

https://dspace.org/

Information feedback to local communities of HDSS sites using infographics: the case of niakhar HDSS site

Dione, Djibril

2017
http://hdl.handle.net/10539/23143

Downloaded from a DSpace repository. https://dspace.org/

Information Feedback to Local Communities of HDSS Sites using Infographics: The Case of Niakhar HDSS Site



DJIBRIL DIONE

STUDENT NUMBER: 981721

A research report submitted to the Faculty of Health Science in partial fulfillment of the requirements for the degree of *Master of Science (MSc)* in Epidemiology - Research Data Management

June 2017

Djibril Dione . *2017.*
Information Feedback to Local Communities of HDSS Sites using Infographics: The Case of Niakhar HDSS Site.

Copyright © University of the Witwatersrand, Johannesburg, South Africa.

All rights reserved. No part of this protocol may be stored in a retrieval system, transmitted, or reproduced, in any form or by any means, including but not limited to photocopy, photograph, magnetic or other record, without prior agreement and written permission of the copyright holder.

SUPERVISOR:
Dr. Gideon Nimako, University of The Witwatersrand

SUPPORTED BY:
INDEPTH Network
University of the Witwatersrand
Niakhar HDSS

DECLARATION

I hereby declare that this research work is project work carried out by me under the guidance of Dr. Gideon Nimako. I have taken care in all respect to honour the intellectual property right and have acknowledged the contribution of others for using them in this work and further declare that the work reported in this project has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

Parktown, Johannesburg, June 2017



Djibril Dione

DEDICATION

I dedicate my research work to To my wife who has been the source of my strength throughout this program, for her patience and support during my study, my family and friends. I also dedicate this work to my Father who has encouraged me all the way and whose encouragement has ensured this completion. To my Daughter Khadija who has been affected in every way possible. Thank you all, God bless you.

ABSTRACT

Health and Demographic Surveillance System (HDSS) sites are institutions that primarily collect periodic demographic and health related data within a defined geographical area. The data are then analysed and the research findings are published in peer-reviewed journals or presented at conferences and workshops. However, In some sites there are processes for disseminating the research findings to the participating communities. The dissemination of research findings using infographics can enable health care service providers and community decision makers to make effective use and incorporate such findings into their strategies, policies and planning to improve heath outcome of the population. In this project, we have implemented a visualization web platform that can be used by researchers, community decision makers and public health policy-makers to better identify trends associated with the research findings. We implemented this platform on top of the core HDSS dataset. An Extract Transform and Load (ETL) process feeds processed data to our layer which then provides the requisite utility tools for visualization. We have provided a framework that allows other cohort studies that utilize the HDSS core dataset as baseline data to be plugged in for additional infographics displays. We tested and prototyped our tool using Niakhar HDSS Site core HDSS datasets. The project also provides some generic guidelines that allow this tool to be used in other HDSS sites.

ACKNOWLEDGMENTS

I would like to express my gratitude to the INDEPTH Network who supported and sponsored the coursework of this master programme. I am also grateful to my supervisor Dr Gideon Nimako for his guidance during this program. To URMITE - Research Institute for Sustainable Development for providing their Data. To My wife and my daughter for their support and for allowing me to progress in my career development. I would also like to thank the R community for their contribution and package development. I would also Like to thank Dr. Valerie Delaunay for her guidance during this project and for her availability and all the URMITE Team.

CONTENTS

DECLARATION iii
DEDICATION iv
ABSTRACT v
ACKNOWLEDGMENTS vi
ACRONYMS viii
List of Figures viii
List of Tables viii
1 INTRODUCTION 1
1.1 Problem Statement . . . 2
1.2 Motivation . . . 3
1.3 Contributions . . . 4
1.4 Outline of Research Report . . . 4
2 BACKGROUND AND RELATED WORK 6
2.1 Use Modern Technology in Research Findings Dissemination . . . 6
2.2 Use of Visualisation in Communicating Research Findings . . . 7
2.3 Impact of Dissemination of Research Result for Public Health Policies . . . 8
2.4 Overview of Dissemination Channels in Health Research . . . 8
2.4.1 Websites or Web Applications . . . 9
2.4.2 Scientific Journal Publications . . . 10
2.4.3 Mass Media . . . 12
2.4.4 Community Meetings . . . 13
2.4.5 Summary of the Literature . . . 13
3 AN INFOGRAPHICS PLATFORM FOR RESEARCH INFORMATION FEEDBACK 14
3.1 Overview of Technologies Used . . . 14
3.1.1 Statistical Software, R . . . 14
3.1.2 R Shiny . . . 15
3.1.3 Microsoft SQL Server . . . 15
3.2 Application Architecture . . . 15
3.3 Niakhar HDSS . . . 16
3.4 Implementation and Pre-processing Considerations . . . 17
3.5 An Overview of the Infographic Platform . . . 17
3.5.1 Module 1: Mortality Visualisation Utility . . . 18
3.5.2 Module 2: Fertility Visualisation Utility . . . 20
3.5.3 Module 3: Migration Visualisation Utility . . . 21
3.5.4 Module 4: Population Visualisation Utility . . . 21
3.5.5 Module 5: Cohort Studies Visualisation Utility . . . 22
3.6 Variables within HDSS Dataset . . . 23
3.7 Ethical consideration . . . 23
4 PRESENTATION LAYER OF THE INFOGRAPHICS PLATFORM 25
4.1 Presentation of the Shiny App . . . 25
4.2 Descriptive Visualisations . . . 26
4.2.1 Use Case: Mortality Data Visualisation . . . 26
4.2.2 Use Case: Fertility Visualisation . . . 29
4.2.3 Use Case: Immigration Visualisation . . . 31
4.3 Plugin for Cohort Studies . . . 35

4.4 Combining Dissemination Strategies 38
5 IMPLEMENTATION AND DEPLOYMENT CONSIDERATIONS 40
5.1 Deployment Considerations in other HDSS Sites 40
5.2 Infrastructure Consideration 42
5.3 Use of the Infographic Platform to Enhance Community Engagement 42
6 CONCLUSION, RECOMMENDATIONS AND FUTURE DIRECTIONS 44
A PLAGIARISM DECLARATION 46
B ETHICS CLEARANCE CERTIFICATE 47
C SOURCE CODE 48

ACRONYMS

API Application Programming Interface

INDEPTH International Network for the Demographic Evaluation of Populations and Their Health

HDSS Health and Demographic Surveillance System

ICT Information Communication Technologies

CDC The Centre for Disease Control and Prevention

GIS Geographic Information System

DHS Demographic and Health Survey

ANSD Agence National de la Statistique et de la Demographie

DBMS Database Management System

SQL Structured Query Language

ODBC Open Database Connectivity

RODBC R plugin for Open Database Connectivity

GPL General Public License

HTML Hyper-Text Marckup Language

ETL Extract Transform and Load

LIST OF FIGURES

Figure 1 Application Architecture 16
Figure 2 Mortality rate trend 26
Figure 3 Crude Mortality Rate 27
Figure 4 Age-specific Mortality 2000 27
Figure 5 Age-specific Mortality 2005 28
Figure 6 Age-specific Mortality 2010 28
Figure 7 Age-specific Mortality 2014 29
Figure 8 Crude Birth Rate Trend - Niakhar, 2000-2014 30
Figure 9 Total Fertility Rate Trend - Niakhar, 2000-2014 30
Figure 10 Fertility Rate Per Age Group -Niakhar, 2000-2014 31
Figure 11 Crude In-Migration Rate - Niakhar, 2000-2014 32
Figure 12 Crude Out-Migration Rate - Niakhar, 2000-2014 32

Figure 13 In-Migration Rate by Age Group - Niakhar, 2000 33
Figure 14 In-Migration Rate by Age Group - Niakhar, 2005 33
Figure 15 In-Migration Rate by Age Group - Niakhar, 2010 34
Figure 16 In-Migration Rate by Age Group - Niakhar, 2014 34
Figure 17 Out-Migration Rate by Age Group - Niakhar, 2000 35
Figure 18 Out-Migration Rate by Age Group - Niakhar, 2005 35
Figure 19 Out-Migration Rate by Age Group - Niakhar, 2010 36
Figure 20 Out-Migration Rate by Age Group - Niakhar, 2014 36
Figure 21 Population Age Pyramid 37
Figure 22 Cohort Study Plugin 37
Figure 23 Malaria Incidence Rate from 2000-2016 38
Figure 24 INDEPTH HDSS from Dynamic Cohort Model 41
Figure 25 Plagiarism Declaration 46
Figure 26 Ethics Clearance Certificate 47

LIST OF TABLES

Table 1 Dissemination Channels 10
Table 2 Dissemination Channels 11
Table 3 List of Variables 23
Table 4 INDEPTH - iSHARE Data Dictionary 41

1

INTRODUCTION

A Health and Demographic Surveillance System (HDSS) consists of field operations for longitudinal follow-up of well-defined entities or primary subjects (individuals, households, and residential units) and all related demographic and health outcomes within a circumscribed geographic area [27]. HDSS sites belong to research institutions and primarily monitor (surveillance) demographic and health-related challenges in the surrounding communities for an extended period [8]. They play an important role in generating high quality of demographic and heath-related data to fill the gap created by the lack of comprehensive and reliable national civil registration and health information system [30].

Niakhar HDSS is one of the 52 HDSS sites within the International Network for the Demographic Evaluation of Populations and Their Health (INDEPTH) [27]. It was established in 1962 and had a second enumeration and extension of the study area from 8 to 30 villages in 1982. The HDSS collect longitudinal population-based, demographic (births, deaths, marriages, pregnancies, and migrations in population), health related and socio-economic data from the surveillance population. These data are then analysed by epidemiologists and population health scientists and the results are usually published in journals or presented at conferences.

