

Mapping Livestock Keepers and Their Herds Across Africa Based on Household Surveys

Stephen Oloo
Catherine Pfeifer

FOSS4G 2018
Dar es Salaam
30th August, 2018



RESEARCH
PROGRAM ON
Livestock



Content

Background

- Introduction
- DHS data
- Automatic Processing

Demonstration

- DHS Data Download
- Data Extraction
- Study Area
- DHS Data Preparation
- Data modeling and prediction

Way Forward

- Data Adjustment
- Other Applications

Content

Background

- Introduction
- DHS data
- Automatic Processing

Demonstration

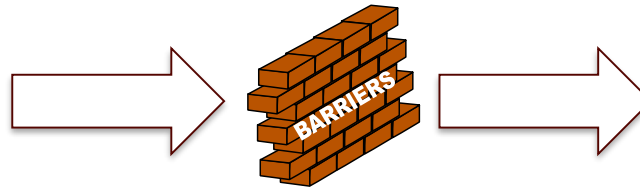
- DHS Data Download
- Data Extraction
- Study Area
- DHS Data Preparation
- Data modeling and prediction

Way Forward

- Data Adjustment
- Other Applications

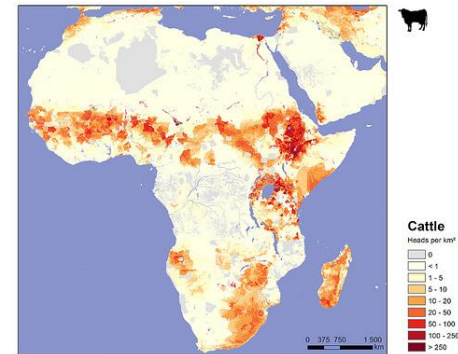
Introduction

Socioeconomic Data



- Cost
- Time

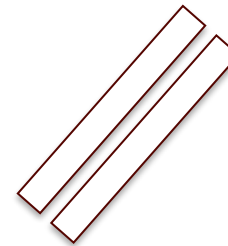
Map



Free Data



Open source software



Demographic Health Survey (DHS) data

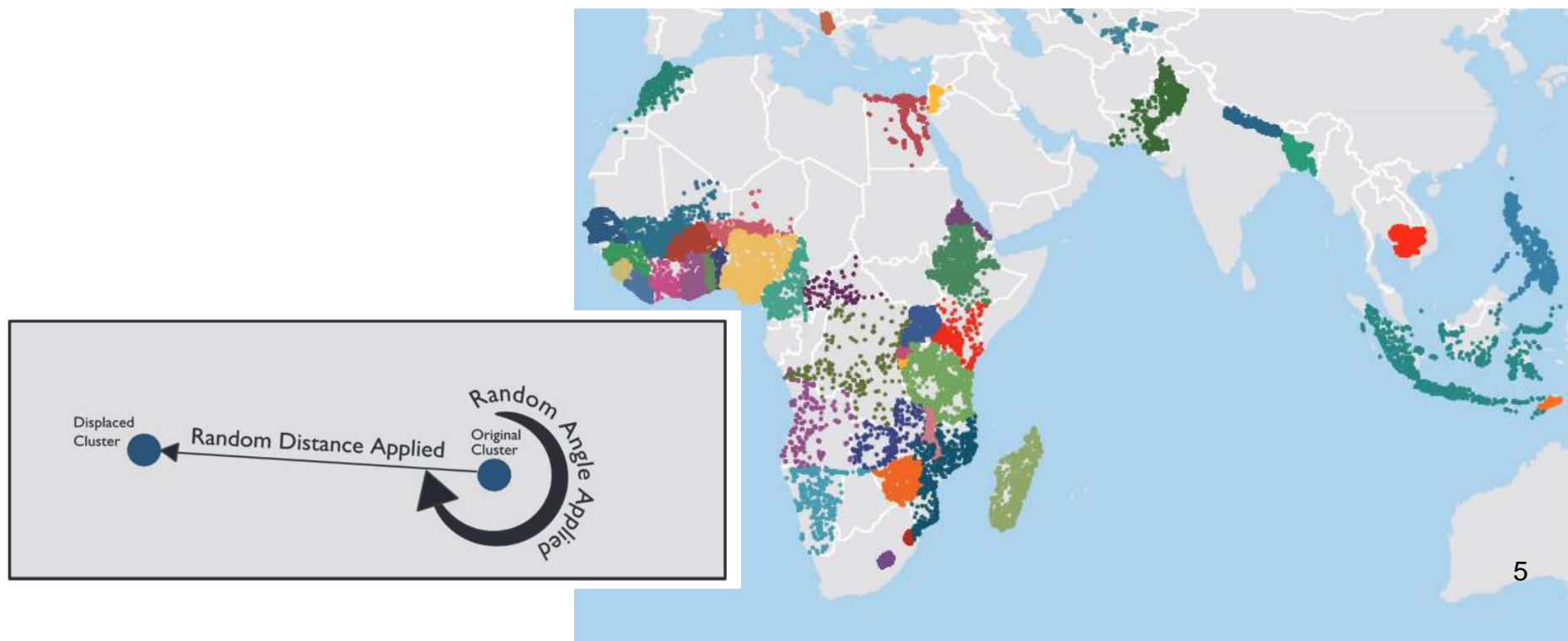
- DHS program was initiated in 1984 by USAID
- DHS program carries out nationally representative demographic and health surveys in developing countries in collaboration with the statistical body of the respective countries
- Main aim of DHS data is to support informed decision making in developing countries
- So far, more than 300 nationally representative surveys have been carried out in about 90 countries

