Cycle
2012-03-09	2016-02-02	N/A

Data Processing

No content available

Data Appraisal

No content available

File Description

Variable List

Locations dataset

Content	Locations Dataset
Cases	7568
Variable(s)	8
Structure	Type: Keys: ()
Version	1.1.0
Producer	
Missing Data	

Variables

ID	Name	Label	Type	Format	Question
V1	StudyId	Study Identifier	discrete	character	
V2	LocationId	Location Identifier	contin	numeric	
V3	LocationType	Location Type	discrete	character	
V4	CoordLongX	Coordinates Longitude X-axis	contin	numeric	
V5	CoordLatY	Coordinates Latitude Y-axis	contin	numeric	
V6	HexagonId	Hexagon Identifier	discrete	numeric	
V7	Num_Individuals	Number of Individuals per LocationId	contin	numeric	
V8	Num_Households	Number of Households per LocationId	contin	numeric	

Individuals dataset

Content	Individuals dataset
Cases	28626
Variable(s)	11
Structure	Type: Keys: ()
Version	1.1.0
Producer	
Missing Data	

Variables

ID	Name	Label	Type	Format	Question
V9	StudyId	Study Identifier	discrete	character	
V10	IndividualId	Individual Identifier	contin	numeric	
V11	HouseholdId	Household Identifier	contin	numeric	
V13	Sex	Sex	discrete	character	
V14	DateOfBirth	DateOfBirth	discrete	character	
V15	HighestEducationLevel	Highest Education Level Code	contin	numeric	
V16	HighestEducationLevelDesc	Highest Education Level Description	discrete	character	
V17	CurrentMaritalStatus	Current Marital Status Code	discrete	numeric	
V18	CurrentMaritalStatusDesc	Current Marital Status Description	discrete	character	
V19	CurrentlyInEmployment	Currently In Employment Code	discrete	numeric	
V20	CurrentlyInEmploymentDesc	Currently In Employment Description	discrete	character	

Household dataset

Content	Household datasets
Cases	7803
Variable(s)	6
Structure	Type: Keys: ()
Version	1.1.0
Producer	
Missing Data	

Variables

ID	Name	Label	Type	Format	Question
V21	StudyId	Study Identifier	discrete	character	
V22	HouseholdId	Household Identifier	contin	numeric	
V27	LocationId	Location Identifier	discrete	numeric	
V24	ClusterId	Cluster Identifier	contin	numeric	
V25	AssetScore	Asset Score	discrete	numeric	
V26	AssetScoreIndex	Asset Score Index	discrete	character	

Linkage to Care Dataset.

Content

Cases 26393

Variable(s) 16

Structure Type:
Keys: ()

Version 1.1.0

Producer

Missing Data

Variables

ID	Name	Label	Type	Format	Question
V28	IndividualId	Individual Identifier	contin	numeric	
V53	TestAgreedatFirstPositive	Test Agreed at First Positive Test	discrete	numeric	
V30	FirstPositiveDate	Date of First Positive Test	discrete	character	
V54	TestAgreedatLastContact	Test Agreed at Last Contact	discrete	numeric	
V32	LastHHContactDate	Date of Last Household Contact	discrete	character	
V33	FirstClinicDate	Date of First Clinic Visit	discrete	character	
V34	FirstClinicLocationId	First Clinic LocationId	contin	numeric	
V35	LastClinicDate	Daet of Last Clinic Visit	discrete	character	
V36	LastClinicLocationId	Last Clinic LocationId	contin	numeric	
V37	ClinicVisited	Clinic Visited	discrete	character	
V38	CurrentDate	CurrentDate	discrete	numeric	
V39	DaystoClinicLink	Number of Days to link to Clinic	contin	numeric	
V40	DaysSinceLastContactHome	Number of Days Since Last Home Contact	contin	numeric	
V41	DaysSinceLastCLinicVisit	Number of Days Since Last CLinic Visit	contin	numeric	
V42	DaysNotLinkedtoClinic	Number of Days Not Linked to Clinic	contin	numeric	
V52	HIVStatus	HIV Status	discrete	character	

DataCollectionCycle dataset

Content

Cases 112

Variable(s) 5

Structure Type:
Keys: ()