However, in most HDSS sites, the dissemination of these research findings to local communities is almost non-existent. Some HDSS sites have community engagement departments or units. For instance, the Agincourt HDSS in northern rural South Africa has the Learning, Information Dissemination and Networking with the Community Division [17]. Such communities and their leaders normally establish good relations with the research institu-

tion. Such community engagement units usually do not communicate research findings to community members and their leaders [33].

For instance, the use of a web-based data visualisation platform associated with Geographic Information System (GIS) can serve as a visual link between the researchers and these community decision makers to better visualise public health concerns and interpret trends based on local knowledge [11]. Data visualisation or infographics is an emerging area that use visual means to represent research findings [25]. It constitutes a common way of communicating quantitative and qualitative information [1]. GIS visualisation has been used in public health as an analytic tool to organise, analyse and visualise patterns in health-related data. Research has shown that disseminating research findings using infographics will help community decision-makers to better understand the research activities of research institutions, support policy and help them to understand where service could be improved and the burden of disease [32].

It is imperative that once scientists conclude research studies, the results and findings are shared with different stakeholders. The The Centre for Disease Control and Prevention (CDC) defines this process as the process of communicating evaluation procedures or lessons learned to relevant audiences in a timely, unbiased, and consistent manner, with the ultimate goal of achieving full disclosure and impartial reporting [5]. Dissemination of research findings can enable health care service providers and community decision makers to make effective use and incorporate such findings in their strategies and policies to improve health outcomes of the population.

1.1 PROBLEM STATEMENT

At Niakhar HDSS, a large proportion of the surveillance dataset are not analysed due to the lack of research scientists. However, some proportion of the dataset collected from Demographic surveillance area is usually analysed for scientific publication purposes. The research

result can enable policy makers and community leaders to make appropriate decisions that will positively affect the local community. However, in many instances, their research findings and results only are published in academic and research journals or presented at conferences.

The main objective of HDSS sites is to provide reliable health and demographics information to the local communities and inform heath policy and planning at the local and national level [30]. This objective is normally not realised, and the community residents and decision makers are left out from the feedback loop and are not usually aware of such findings. This can partly be attributed to the fact that there are no existing dissemination strategies that allow local communities to make use of these findings in the HDSS and enhance community engagement. Community engagement is an increasingly important requirement for HDSS sites.

To contribute to policy and health-related decision making more widely, this project aimed to explore infographics methodologies and strategies that can be employed to enhance the utilisation of research findings at HDSS sites to inform decision-making. In this work, we developed a data visualisation platform that enables HDSS institutions to convey their research findings to policy makers and community leaders. This platform will aid the HDSS site to develop strategies and modalities for disseminating research information to communities via their leaders. We envisage that such a framework will enhance the various community engagement and outreach programs normally conducted by these institutions.

1.2 MOTIVATION

The HDSS produces a high volume of disparate data, and researchers are faced with the challenge of integrating datasets coming from diverse data source such as clinical data, mortality data, demographic data, socio-economic data, etc. as well as communicating research findings to the local community and inform decision makers. This minimises the potential of their research findings to influence public health practice and make appropriate decision

to improve heath outcomes of the local communities. This lack of appropriate channel of disseminating research results has motivated this research project. Unlike other HDSSs, Niakhar site does not have any form of community engagement activities and dissemination strategy. This research project provides a step towards bridging this gap by providing a platform for HDSS sites that aid researchers in disseminating research findings to local communities. The work also discusses other dissemination methods that can be coupled with infographics to reach the local community.

1.3 CONTRIBUTIONS

The major contribution of this project is the implementation of methods of communicating research findings of Niakhar HDSS site to local communities using infographics. The use of infografics will help in the improvement of information dissemination with the aim of raising the awareness of health conditions within the local community. The main results reported in this research project are:

1. An implementation of an infographic platform that enables community leaders to have access and make an informative decision on the research findings of the HDSS. This can ensure that the community understands the issues addressed by the research program.

2. Illustrations of how the infographic platform can be adapted at other HDSS sites and for specific cohort studies that utilise HDSS platform.

3. a discussion of other dissemination strategies that can be adapted by HDSS sites in reaching local communities.

1.4 OUTLINE OF RESEARCH REPORT

The remainder of this research report is organised as follows. In the next chapter (i.e., Chapter 2), we review the relevant works that have been conducted in research output dissemination. In Chapter 3, we give a description of the research methodology and technology used to develop the data visualisation platform. We then discussed the functionalities of the visualisation platform in Chapter 4. Also in Chapter 4 we discuss how the platform can be coupled with other communication media or strategies to reach local leader and policymakers. In Chapter 5, we gave some deployment considerations for other HDSS sites. We finally gave recommendations, conclusion and future directions in Chapter 6.

2

BACKGROUND AND RELATED WORK

In this chapter, we provide a review of related work in the dissemination of research information to local communities. First we explore the importance of the use of modern technology in disseminating research findings to the communities. We then discussed the impact of dissemination of research findings on public health policies and described the channels used in research findings. We concluded this chapter by discussing the challenges of information dissemination at INDEPTH Network sites.

2.1 USE MODERN TECHNOLOGY IN RESEARCH FINDINGS DISSEMINATION

With the eminent growth of the internet and web technologies, the web has naturally become the predominant source of information and communication as well as an underlying delivery mechanism for interactive information visualisation and dissemination [25]. The demographic and health related dataset collected in research institutions and the need to draw inferences and to analyse such large datasets have necessitated the need for appropriate channels of dissemination [3, 25]. These large datasets need to be represented in simplified formats for a better understanding and interpretation using visualisation applications[7]. Infographics have played important roles in epidemiology and public health, supporting dissemination of findings, data aggregation and analysis, and use of data for hypothesis testing and decision making [7]. Data visualisation can be used to reach a variety of stakeholders, including those with less technical expertise and low literacy and promote dialogue among researchers, policy-makers and local community leaders and also create mutual benefits [16].

It can be of great use in the research process cycle, enabling researchers and policy makers to explore datasets to identify patterns, associations, trends, etc, [32].

2.2 USE OF VISUALISATION IN COMMUNICATING RESEARCH FINDINGS

C. Gatto et al. [16] discussed the fact that visualisation can improve understanding and comprehension as well as has the potential to increase the use of research findings to benefit the community. The use of summary charts, graphs, maps, etc. among other, is useful to transform data into visual meaning and represent a key component of interpreting and bringing meaning to the dataset. For instance, John Snow, the father of epidemiology, graphed and mapped the cholera outbreak in London in 1854 and was able to demonstrate that its source were the local water sources [13] and also understood how the germs were spread [14]. The use of visual representation can allow the reader of the findings to know the change that has occurred over the time as opposed to the representation of data using tables of values which limit the reader in understanding trends as well as compare events over time [14].

Research studies have shown that with the large dataset emanating from heath research, web visualisation represents a better way of disseminating research findings in real time and a way of reaching not only community leaders but also those who are interested in these findings through the web [25]. The web is being considered as the most preferred option against the traditional communication media due to its flexibility and speed of information exchange and reaching a large target. These characteristics make the internet an obvious way for research findings dissemination [12].

2.3 IMPACT OF DISSEMINATION OF RESEARCH RESULT FOR PUBLIC HEALTH POLICIES

An important goal of research is to generate evidence to guide decision makers in improving policies and practices [22]. Despite this goal, there is still a large gap between the evidence generated from research and the application of those findings in community setting [4]. A research done by Brownson et al. [4] found that in an analysis of the content of 1210 articles from 12 public health journals, only 1% of the article were classified in the dissemination category. Actually, the gap of translating research-base into public health action is a challenge in research institutions. This gap can be filled through the use of effective dissemination channels that can reach the stakeholders to make effective use of the findings and also increase the interaction between researchers and policymakers.

The challenge of dissemination is to improve the accessibility of research findings. To find and use proven approaches, policymakers from government and development agencies usually refer to research findings to better understand problems of the communities, to inform decision making and to identify effective solutions [22].

2.4 OVERVIEW OF DISSEMINATION CHANNELS IN HEALTH RESEARCH

Various dissemination channels can be used to give feedback to the local community on the findings of the research conducted, and the consideration of the appropriate channel is essential in reaching the targeted audience [24]. Each of these dissemination tools has different strengths and weaknesses of reaching the audience and in certain scenarios, it is imperative to combine different dissemination strategies to complement one another to reach the local communities [24]. Mueller et al. [23], in their study found that dissemination of research findings through multiple channels increase the usefulness and stakeholder satisfaction and further the research findings sharing. However, there are limited studies conducted to em-

phasize how dissemination channels may affect the uptake of research findings [23]. Several modes exist for disseminating findings, and we only described the most widely used in this section for the purpose of comparison.

2.4.1 *Websites or Web Applications*

This channel of dissemination has become indispensabl, and it allows easy access to the information about the research findings and provide easy updates [12]. Institutional websites offer more flexibility and speed of communicating findings to the audience because the visitor can easily share the link with other visitors to increase the visibility of results and the link could be emailed to stakeholders. The web represents a delivery mechanism that offers a large number of advantages but also has limitation as indicated in Table 3 [25]. Many HDSS sites including Niakhar HDSS are located in rural area with a limited internet coverage of the area and poor connectivity. This represents a limitation of this channel for reaching the population in which the research was conducted. Also, this channel requires a high level of literacy to comprehend the content. For instance, many organisation like United Nations, Demographic and Health Survey (DHS) Program, Word Bank among other disseminate research results through a data visualisation web application. These web applications allow researchers around the world and local communities to find relevant information in these research datasets. Also, the INDEPTH Network through its iShare project uses a web platform called IndepthStat to share the indicators of the affiliated HDSS to the global and local community [18]. These platforms are not only limited to sharing indicators but also share well-documented dataset for secondary data analysis. On web platforms, dissemination can take the form of interactives visual charts and graphs representing the visual meaning of the data.