Spatiality of DHS data

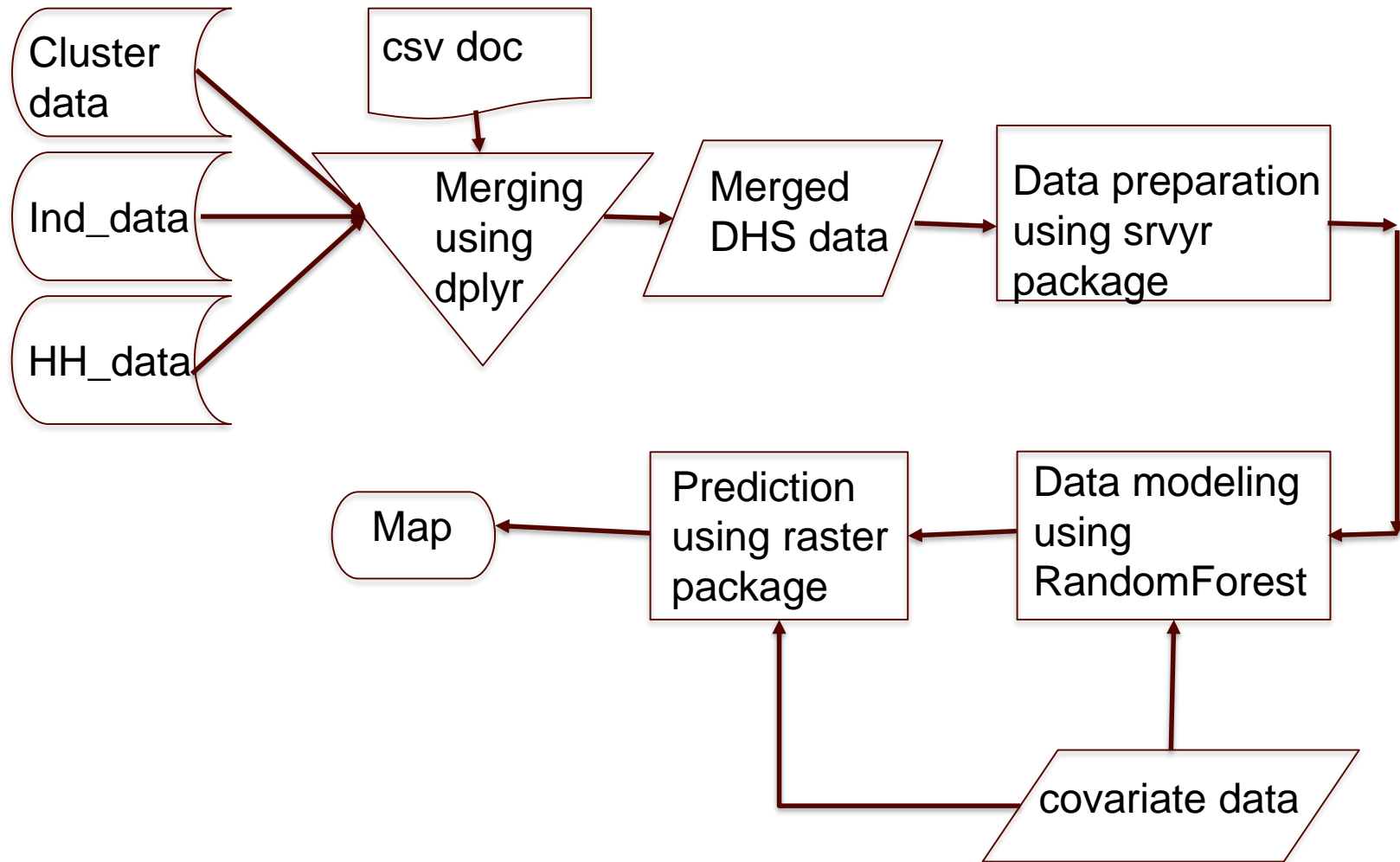
Cluster coordinates :

Cluster is a group of 20-40 households

For the participants to remain anonymous the GPS coordinates for the cluster are randomly shifted (2km in urban, 5km in rural area)



Automatic Processing



Content

Background

- Introduction
- DHS data
- Automatic Processing

Demonstration

- DHS Data Download
- Data Extraction
- Study Area
- DHS Data Preparation
- Data modeling and prediction

Way Forward

- Data Adjustment
- Other Applications

DHS Data Download

The DHS Program > Data > Available Datasets

Data Available Datasets

ABOUT THE DATA

Getting Started

Data Collection

Data Processing

Recommended Citations

UNDERSTANDING SURVEY STATISTICS

Data Quality and Use

Data Tabulation Plan

Survey Indicators

Data Tools and Manuals

Online Guide to DHS Statistics

WORKING WITH DATASETS

Using Datasets for Analysis

Model Datasets

Data Variables & Definitions

Dataset Types

File Formats

File Types & Names

Merging Datasets

Calculating the Age of Children

Contraceptive Calendar Tutorial

experience on our website. [Learn more](#)

AFGHANISTAN

Survey	Type	Phase	Recode	Survey Datasets	GPS Datasets	HIV/Other Biomarkers Datasets	SPA Datasets
--------	------	-------	--------	-----------------	--------------	-------------------------------	--------------