Version 1.1.0

Producer

Missing Data

Variables

ID	Name	Label	Type	Format	Question
V43	StudyId	Study Identifier	discrete	character	
V44	DataCollectionCycle	DataCollectionCycle	discrete	numeric	
V45	ClusterId	Cluster Identifier	contin	numeric	
V46	StartDate	Start Date	discrete	character	
V47	EndDate	End Date	discrete	character	

Migration dataset

Content

Cases 11642

Variable(s) 4

Structure Type:
Keys: ()

Version 1.1.0

Producer

Missing Data

Variables

ID	Name	Label	Type	Format	Question
V48	IndividualId	Individual Identifier	contin	numeric	
V49	MigrateDate	Date of Migration	discrete	character	
V50	MigrateVPDate	Date of Migration - Text	discrete	character	
V51	DestinationCode	Destination Code	discrete	character	

Study Identifier (StudyId)

File: Locations dataset

Overview

Type: Discrete	Valid cases: 7568
Format: character	Invalid: 0
Width: 4	

Location Identifier (LocationId)

File: Locations dataset

Overview

Type: Continuous	Valid cases: 7568
Format: numeric	Invalid: 0
Decimals: 0	Minimum: 2
Range: 1-12094	Maximum: 12101
	Mean: 6025.9
	Standard deviation: 3496.7

Location Type (LocationType)

File: Locations dataset

Overview

Type: Discrete	Valid cases: 7568
Format: character	Invalid: 0
Width: 3	

Coordinates Longitude X-axis (CoordLongX)

File: Locations dataset

Overview

Type: Continuous	Valid cases: 7568
Format: numeric	Invalid: 0
Decimals: 8	Minimum: 0
Range: 0-32.37347	Maximum: 32.4
	Mean: 32.3
	Standard deviation: 0.4

Coordinates Latitude Y-axis (CoordLatY)

File: Locations dataset

Overview

Type: Continuous	Valid cases: 7568
Format: numeric	Invalid: 0
Decimals: 8	Minimum: -28.3
Range: -28.3271694087-0	Maximum: 0
	Mean: -28.2
	Standard deviation: 0.3

Hexagon Identifier (HexagonId)

File: Locations dataset

Hexagon Identifier (HexagonId)

File: Locations dataset

Overview

Type: Discrete	Valid cases: 7567
Format: numeric	Invalid: 1
Decimals: 0	

Number of Individuals per LocationId (Num_Individuals)

File: Locations dataset

Overview

Type: Continuous	Valid cases: 7568
Format: numeric	Invalid: 0
Decimals: 0	Minimum: 0
Range: 1-43	Maximum: 21
	Mean: 3.9
	Standard deviation: 2.4

Number of Households per LocationId (Num_Households)

File: Locations dataset

Overview

Type: Continuous	Valid cases: 7568
Format: numeric	Invalid: 0
Decimals: 0	Minimum: 0
Range: 1-40	Maximum: 10
	Mean: 1.1
	Standard deviation: 0.6

Study Identifier (StudyId)

File: Individuals dataset

Overview

Type: Discrete
Format: character
Width: 4

Valid cases: 28626
Invalid: 0

Individual Identifier (IndividualId)

File: Individuals dataset

Overview

Type: Continuous
Format: numeric
Decimals: 0
Range: 1-28596

Valid cases: 28626
Invalid: 0
Minimum: 1
Maximum: 28613
Mean: 14306.1
Standard deviation: 8260.4

Household Identifier (HouseholdId)

File: Individuals dataset

Overview

Type: Continuous
Format: numeric
Decimals: 0
Range: 1-7799

Valid cases: 28626
Invalid: 0
Minimum: 1
Maximum: 7800
Mean: 3894.4
Standard deviation: 2252.4

Sex (Sex)

File: Individuals dataset

Overview

Type: Discrete
Format: character
Width: 3

Valid cases: 28626
Invalid: 0

DateOfBirth (DateOfBirth)

File: Individuals dataset

Overview

Type: Discrete
Format: character

Valid cases: 28625
Minimum: NaN
Maximum: NaN

Highest Education Level Code (HighestEducationLevel)

File: Individuals dataset

Overview

Highest Education Level Code (HighestEducationLevel)

File: Individuals dataset

Type: Continuous
Format: numeric
Decimals: 0
Range: 1-99

Valid cases: 25812
Invalid: 2814
Minimum: 1
Maximum: 99
Mean: 6.6
Standard deviation: 16.9