Dissemination Channels	Advantages	Disadvantage
Internet & E-mail	- Flexibility - Electronic process can reach a large number of people quickly - The information being conveyed can be made and propagate quickly	- Cost - Reliable access to the information and communication technology - Reach audience with computer literacy
Journal articles	Wide impact on scientific network	- Limited audience - Oriented audience - Not jargon-free information - Not easy to access
Research report		Limited audience
Community meeting	- can be structured achieve some outcomes - time limited - can invite a particular targeted group	- may not attract participants who are representative of the community - can frustrate participants as discussion is often artificially constrained to some limited priorities - is not suitable for topics around which there is significant controversy or negative opinion
Mass media	- Reach a large number of people - Raise public awareness and understanding of open access to research data - prompt policy-makers to support the consortium's recommendations.	- Expensive for private research institution - Can only convey a limited amount of information - Promote One-way communication

Table 1: Dissemination Channels

2.4.2 *Scientific Journal Publications*

Scientific publication in peer-reviewed journals is a peak achievement for many researchers and represents a visible contribution to the research field in the scientific community [24]. Peer-reviewed repositories for the scientific community show that those studies contain in them are credible. [29]. The traditional way of communicating academic research findings is through journals or articles. However, this channel does not meet most users needs in term of easy access and understanding [29]. For instance, articles are unlikely to reach the

non-technical targeted audience and local communities[24]. Although researchers use visual representation of the articles to explain their research findings, the use of infographics in scientific publications is limited due to the required format, and these visualisations are sometimes too detailed and less simple to allow the layperson to have a better understanding.

2.4.3 *Mass Media*

Mass communication includes electronic and print media. Electronic media includes radio, television, teletext, videotext, and satellite telecommunications. Print media encompasses books, newspapers, magazines, newsletters, and comics.

The effective area of electronic mass media is television and radio. Mass media are channels that can be used in research findings dissemination. They are very useful in reaching a wide audience effectively [2] but represent the most expensive channel[1]. Information is shared through talk shows and can be utilised for behaviour change in local communities. For instance, compared to the last two decades, the Senegalese media landscape has now had multiple channels and users have now the choice of switching from one channel to another. A study revealed that the radio and television had become the most frequented media in Senegal with a national average rate of 75% (60% in the rural area against 97% in urban area)[2]. Also, according to the National Statistics and Demographics agency Agence National de la Statistique et de la Demographie (ANSD), 8 out of 10 households had a TV in 2011. The rural electrification plan has made the television more accessible to communities and can be used to disseminate research findings. However, the internet penetration remains low with an overall rate of 61% in 2015[3].

To disseminate research finding using TV, the use of image representation on advertisement or communication regarding the finding can be a great way for the audience to understand and be aware of the research data. Cartoon videos in the local language can

1 http://codi.tamucc.edu/archives/pubs/articles/.edwards/.chap5.htm
2 http://www.cnra.sn/do/radio-et-television-faits-et-chiffres/
3 http://www.osiris.sn/Internet.html

also be used as an infographic representation to sensitize the community about the finding and how to improve their health condition.

2.4.4 *Community Meetings*

In rural communities, social events are part of the lives of the people. These events or meeting can be exploited to disseminates research findings to the community [20]. For instance, Agincourt HDSS has created a Learning, Information Dissemination and Networking with the Community Office (LINC) to build a community relationship and disseminate information [17]. In addition, Agincourt HDSS organise routine meetings to inform the village leaders and communities of research project update and discuss the research findings. The limitations of this channel are listed in Table 1.

2.4.5 *Summary of the Literature*

Information sharing between stockholders has become a crucial part of health research. For a better communication with the communities and partners, the researchers need sufficient tools aligned with modern technology to reach the communities. These modern technologies include web application that can translate research data into a visual meaning and makes information easy to access and update. However, most HDSS lack such tool to communicate or disseminate research result to their study participants. To fill this gap, we aimed to develop an infographics platform and help the HDSS to disseminate their findings and to share data and knowledge with policymakers and local communities.

3

AN INFOGRAPHICS PLATFORM FOR RESEARCH INFORMATION FEEDBACK

This chapter explains the techniques and technologies used in developing the infographics platform for the research information feedback. This developmental research project was conducted at the Niakhar HDSS site. We will first describe the technologies and software that were used to develop the infographic platform then describe the study setting and its implication within the institution. We will finish by demonstrating the use with cases of the module by using sample dataset from the Niakhar HDSS site.

3.1 OVERVIEW OF TECHNOLOGIES USED

The data visualisation platform was implemented using the statistical software R, coupled with Shiny web application framework. SQL server was also used as a Database Management System (DBMS) for storing the staging data collated from the HDSS data store. The link between the R and Structured Query Language (SQL) server was done using the R plugin for Open Database Connectivity (RODBC) package that allow R to access objects of the database through the Open Database Connectivity (ODBC) connection.

3.1.1 *Statistical Software, R*

R is a programming language and software environment for statistical computing and graphics supported by the R Foundation for the Statistical Computing[1]. R is an open-source soft-

1 https://en.wikipedia.org/wiki/R_(programming_language)

ware available under GNU General Public License (GPL) and its capabilities are extended through user created-packages developed and published by the R community. We chose R primarily because it is a free software and has advanced data management and statistical capabilities to manipulate large dataset and produce analytic results.

3.1.2 *R Shiny*

Shiny is an interactive web application framework that combines computational power of R with the interactivity of the modern web technologies. Shiny Application Programming Interface (API) implement the required Hyper-Text Marckup Language (HTML), CSS and JavaScript needed to build an interaction web application without any knowledge of HTML. Shiny provide a simple way to display the infographics in a web browser and provide many input function for a variety of user interactions.

3.1.3 *Microsoft SQL Server*

Microsoft SQL Server is a relational Database Management System (DBMS) from Microsoft for easy development and management of databases. Niakhar HDSS use a data entry application connected to an MS Access Database. Once the field surveillance round finalised, the data are then exported into an MS SQL Server Database for the data management process. MS SQL Server is also used to export dataset for analysis and as staging platform in our infographics platform.

3.2 APPLICATION ARCHITECTURE

The developed platform used a three-tier architecture that includes a database server that hosts the dataset for the routine yearly HDSS surveillance. In order to process the stag-

ing dataset, an R engine that represents a bridge between the Shiny web server and the database was used to extract intermediate dataset used for the analytics in the platform (see Figure 1 below). The principle is that the R engine extracts the data required for the data visualisation from the database, performs the necessary data management and data transformations, performs the analytics required to be visualised and passes the outputs to the shiny web server. The last part of the architecture is a Shiny web server which is used as the presentation layer in the platform. The role of the web server is to receive requests from the end user, transfer them to the R engine for processing through the web browser and bring back the requested results.



Figure 1: Application Architecture

3.3 NIAKHAR HDSS

This work was conducted using the dataset from the Niakhar HDSS, one of the 53 HDSS sites affiliated with the INDEPTH Network. Niakhar HDSS located in the centre of Senegal, at 135 Km southeast of Dakar, the Senegalese capital, was established 1962 [10]. The HDSS collect demographic, health related and socio-economic data in over 3300 households and covers a total area of 30 villages, with a population of about 45000 residents in 2014. The majority of the population belong to the Serere ethnic group and a minority of the Wolof and

Poular ethnic group. To start this research project, we conducted informal investigations into the existing circumstances of community engagement with the research institution, followed by the data management processes. We finally planned and developed the data visualisation platform to improve the process of disseminating research results and finished by discussing other methods of dissemination out to the communities.

3.4 IMPLEMENTATION AND PRE-PROCESSING CONSIDERATIONS

We utilise surveillance data collected from 1983 to 2015 to build the visualisation platform. The data was cleaned according to the following specified criteria. The first criterion was to select a period in which the indicator had to be calculated. We decided to use the period between 2000 to 2014 resulting in up to 14 years of follow-up surveillance because the data collection was computerised in 2005 and that period give more accurate data to analyse. The second was to remove all the negative ages after computation of the age and then remove all the internal visit that had occurred in the HDSS area. The last criterion was to remove all the individuals that did not fall under the censoring dates. The indicators were then produced using the cleaned dataset. All the indicators were computed either by sex and/or by year and can be compared across the different sex.

3.5 AN OVERVIEW OF THE INFOGRAPHIC PLATFORM

This is a visualisation platform that will aid the dissemination and sharing of relevant information from the research findings at Niakhar HDSS. The dissemination tool was implemented based on longitudinal demographic data collected routinely in two or three round per year depending on the available of funding. There are five main utilities/modules of analysis within the platform and each module represents an analytic entity. Each module implements several indicators that give information about the trend in the subject matter.

In this section, we will discuss these utilities/modules of analysis and indicators and also how they were obtained and utilised.

3.5.1 *Module 1: Mortality Visualisation Utility*

The mortality utility was developed to display all the mortality indicators and show the mortality trends using the following demographic calculation. The computation of these indicators was done according to the following INDEPTH definitions [19]:

- **Crude Death Rate (CDR)**

  The CDR represents the total number of deaths occurring among the population divided by the person-years of all ages. This indicator is expressed in deaths per 1000 person-years.

- **Age-specific mortality rates (ASMR)**

  ASMR is the number of deaths in a specific age group divided by the person-years lived in that age-group, expressed in deaths per 1000 person years

- **Age Standardised Crude Death Rate (ASCDR)**

  ASCDR is a weighted average of the age-specific mortality rates, expressed in deaths per 1000 person-years. The weights are the proportions of person-years in the corresponding age over the entire period of HDSS observation.

- **Neonatal Mortality Rate per** 1000 **person-years (NMR)**

  NMR represent the number of deaths in the first 28 completed days of life divided by the number of person-years lived in the HDSS within the first 28 days, expressed in deaths per 1000 person-years.

- **Neonatal Mortality Rate per** 1000 **Live Births (NMRLB)**

NMRLB is the standard rate obtained by dividing the number of deaths in the first 28 completed days of life by the number of live births in the HDSS. It is expressed in deaths per 1000 live births.