Survey	Type	Name	Date modified	Type	Size
--------	------	------	---------------	------	------

Afghanistan 2015	Standard DHS	ALB_2008_50	1/18/2017 10:13 AM	File folder	
------------------	--------------	-------------	--------------------	-------------	--

Afghanistan 2010	Special	ARM_2010_61	1/18/2017 10:12 AM	File folder	
------------------	---------	-------------	--------------------	-------------	--

		BDI_2010_61	9/8/2017 3:33 PM	File folder	
--	--	-------------	------------------	-------------	--

		BEN_2006_51	11/23/2016 12:19 ...	File folder	
--	--	-------------	----------------------	-------------	--

		BEN_2011_61	11/23/2016 12:19 ...	File folder	
--	--	-------------	----------------------	-------------	--

		BFA_2010_62	5/12/2017 4:36 PM	File folder	
--	--	-------------	-------------------	-------------	--

		BFA_2014_70_MIS	11/7/2017 12:24 PM	File folder	
--	--	-----------------	--------------------	-------------	--

		BGD_2007_51	1/18/2017 10:14 AM	File folder	
--	--	-------------	--------------------	-------------	--

		BGD_2011_61	1/18/2017 10:15 AM	File folder	
--	--	-------------	--------------------	-------------	--

		BGD_2014_70	1/18/2017 10:15 AM	File folder	
--	--	-------------	--------------------	-------------	--

		BOL_2008_51	1/18/2017 10:17 AM	File folder	
--	--	-------------	--------------------	-------------	--

		CIV_2011_62			
--	--	-------------	--	--	--

		CMR_2011_61			
--	--	-------------	--	--	--

		COD_2007_50			
--	--	-------------	--	--	--

		COD_2013_61			
--	--	-------------	--	--	--

		COG_2005_51			
--	--	-------------	--	--	--

		COG_2011_60			
--	--	-------------	--	--	--

		COL_2010_61			
--	--	-------------	--	--	--

		COM_2012_61			
--	--	-------------	--	--	--

		DOM_2007_52			
--	--	-------------	--	--	--

		DOM_2013_61			
--	--	-------------	--	--	--

		EGY_2008_5A			
--	--	-------------	--	--	--

		EGY_2014_61			
--	--	-------------	--	--	--

		ETH_2011_61			
--	--	-------------	--	--	--

		ETH_2016_70			
--	--	-------------	--	--	--

		GAB_2012_61			
--	--	-------------	--	--	--

		GHA_2008_5A			
--	--	-------------	--	--	--

		GHA_2014_71			
--	--	-------------	--	--	--

		GIN_2012_62			
--	--	-------------	--	--	--

		GMB_2013_60			
--	--	-------------	--	--	--

		GUY_2009_51			
--	--	-------------	--	--	--

Data download

[www.dhsprogram.com]

How the data is saved

Inside country folders

Name

BJIR61FL.DOC

GE.cpg

GE.dbf

GE.prj

GE.sbn

GE.sbx

GE.shp

GE.shp.xml

GE.shx

HR.SAV

IR.SAV

sdr_subnational_boundaries.cpg

sdr_subnational_boundaries.dbf

sdr_subnational_boundaries.prj

sdr_subnational_boundaries.sbn

sdr_subnational_boundaries.sbx

sdr_subnational_boundaries.shp

sdr_subnational_boundaries.shp.xml

sdr_subnational_boundaries.shx

Date modified

2/12/2014 5:53 PM

1/31/2014 3:05 PM

1/31/2014 3:05 PM

1/31/2014 3:05 PM

1/31/2014 3:05 PM

1/31/2014 3:05 PM

1/31/2014 3:05 PM

2/12/2014 3:08 PM

1/31/2014 3:05 PM

2/11/2014 2:17 PM

2/11/2014 2:20 PM

4/23/2015 4:29 PM

4/23/2015 4:29 PM

4/23/2015 4:29 PM

4/23/2015 4:29 PM

4/23/2015 4:29 PM

4/23/2015 4:29 PM

4/23/2015 4:29 PM

4/23/2015 4:29 PM

Type

Microsoft Word 9...