Highest Education Level Description (HighestEducationLevelDesc)

File: Individuals dataset

Overview

Type: Discrete
Format: character
Width: 30

Valid cases: 25812
Invalid: 0

Current Marital Status Code (CurrentMaritalStatus)

File: Individuals dataset

Overview

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

Valid cases: 25812
Invalid: 2814

Current Marital Status Description (CurrentMaritalStatusDesc)

File: Individuals dataset

Overview

Type: Discrete
Format: character
Width: 20

Valid cases: 25812
Invalid: 0

Currently In Employment Code (CurrentlyInEmployment)

File: Individuals dataset

Overview

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

Valid cases: 25812
Invalid: 2814

Currently In Employment Description (CurrentlyInEmploymentDesc)

File: Individuals dataset

Overview

Type: Discrete
Format: character
Width: 16

Valid cases: 25812
Invalid: 0

Study Identifier (StudyId)

File: Household dataset

Overview

Type: Discrete
Format: character
Width: 4

Valid cases: 7803
Invalid: 0

Household Identifier (HouseholdId)

File: Household dataset

Overview

Type: Continuous
Format: numeric
Decimals: 0
Range: 1-7799

Valid cases: 7803
Invalid: 0
Minimum: 1
Maximum: 7800
Mean: 3899.6
Standard deviation: 2251.9

Location Identifier (LocationId)

File: Household dataset

Overview

Type: Discrete
Format: numeric
Decimals: 0

Valid cases: 7802
Invalid: 1

Cluster Identifier (ClusterId)

File: Household dataset

Overview

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

Valid cases: 7802
Invalid: 1
Minimum: 1
Maximum: 22
Mean: 11.2
Standard deviation: 5.5

Asset Score (AssetScore)

File: Household dataset

Overview

Type: Discrete
Format: numeric
Decimals: 0
Range: 1-5

Valid cases: 7736
Invalid: 67

Asset Score Index (AssetScoreIndex)

File: Household dataset

Overview

Asset Score Index (AssetScoreIndex)

File: Household dataset

Type: Discrete

Format: character

Width: 7

Valid cases: 7736

Invalid: 0

Individual Identifier (IndividualId)

File: Linkage to Care Dataset.

Overview

Type: Continuous	Valid cases: 26393
Format: numeric	Invalid: 0
Decimals: 0	Minimum: 1
Range: 1-28606	Maximum: 28613
	Mean: 14321.9
	Standard deviation: 8256.7

Test Agreed at First Positive Test (TestAgreedatFirstPositive)

File: Linkage to Care Dataset.

Overview

Type: Discrete	Valid cases: 7502
Format: numeric	Invalid: 18891
Decimals: 0	

Date of First Positive Test (FirstPositiveDate)

File: Linkage to Care Dataset.

Overview

Type: Discrete	Valid cases: 7502
Format: character	Minimum: NaN
	Maximum: NaN

Test Agreed at Last Contact (TestAgreedatLastContact)

File: Linkage to Care Dataset.

Overview

Type: Discrete	Valid cases: 26393
Format: numeric	Invalid: 0
Decimals: 0	

Date of Last Household Contact (LastHHContactDate)

File: Linkage to Care Dataset.

Overview

Type: Discrete	Valid cases: 26393
Format: character	Minimum: NaN
	Maximum: NaN

Date of First Clinic Visit (FirstClinicDate)

File: Linkage to Care Dataset.

Overview

Type: Discrete	Valid cases: 3001
Format: character	Minimum: NaN
	Maximum: NaN

First Clinic LocationId (FirstClinicLocationId)

File: Linkage to Care Dataset.

Overview

Type: Continuous	Valid cases: 3001
Format: numeric	Invalid: 23392
Decimals: 0	Minimum: 50
Range: 177-11633	Maximum: 11148
	Mean: 4607.7
	Standard deviation: 4043.1

Daet of Last Clinic Visit (LastClinicDate)

File: Linkage to Care Dataset.

Overview

Type: Discrete	Valid cases: 3001
Format: character	Minimum: NaN
	Maximum: NaN

Last Clinic LocationId (LastClinicLocationId)

File: Linkage to Care Dataset.