- **Infant Mortality Rate**

  Infant mortality rate is the life-table (Kaplan-Meier) estimate using exact dates of death and censoring. It is expressed in deaths per 1000 live births and interpreted as the probability to die before 1st birthday.

- **Child Mortality Rate**

  Child mortality rate is the life-table (Kaplan-Meier) estimate using exact dates of death and censoring. It is expressed in deaths per 1000 person-years surviving their 1st birthday and interpreted as the probability to die between 1st birthday and 5th birthday.

- **Under-Five Mortality Rate**

  Under-five mortality rate is the life-table (Kaplan-Meier) estimate using exact dates of death and censoring. It is expressed in deaths per 1000 live births and interpreted as the probability to die before 5th birthday

- **Adult Mortality Rate**

  The adult mortality rate is the life-table (Kaplan-Meier) estimate using exact dates of death and censoring. It is expressed in deaths per 1000 people surviving their 15th birthday and interpreted as the probability to die between 15th birthday and 60th birthday.

- **Life Expectancy at Birth**

  Life expectancy is the life-table (Kaplan-Meier) estimate using exact dates of death and censoring. It is expressed as an average number of years that a hypothetical cohort would have lived by experiencing the health conditions of a particular calendar year over its lifetime.

3.5.2 *Module 2: Fertility Visualisation Utility*

The fertility utility/module was added to display all the indicators related to fertility. It was computed using the number of birth to the HDSS Female residents in the reproductive age groups i.e. aged between 15 to 49 years. The indicators include the following:

- **Crude Birth Rate (CBR)**

  CBR is the number of births divided by the person-years of all ages, expressed in births per 1000 person-years.

- **Age-Specific Fertility Rates (ASFR)**

  ASFR is the number of births in a specific age group divided by the person-years lived in that age-group, expressed in births per 1000 people. The seven age-groups are $15-19$, $20-24$, $25-29$, $30-34$, $35-39$, $40-44$, $45-49$.

- **Total Fertility Rate (TFR)**

  TFR is the sum of ASFR multiply by 5. It is interpreted as the mean number of births that a hypothetical cohort of women would have had by experiencing the fertility conditions of a particular calendar year over its lifetime.

- **Mean Age at Childbearing (MAC)**

  MAC is computed as the sum of ASFR weighted by the distribution of females at the mid-point of each age group as per the life-table estimate, divided by the sum of the age-specific rates. It is interpreted as the mean age of mothers at the birth of their children if women were subjected throughout their lives to the age-specific fertility rates observed in a given year.

- **Sex Ratio at Birth (SRB)**

  SRB is the number of male births divided by the number of female births, expressed in male births per 100 female births.

3.5.3 *Module 3: Migration Visualisation Utility*

Migration is defined as the change of residences of within or outside the surveillance boundaries of the HDSS site we have two type of migration: The In-Migration defined as the change of residence to the surveillance boundaries and Out-Migration that represent the change of residence from the boundaries of residence to another. These are the indicators used in this utility:

- **Crude In-Migration Rate (CIMR)**

  CIMR is the number of in-migrations divided by the person-years of all ages, expressed in in-migrations per 1000 person-years.

- **Crude Out-Migration Rate (COMR)**

  COMR is the number of out-migrations divided by the person-years of all ages, expressed in out-migrations per 1000 person-years.

- **Age-Specific In-Migration Rates (ASIMR)**

  ASIMR is the number of in-migrations in a specific age group divided by the person-years lived in that age-group, expressed in in-migrations per 1000 person-years.

- **Age-Specific Out-Migration Rates (ASOMR)**

  ASOMR is the number of out-migrations in a specific age group divided by the person-years lived in that age-group, expressed in out-migrations per 1000 person-years.

3.5.4 *Module 4: Population Visualisation Utility*

The population age pyramid presents the distribution of the population across the age groups per five years and by sex, also the under one population was considered. It is computed using

the number of individual residents in the defined age group. The number of individuals was calculated taking into account the period of observation only since we had a dynamic cohort and the time of observation represent the Person-Years.

- **Sex Ratio**

  Sex ratio is the ratio of males to females. It is computed as the number of HDSS male residents divided by the number of HDSS female residents, expressed in number of males per 100 females.

- **Proportion** < 15

  It is the number person-years lived before the 15th birthday divided by the total number of person-years.

- **Proportion** $15 - 59$

  It is the number person-years lived between the 15th and 60th birthdays divided by the total number of person-years.

- **Proportion** $60+$

  It is the number person-years lived after the 60th birthday divided by the total number of person-years

3.5.5 *Module 5: Cohort Studies Visualisation Utility*

The cohort study utility represents a module that was developed as an infographic platform for the visualisation of data from any cohort studies annexed within the HDSS. The purpose of this module was to organised clean dataset and link it to the core HDSS data for visual representation.

3.6 VARIABLES WITHIN HDSS DATASET

The variable that were in the HDSS dataset are presented in the following table:

Variables	Description
Individual Id	Unique Identifier of the followed Induvial
Date of Birth	Date of Birth of the Individual
Sex	Gender of the Individual with option 1: Male and 2: Female
Start Date	Represent the date in which the individual joined the demographic surveillance
Start Event	Event in which the individual enters in the study. Different start event was considered as start event: o Enumeration; o Birth o In-Migration
End Date	Date in which the individual left the demographic surveillance.
End Event:	Date in which the individual joined the demographic surveillance. The different end event was: o Death o Out-migration o End of Observation
Location	Village of the individual a point in time
Mother Id	Unique identifier of the mother of the individual. Also an Individual Id

Table 2: List of Variables

3.7 ETHICAL CONSIDERATION

Niakhar HDSS is a research platform that hosts several studies in many disciplines and each of these studies requires an ethical approval from the ethics committee based at the Department of Health before their starting up. For instance, this data visualisation application

uses the data collected from these studies as input and for that purpose, a letter approving the uses of these data was obtained. However, all the database servers were hosted by the IT infrastructure of the research organisation in a secure and lockable place accessible by authorised personnel. Also, accessing these servers requires authentication and strict privileges were set up to prevent information divulgation by an unauthorised person. There was no personal information shared but only aggregated dataset compiled by village. In addition to HDSS ethical requirement, an ethical approval was obtained from the University of Witwatersrand Human Research Ethics Committee before this project commenced.

4

PRESENTATION LAYER OF THE INFOGRAPHICS PLATFORM

This chapter will first presents the features of the presentation layer of the platform. This is followed by discussions of the indicators resulting from the developed visualisation platform. The description will be done through case scenarios with descriptive analysis. We utilise the mortality module for these case scenarios.

4.1 PRESENTATION OF THE SHINY APP

The interactive interface of the platform contains a dashboard illustrating the health and demographic indicators for Niakhar HDSS using longitudinal dataset collected from 2000 to 2014. The features of the dashboard include a sidebar menu and a display page. The sidebar menu was developed to navigate through the different modules/utilities of the application. This menu has two different parts. The first part is the analysis modules. The different dashboards menu can be accessed using the five modules which are the Mortality, Fertility, Migration, Population and the Upload dataset to display the relevant indicators. The second part includes parameters that allow the users to display indicators by year and the gender and also compare these indicators across these defined parameters as presented in Figure 2 below. The display page represents the main component of this dashboard and it shows the information that is being visualised. To allow future data we have integrated a module that allows the user or researcher to sync newly collected data from the HDSS and display some descriptive statistics. As future work, we intend to implement predictive models using machine learning techniques to draw insight from the newly added dataset.



Figure 2: Mortality rate trend

4.2 DESCRIPTIVE VISUALISATIONS

4.2.1 *Use Case: Mortality Data Visualisation*

Measuring mortality indicators, especially child mortality in a rural area, was one the main objectives of establishing Niakhar as a Democratic surveillance system in 1962 [6]. Sub-Saharan Africa has the highest risk of mortality especially in child mortality [21]. However, this mortality has been seen declining over the years. For instance, the African Leadership for Child Mortality reported that the under-five mortality rate had declined from 163‰ in 1990 to 100‰ in 2011. Despite this considerable decrease, this was insufficient to reach the Millennium Development Goal of 4‰ in 2015. The under-five mortality rate in Senegal is far from the African global rate with 55‰ in 2013 [24]. Figure 4 shows the CDR across the years between 2000 and 2014. A high decrease of the mortality was noted in both male and female and this decrease represents half of the initial value 16 death for 1000 men versus 13 Death for 1000 women in 2000, and respectively 7‰ and 5‰ in 2014 (see Figure 4). This decrease in the mortality can be attributed to the massive establishment of schools