CPG File

DBF File

PRJ File

SBN File

SBX File

SHP File

XML Document

SHX File

SAV File

SAV File

CPG File

DBF File

PRJ File

SBN File

SBX File

SHP File

XML Document

SHX File

Size

114 KB

1 KB

2,346 KB

1 KB

7 KB

1 KB

21 KB

25 KB

6 KB

72,773 KB

76,742 KB

1 KB

48 KB

1 KB

1 KB

1 KB

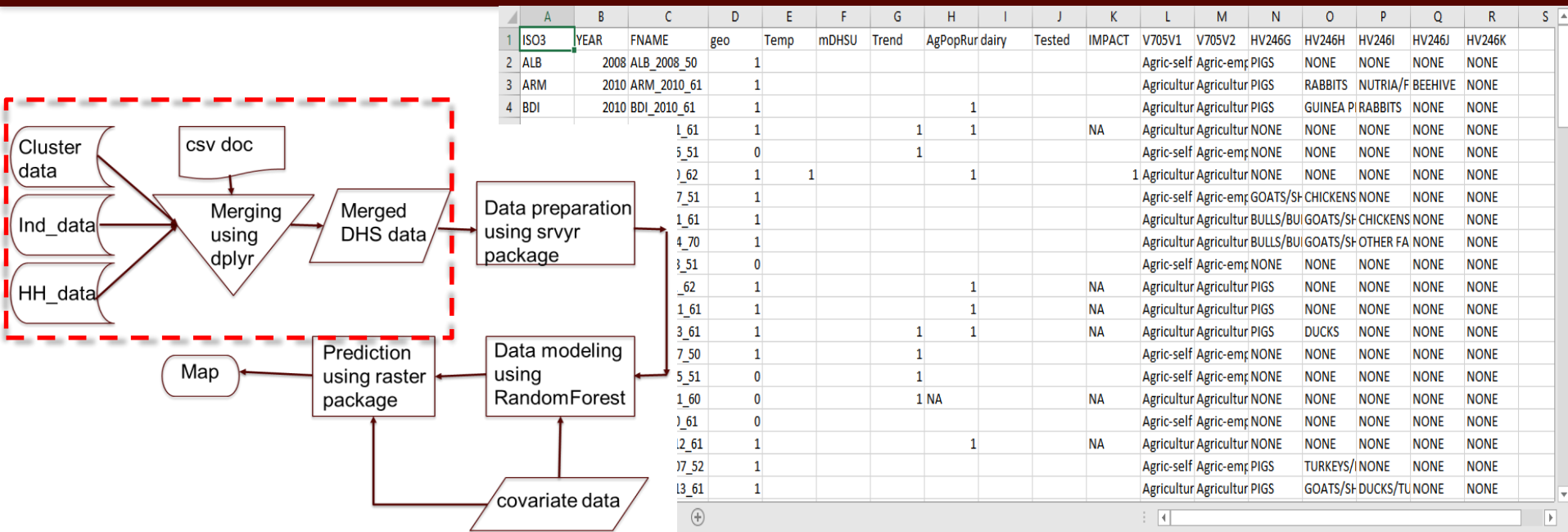
873 KB

12 KB

1 KB

8

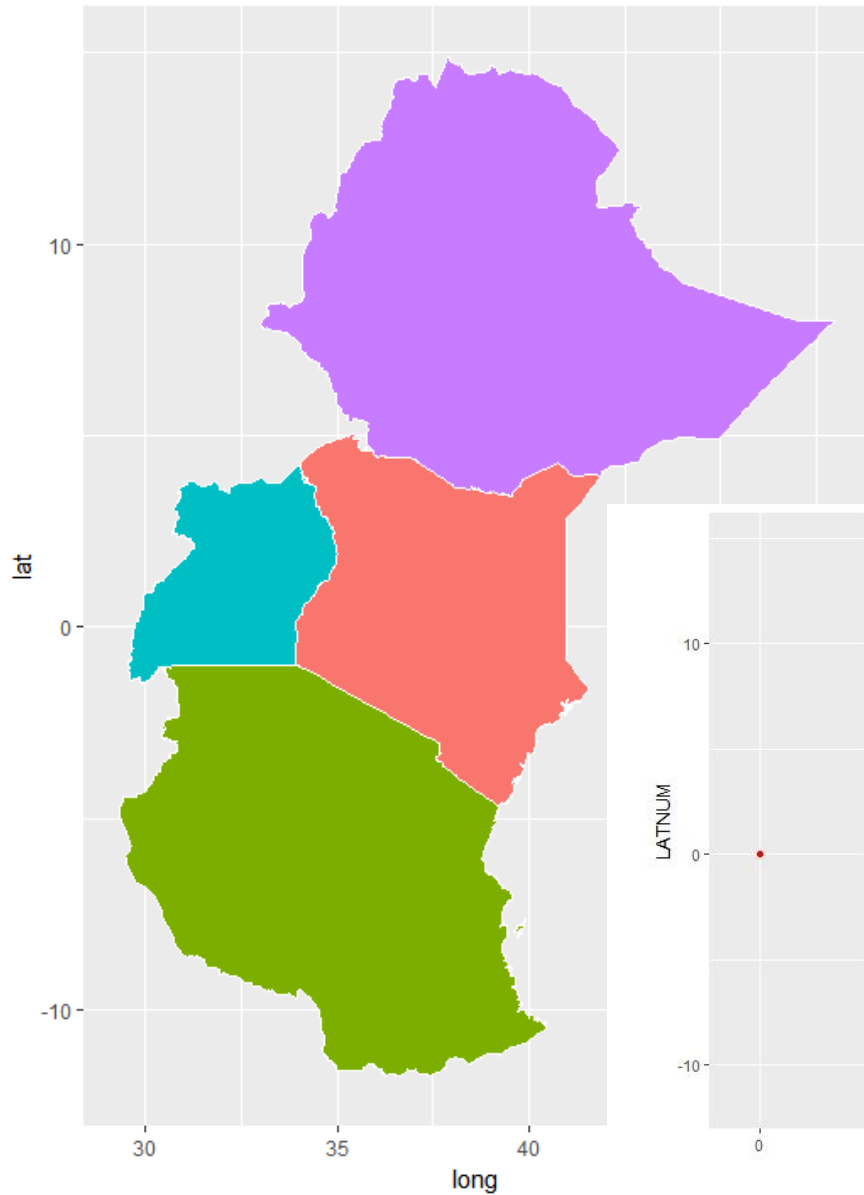
Data Extraction



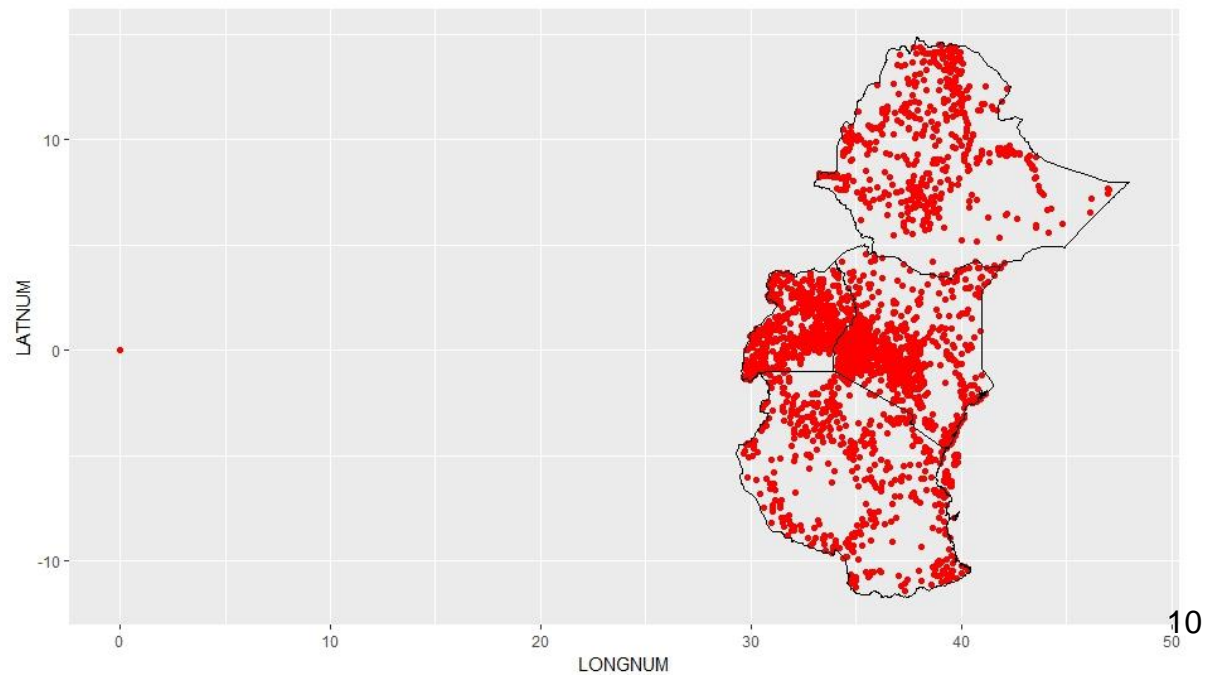
Main data extracted