Overview

Type: Continuous	Valid cases: 261
Format: numeric	Invalid: 26132
Decimals: 0	Minimum: 50
Range: 177-11633	Maximum: 11148
	Mean: 5950.3
	Standard deviation: 4131.1

Clinic Visited (ClinicVisited)

File: Linkage to Care Dataset.

Overview

Type: Discrete	Valid cases: 26393
Format: character	Invalid: 0
Width: 3	

CurrentDate (CurrentDate)

File: Linkage to Care Dataset.

Overview

Type: Discrete	Valid cases: 26393
Format: numeric	Invalid: 0
Decimals: 0	
Range: 20160202-20160202	

Number of Days to link to Clinic (DaystoClinicLink)

File: Linkage to Care Dataset.

Overview

Number of Days to link to Clinic (DaystoClinicLink)

File: Linkage to Care Dataset.

Type: Continuous	Valid cases: 3001
Format: numeric	Invalid: 23392
Decimals: 0	Minimum: -503
Range: -503-1113	Maximum: 1113
	Mean: 94.3
	Standard deviation: 158.8

Number of Days Since Last Home Contact (DaysSinceLastContactHome)

File: Linkage to Care Dataset.

Overview

Type: Continuous	Valid cases: 26393
Format: numeric	Invalid: 0
Decimals: 0	Minimum: 0
Range: 1-1424	Maximum: 1433
	Mean: 273
	Standard deviation: 276.2

Number of Days Since Last CLinic Visit (DaysSinceLastCLinicVisit)

File: Linkage to Care Dataset.

Overview

Type: Continuous	Valid cases: 3001
Format: numeric	Invalid: 23392
Decimals: 0	Minimum: 0
Range: 0-1413	Maximum: 1422
	Mean: 119.3
	Standard deviation: 218

Number of Days Not Linked to Clinic (DaysNotLinkedtoClinic)

File: Linkage to Care Dataset.

Overview

Type: Continuous	Valid cases: 7502
Format: numeric	Invalid: 18891
Decimals: 0	Minimum: 0
Range: 0-1421	Maximum: 1430
	Mean: 323.9
	Standard deviation: 352.8

HIV Status (HIVStatus)

File: Linkage to Care Dataset.

Overview

Type: Discrete	Valid cases: 26393
Format: character	Invalid: 0
Width: 8	

Study Identifier (StudyId)

File: DataCollectionCycle dataset

Overview

Type: Discrete	Valid cases: 112
Format: character	Invalid: 0
Width: 4	

DataCollectionCycle (DataCollectionCycle)

File: DataCollectionCycle dataset

Overview

Type: Discrete	Valid cases: 112
Format: numeric	Invalid: 0
Decimals: 0	
Range: 1-7	

Cluster Identifier (ClusterId)

File: DataCollectionCycle dataset

Overview

Type: Continuous	Valid cases: 112
Format: numeric	Invalid: 0
Decimals: 0	Minimum: 1
Range: 1-22	Maximum: 22
	Mean: 10.3
	Standard deviation: 6.4

Start Date (StartDate)

File: DataCollectionCycle dataset

Overview

Type: Discrete	Valid cases: 112
Format: character	Minimum: NaN
	Maximum: NaN

End Date (EndDate)

File: DataCollectionCycle dataset

Overview

Type: Discrete	Valid cases: 112
Format: character	Minimum: NaN
	Maximum: NaN

Individual Identifier (IndividualId)

File: Migration dataset

Overview

Type: Continuous	Valid cases: 11642
Format: numeric	Invalid: 0
Decimals: 0	Minimum: 1
Range: 4-28603	Maximum: 28611
	Mean: 14197.4
	Standard deviation: 8252.4

Date of Migration (MigrateDate)

File: Migration dataset

Overview

Type: Discrete	Valid cases: 7687
Format: character	Minimum: NaN
	Maximum: NaN

Date of Migration - Text (MigrateVPDate)

File: Migration dataset

Overview

Type: Discrete	Valid cases: 11642
Format: character	Invalid: 0
Width: 8	

Destination Code (DestinationCode)

File: Migration dataset

Overview

Type: Discrete	Valid cases: 11642
Format: character	Invalid: 0
Width: 7	

Documentation

Other materials

Data Everywhere datasets -TasP - Documentation

Title	Data Everywhere datasets -TasP - Documentation
Filename	Data Everywhere datasets -TasP - Documentation.pdf