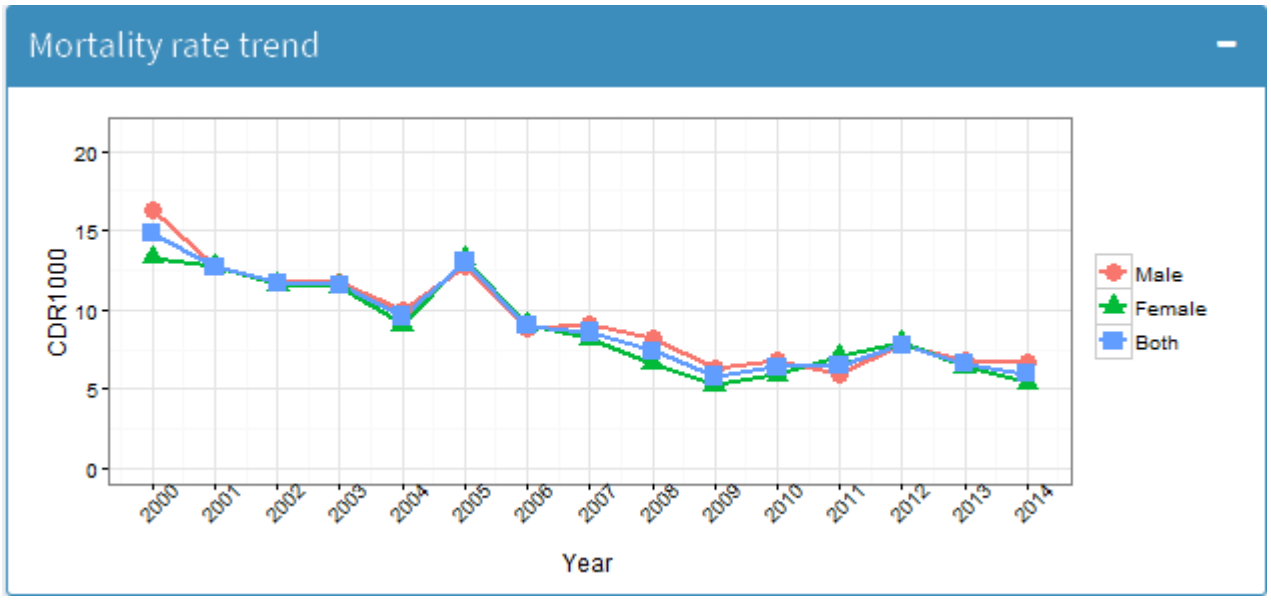

Figure 3: Crude Mortality Rate

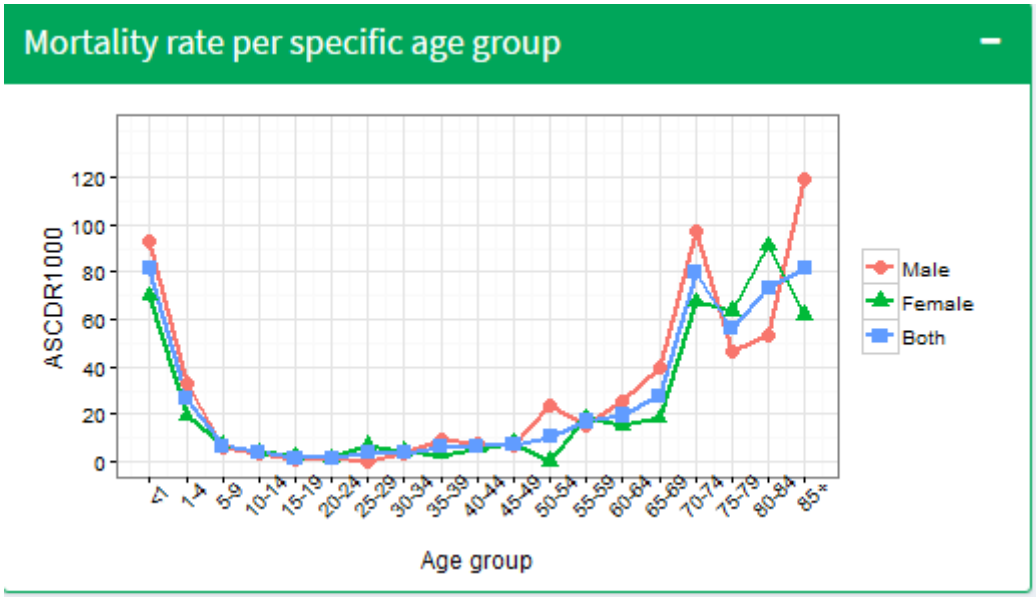

Figure 4: Age-specific Mortality 2000

in rural area in the years 2000 that increased schooling and the level of literacy and affect the behaviour change of the population [10]. Although a high decrease of the mortality was noted in the area between 2000 and 2014, a pick was noted in 2005 with a mortality rate of 13‰ distributed across the under 1-year-old, children and adults. The pick was also noted in 2012 but affect mostly the under 1-year-old with a rate two times more than the previous year. A high mortality rate was noted for the under 1-year age group when comparing

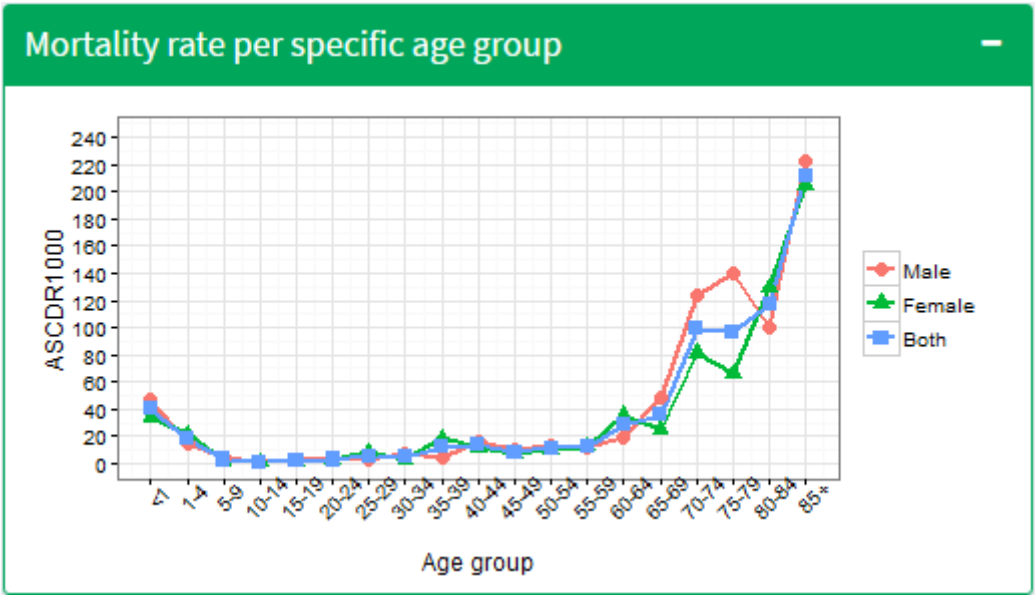

Figure 5: Age-specific Mortality 2005

the mortality rate across different age groups, with an overall of 80‰ in 2000. This rate decreased especially for the under-five age groups if we compare it to the mortality in the same age group from 2000 to 2014 (see Figures 5, 6 and 7).

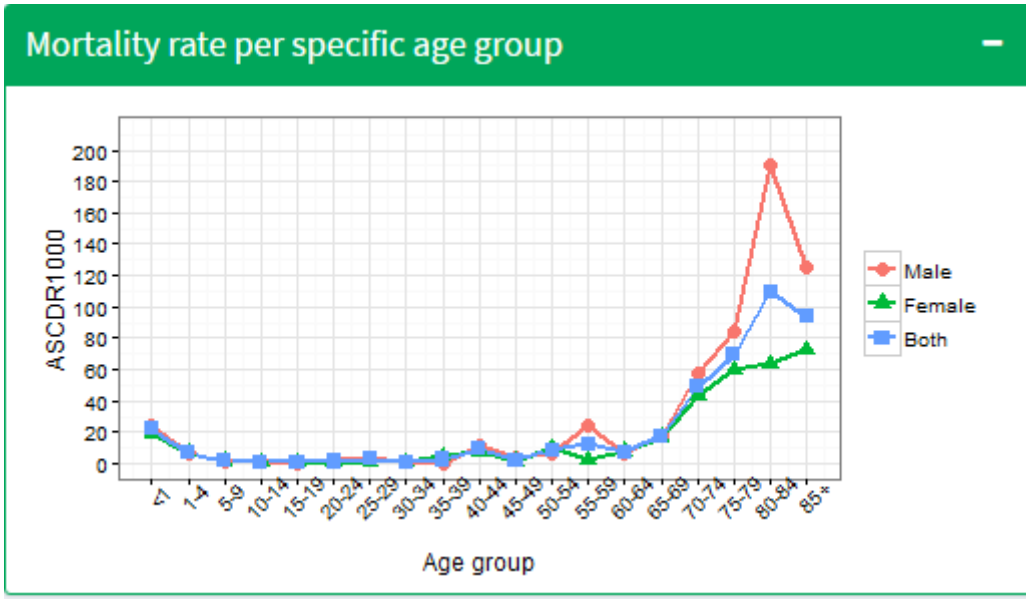

Figure 6: Age-specific Mortality 2010

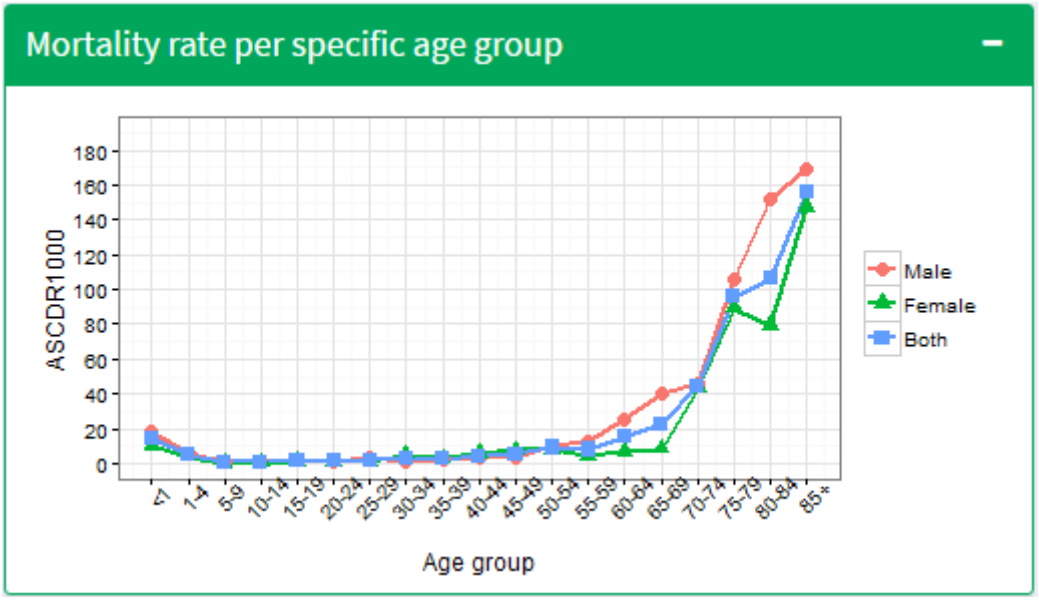

Figure 7: Age-specific Mortality 2014

4.2.2 *Use Case: Fertility Visualisation*

Fertility in sub-Saharan Africa remains the world highest although there is a strong evidence of fertility transitions [31]. For instance, according to the United Nations population division, the cumulative number of children per women was 7.0 in 1960 compared to 5.0 in 2014. In Senegal, fertility has followed the same trend as in Sub-Saharan Africa which was slowly declining over the past years with a higher fertility among women in rural areas compared to women in urban areas [15]. The transition in rural areas was due to the literacy and also the socio-economic transformation based on the economic and agricultural crises [28].