- Household economic activity
- Country specific livestock variable

Study Area

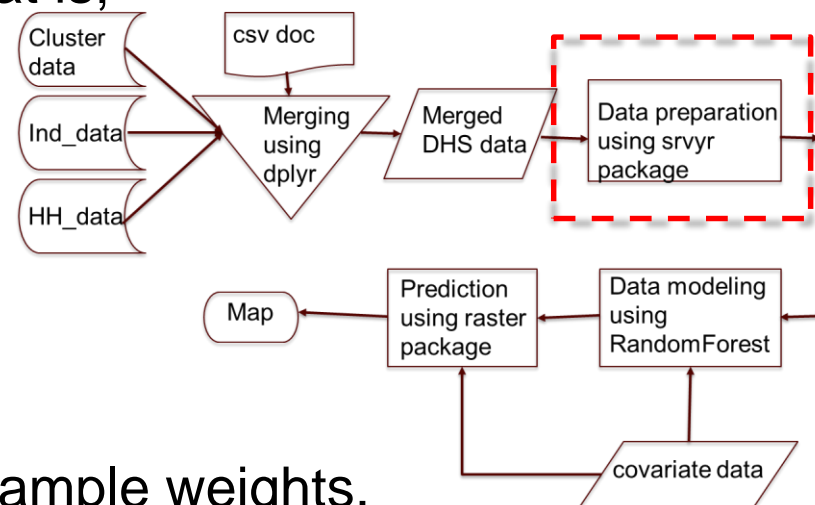


- East Africa
- Cluster map

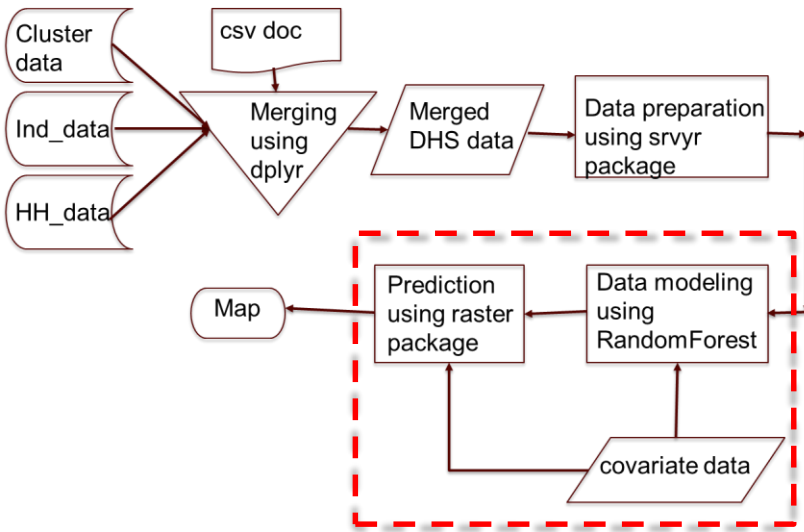


DHS Data Preparation

- DHS surveys are complex surveys that is;
 - They are stratified
 - They are clustered
 - They are random
- Sampling error was corrected using sample weights, this was done using srvyr package
- Data aggregated per cluster for household that have got at least one cattle

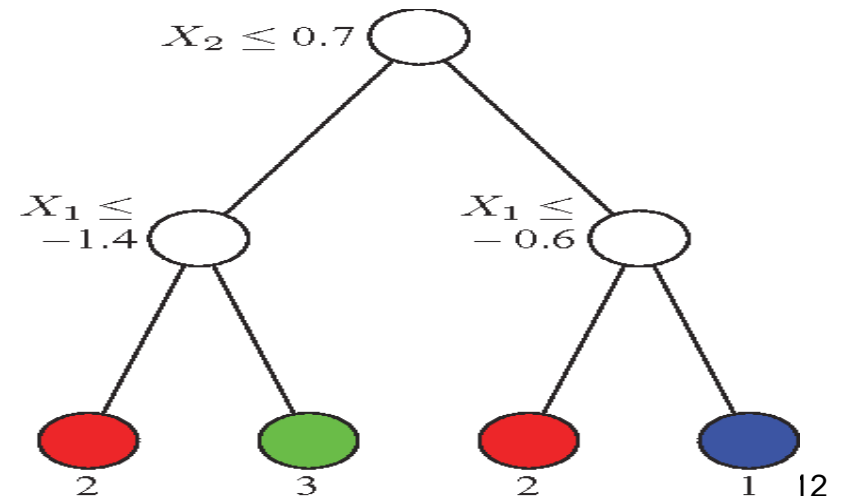
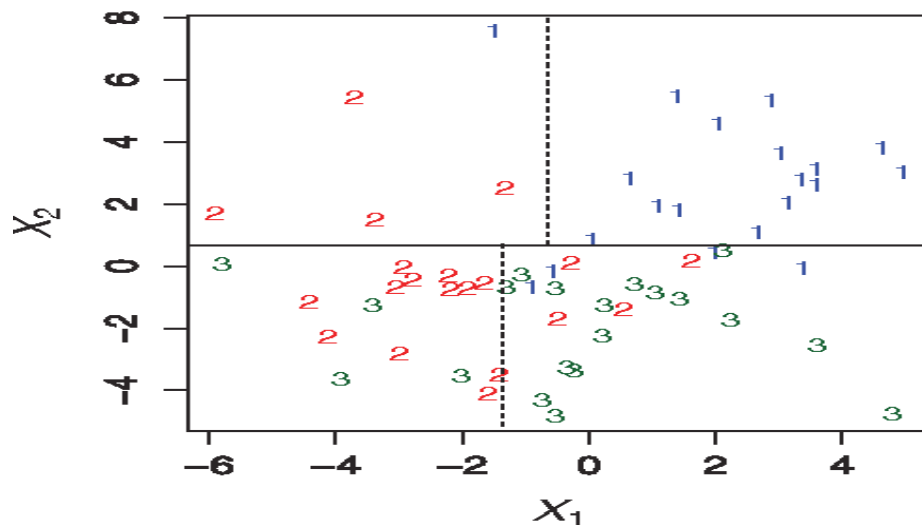