However, in Niakhar HDSS, the total fertility remains high despite the fertility transition and the decrease of the fertility across the years of follow-up surveillance. This fertility was at 7.3 children per women in 2000 and has decreased to 6.3 children per women in 2014 (see Figure 9). We also noted an important decrease of 5.6 children per woman in 2011 and then an increase of the fertility in 2012 going up to 6.7 children per women. A decrease of the Crude Fertility Rate across the observation period showed a rate of 43 children per 1000 women in 2000 compared to 38 children per 1000 women in 2014 (see Figure 7 8). According to Garenne et al., the fertility transition in Niakhar began after the peak in the year 80s

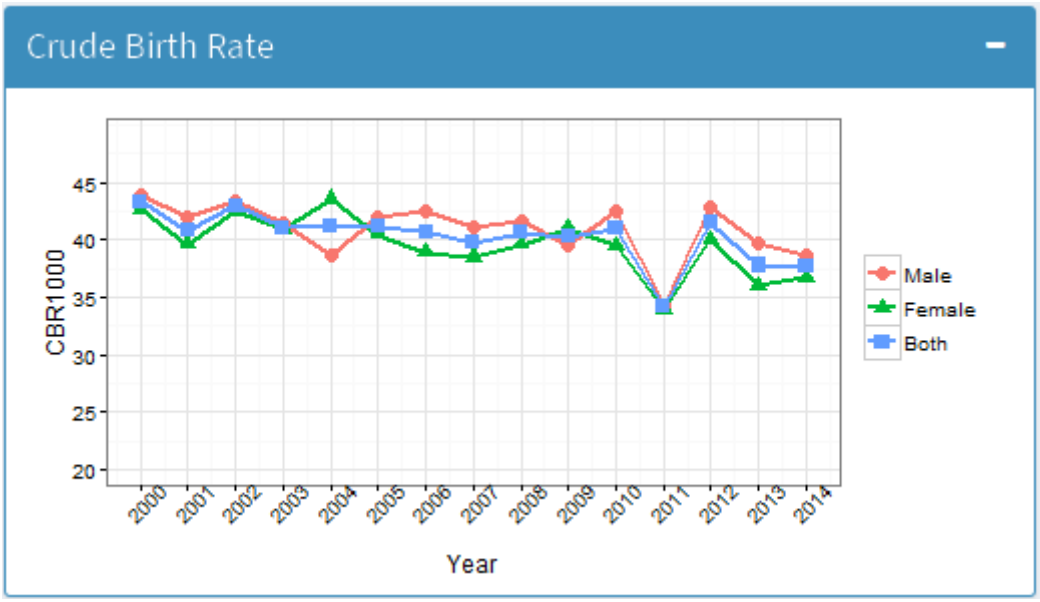

Figure 8: Crude Birth Rate Trend - Niakhar, 2000-2014

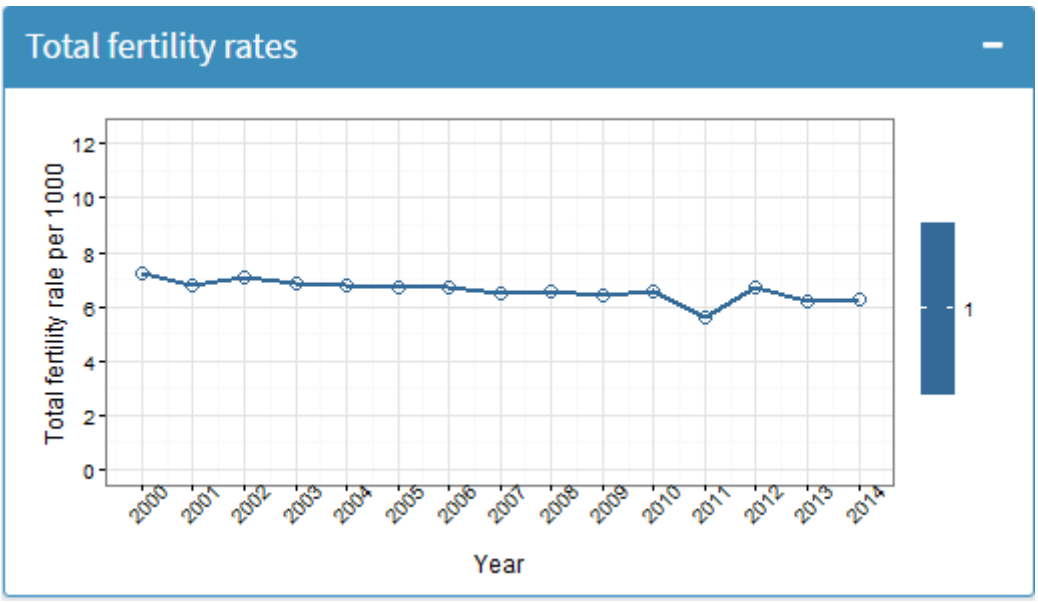

Figure 9: Total Fertility Rate Trend - Niakhar, 2000-2014

with eight children per women and since then this decrease has become regular over the periods [15]. The fertility visualisation across different age groups showed that the fertility transition became more apparent in some age groups than others. For instance, we noted that the fertility rate has considerably decreased for the women aged under 30 years old compared to the women over 30 years old as shown in Figure 10. Garenne et al. [15] assessed

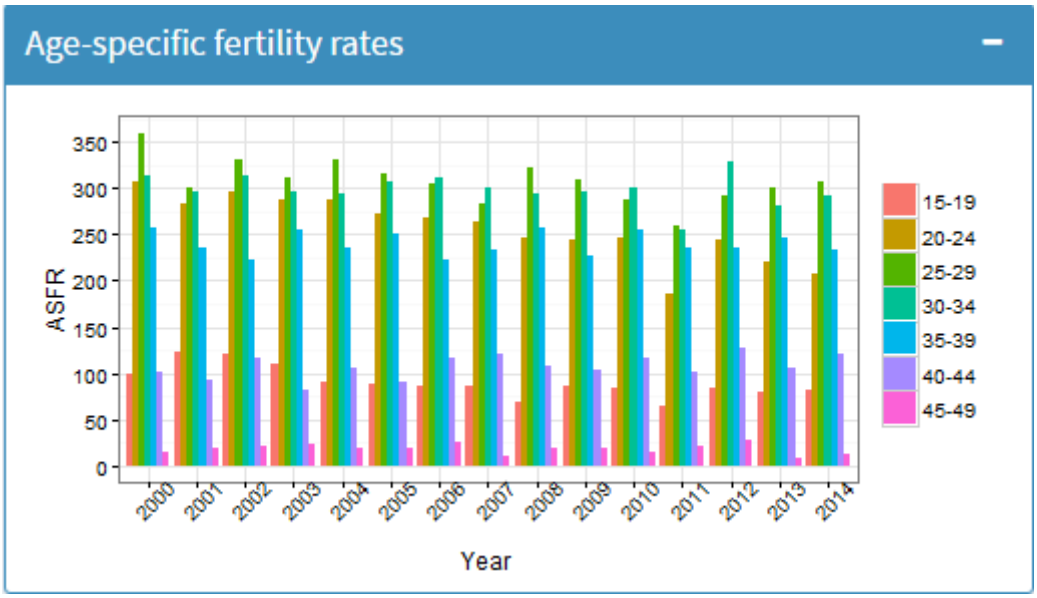

Figure 10: Fertility Rate Per Age Group -Niakhar, 2000-2014

that this is caused by the reduction of marriage for the young women and probably the use of contraceptive methods.

4.2.3 *Use Case: Immigration Visualisation*

Migration to urban area is strictly related to the economic development and urbanisation of the city and this has resulted in the concentration of the economic activities and also the economic growth of those cities [26]. In rural areas where poverty is predominant, the population find migration to urban area as a definitive answer to reduce poverty [10]. However, the migration from rural to urban area weaken the rural economic development by depriving their human and material capital resource, therefore reduce the agricultural productivity [9]. De Hass (2010) added that migration is also seen as an aggravating problem of underdevelopment with the labour force exploited for the benefit of the industrialised regions and countries. Quite on the contrary, migrants contribute to the economic growth of the urban areas via informal activity like commercial and weak remunerated employment but this contribution is not seen as important [26]. In Niakhar HDSS, Valerie et al. [9] noted that migration is strictly related to temporary and informal labour works. These temporary