Data Modelling and Prediction

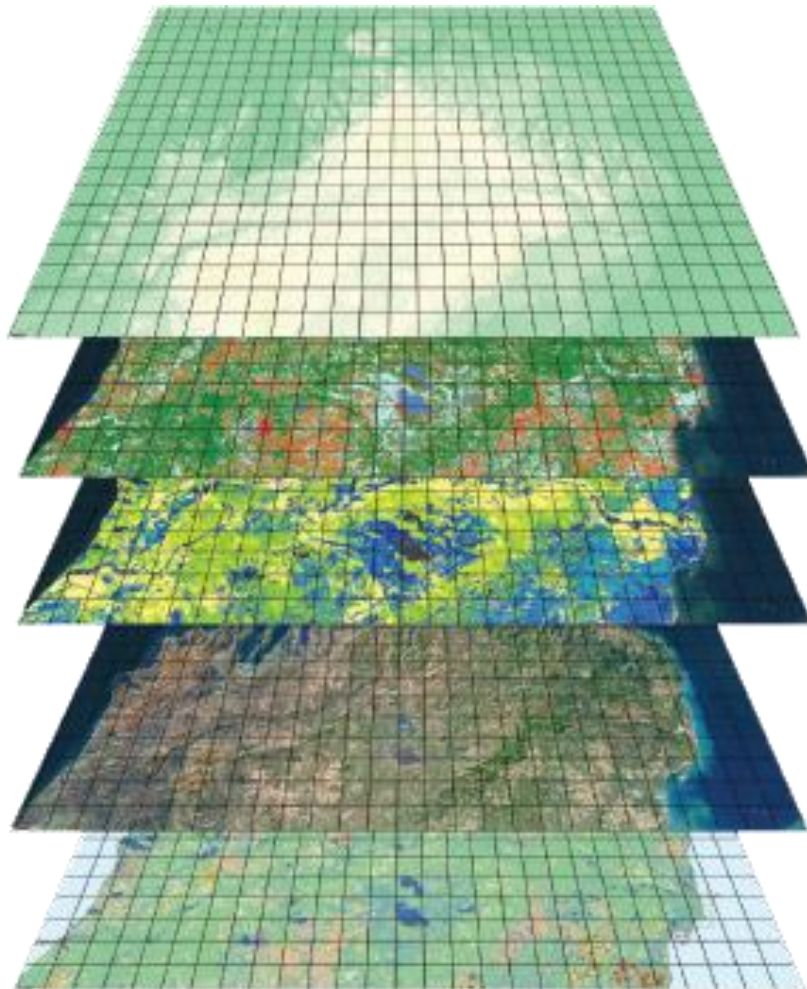


Random Forest

- It is a machine learning technique that uses decision trees
- At each node there is split based on how much the new partition reduces the model error



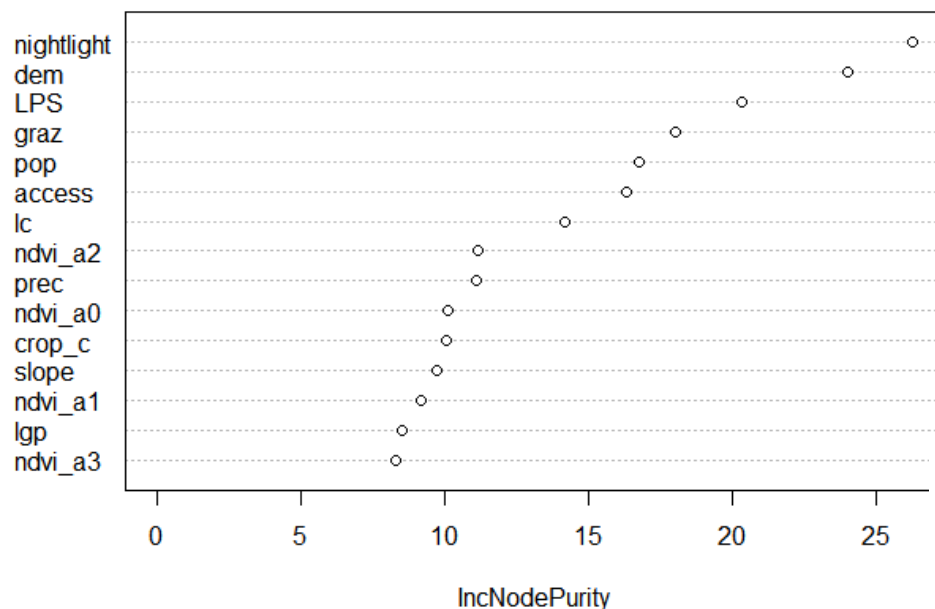
Data Modelling and Prediction: Covariates



- Livestock production systems
- Precipitation
- DEM
- Population
- Land Cover
- Night Light
- NDVI_a0
- NDVI_a1
- NDVI_a2
- NDVI_a3
- Crop Cover
- Travel Time
- Slope