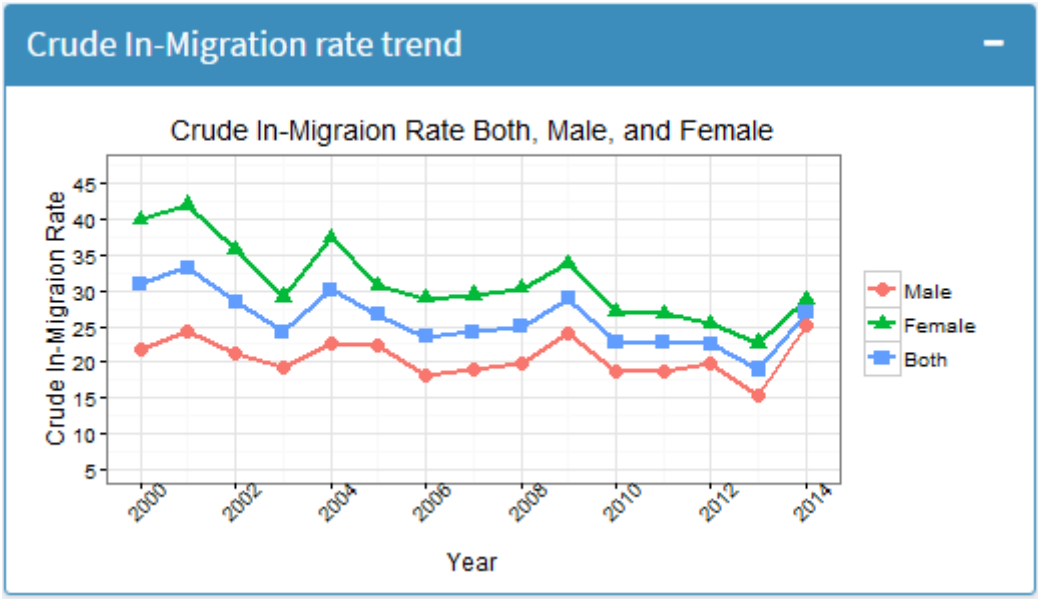

Figure 11: Crude In-Migration Rate - Niakhar, 2000-2014

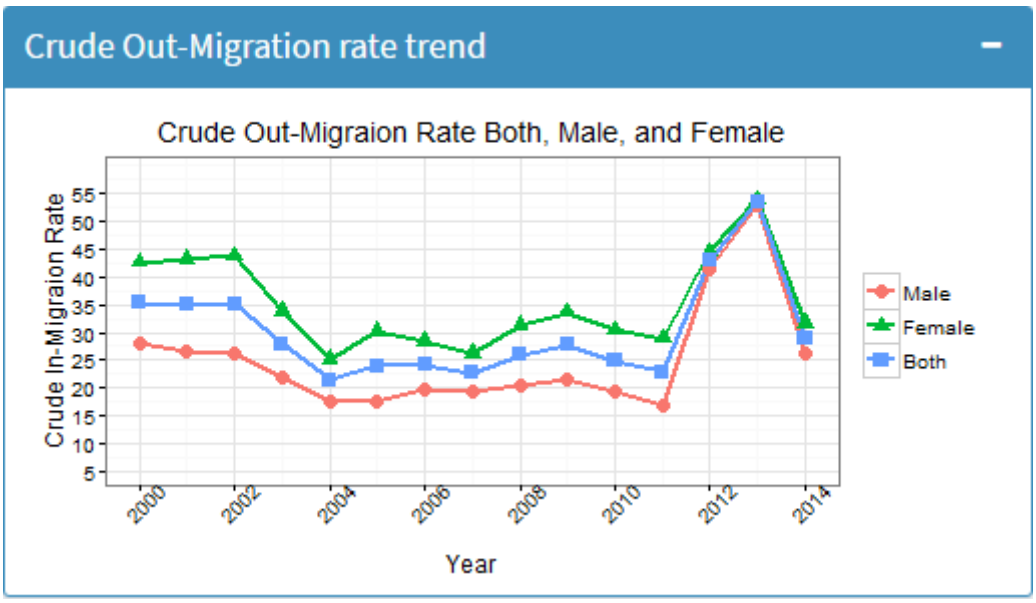

Figure 12: Crude Out-Migration Rate - Niakhar, 2000-2014

migrations started in the 1960s and these happen just after the farming season to increase the household income which was based only on the farming activities [10].

From 2000 to 2014 we noted a decrease in the migration in general in Niakhar going from an overall in-migration and out-migration rates of 32‰ and 35‰ respectively in 2014 against 20‰ for the in-migration in 2013 and 23‰ for out-migration in 2011 (see Figures 12 and 13). Figure 12 also shows an important migration rate in the female population com-

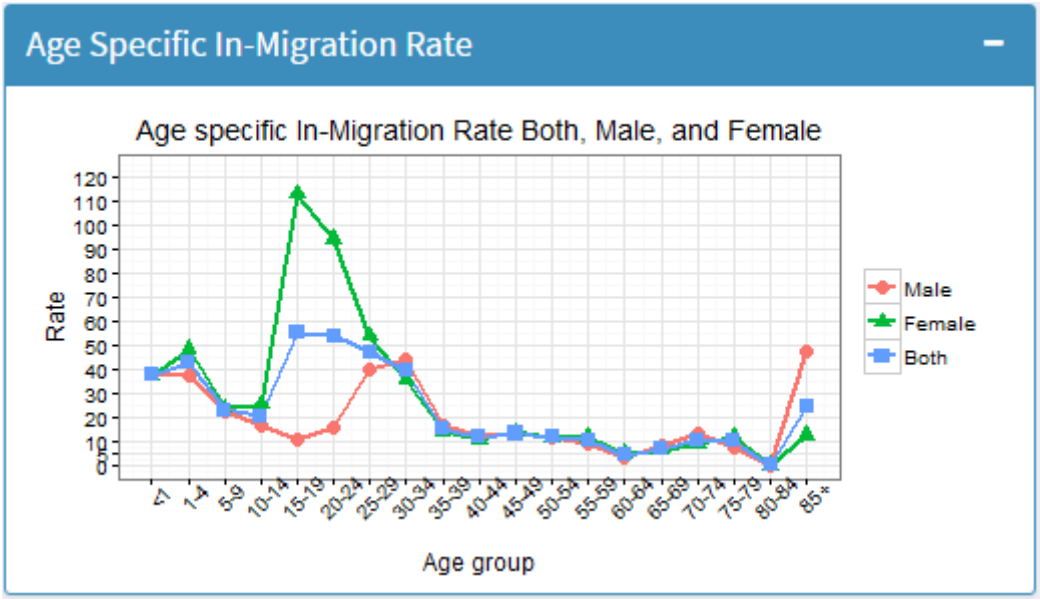

Figure 13: In-Migration Rate by Age Group - Niakhar, 2000

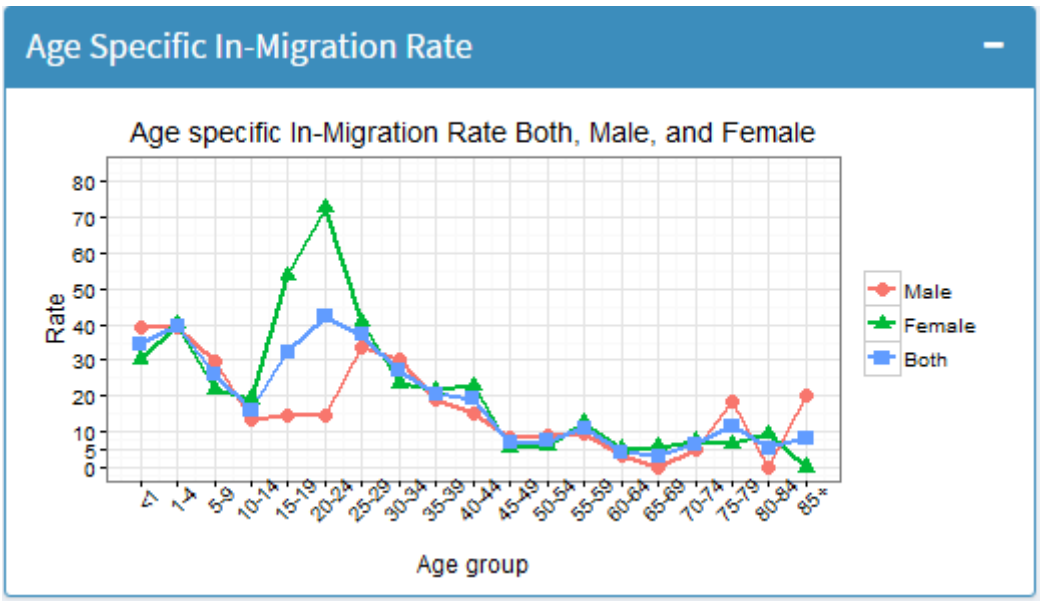

Figure 14: In-Migration Rate by Age Group - Niakhar, 2005

pared to the males across different periods of years. This migration is mostly related family reasons such as marriage, divorce etc. In Senegal and most African countries, the women join the husband after marriage and in case of divorce or death of the husband the women go back to their origin family [10]. Comparing the migration across different age groups showed a decrease among the young (see Figures 14, 15, 16, 17) and (see Figures 18, 19, 20, 21). For instance, in the year 2000, this decrease can be attributed in part to an