Data Modelling and Prediction

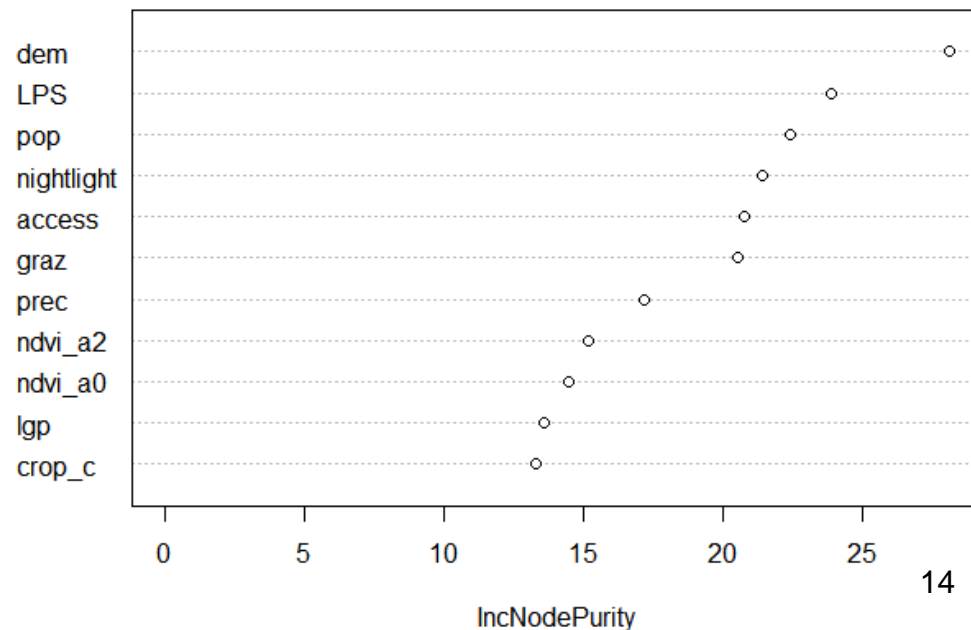
DHS_data_rf



Full Model

- Mean of squared residuals: 0.03818643
- % Var explained: 54.84

DHS_data_rf2



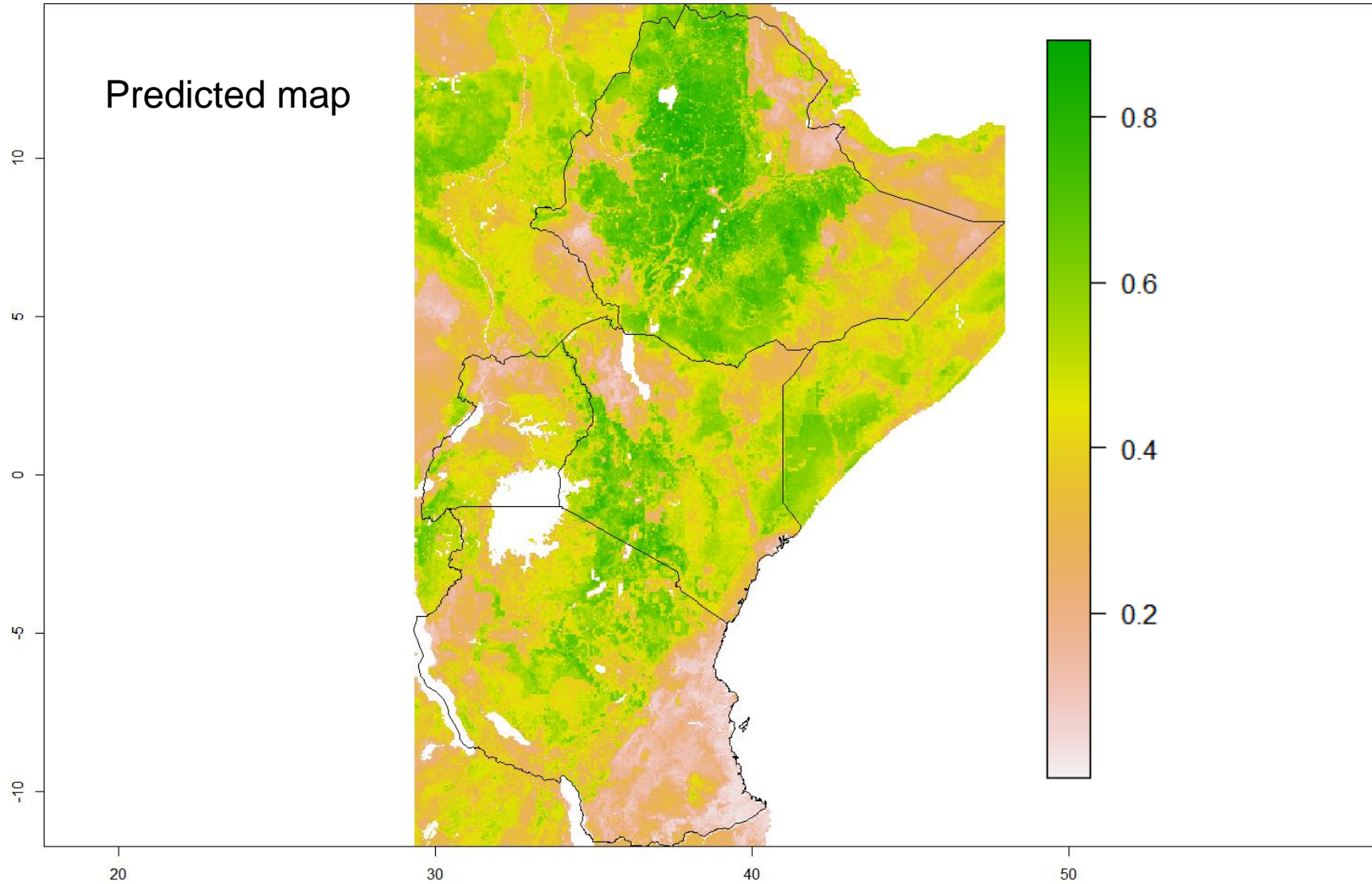
Model after selection using VSURF

package

- Mean of squared residuals: 0.0374028
- % Var explained: 55.76

Map

Predicted map



Content

Background

- Introduction
- DHS data
- Automatic Processing

Demonstration

- DHS Data Download
- Data Extraction
- Study Area
- DHS Data Preparation
- Data modeling and prediction

Way Forward

- Data Adjustment
- Other Applications

Way Forward

Data Adjustment

- Adjust the dataset to conform with the United Nations Population Data
- Where there is recent census data, adjust the dataset to conform with the census data
- Incorporate temporal aspect of the data

Other Application

- The dataset and code can be used to map other socioeconomic such mobile phone ownership, access to information, nutrition

*“The goal is to turn data into information,
and information into insight”*

By Carly Fiorina

better lives through livestock

ilri.org

ILRI thanks all donors and organizations which globally support its work through their contributions to the [CGIAR system](#)

Patron: Professor Peter C Doherty AC, FAA, FRS

Animal scientist, Nobel Prize Laureate for Physiology or Medicine—1996

Box 30709, Nairobi 00100 Kenya
Phone +254 20 422 3000
Fax +254 20 422 3001
Email ilri-kenya@cgiar.org

ilri.org
better lives through livestock

ILRI is a CGIAR research centre

Box 5689, Addis Ababa, Ethiopia
Phone +251 11 617 2000
Fax +251 11 667 6923
Email ilri-ethiopia@cgiar.org

ILRI has offices in East Africa • South Asia • Southeast and East Asia • Southern Africa • West Africa



This presentation is licensed for use under the Creative Commons Attribution 4.0 International Licence.