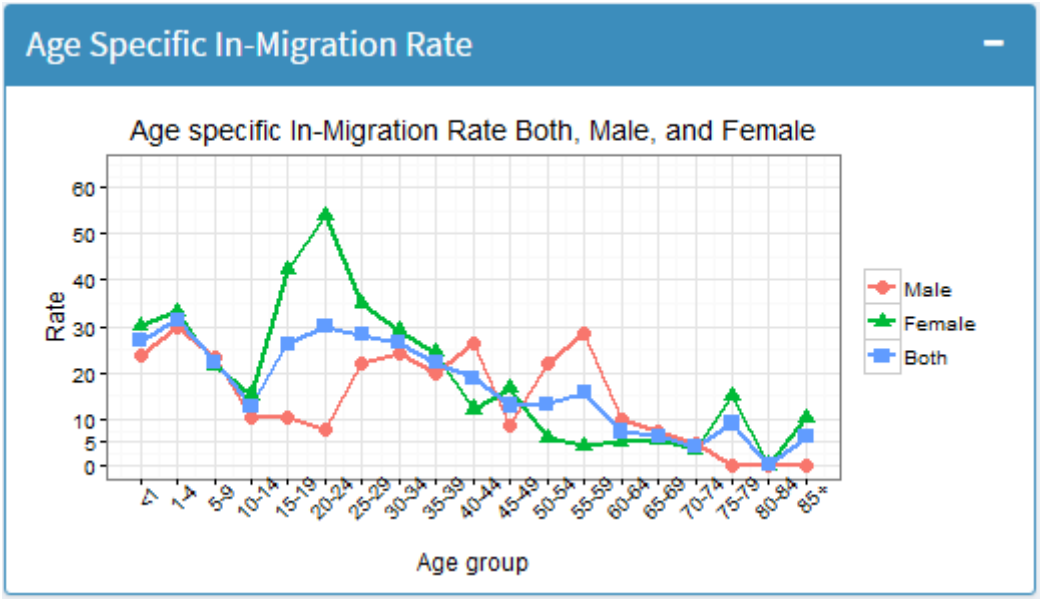

Figure 15: In-Migration Rate by Age Group - Niakhar, 2010

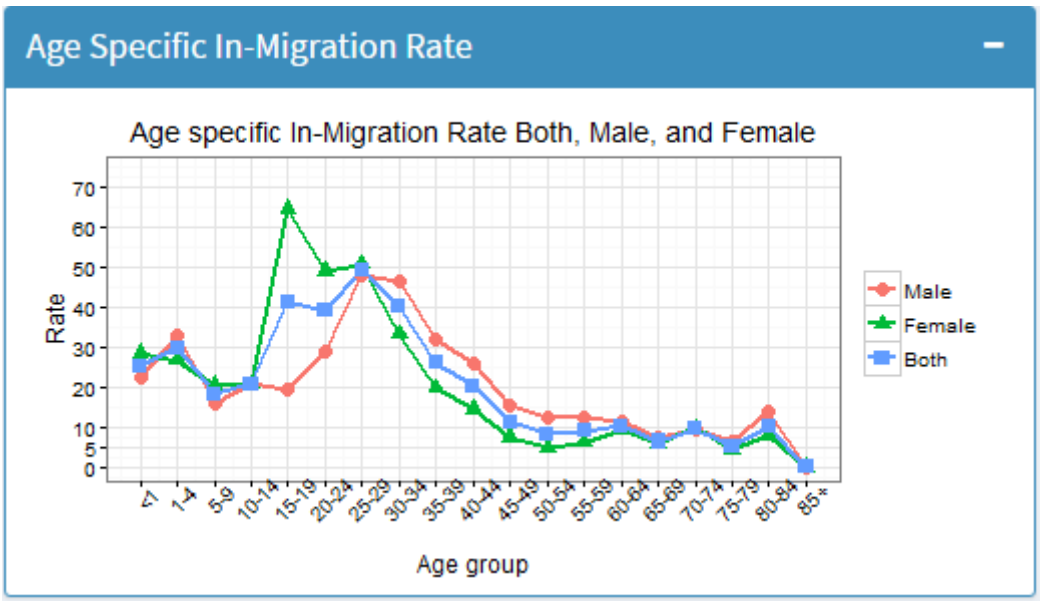

Figure 16: In-Migration Rate by Age Group - Niakhar, 2014

important effort of the government to establish schools in rural areas. This has contributed to the decline of the migration in the young with the schooling.

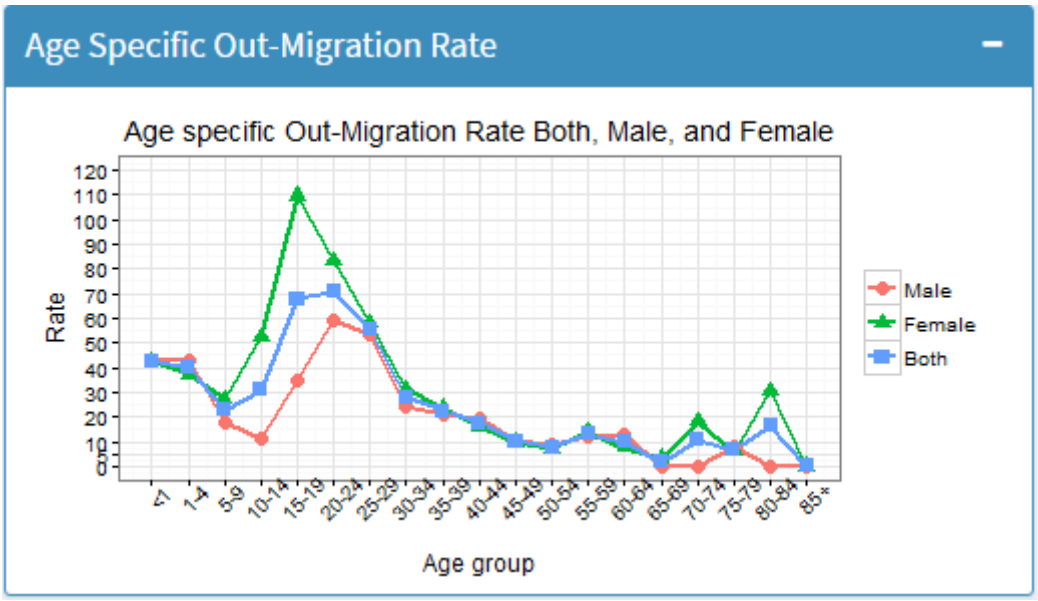

Figure 17: Out-Migration Rate by Age Group - Niakhar, 2000

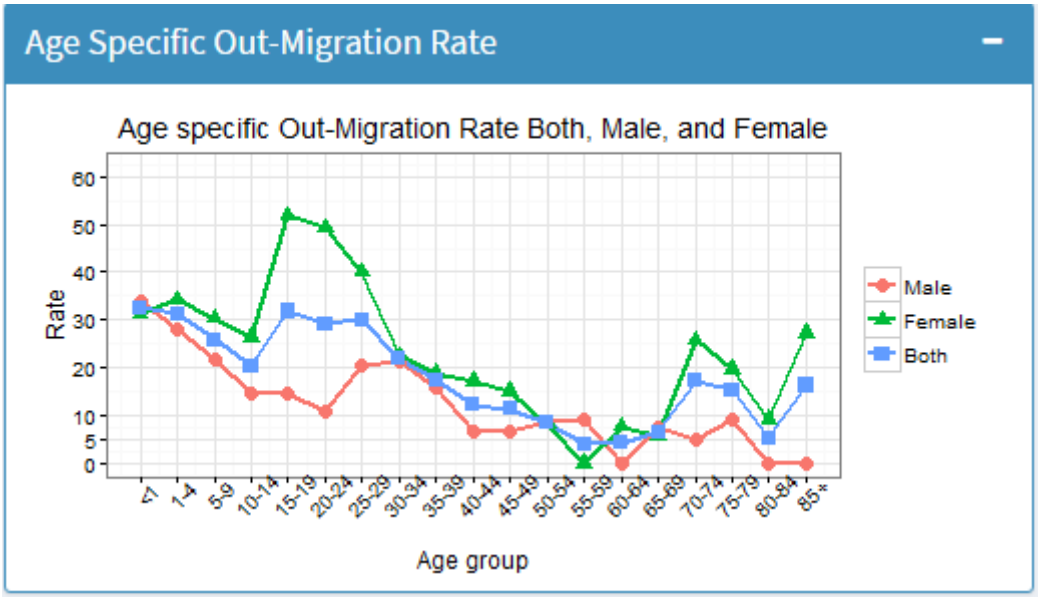

Figure 18: Out-Migration Rate by Age Group - Niakhar, 2005

4.3 PLUGIN FOR COHORT STUDIES

The main focus of HDSS is to conduct demographics surveillance of a particular group of populations in a specific area. However, other longitudinal studies and clinical trials are conducted within the HDSS and those studies used the demographic surveillance data as the baseline. These studies or research are not part of the routine surveillance processes of

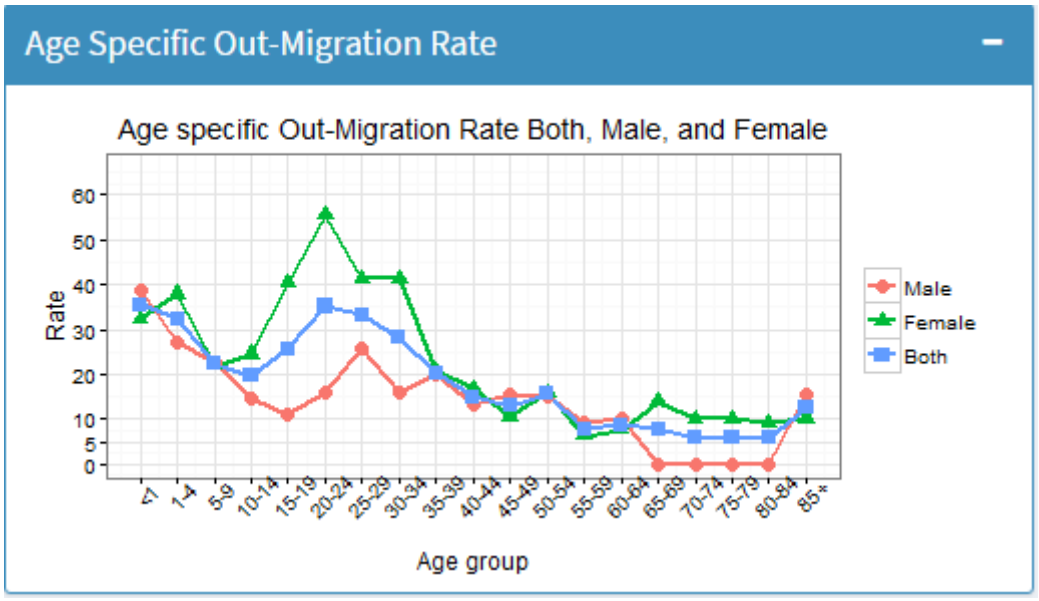

Figure 19: Out-Migration Rate by Age Group - Niakhar, 2010

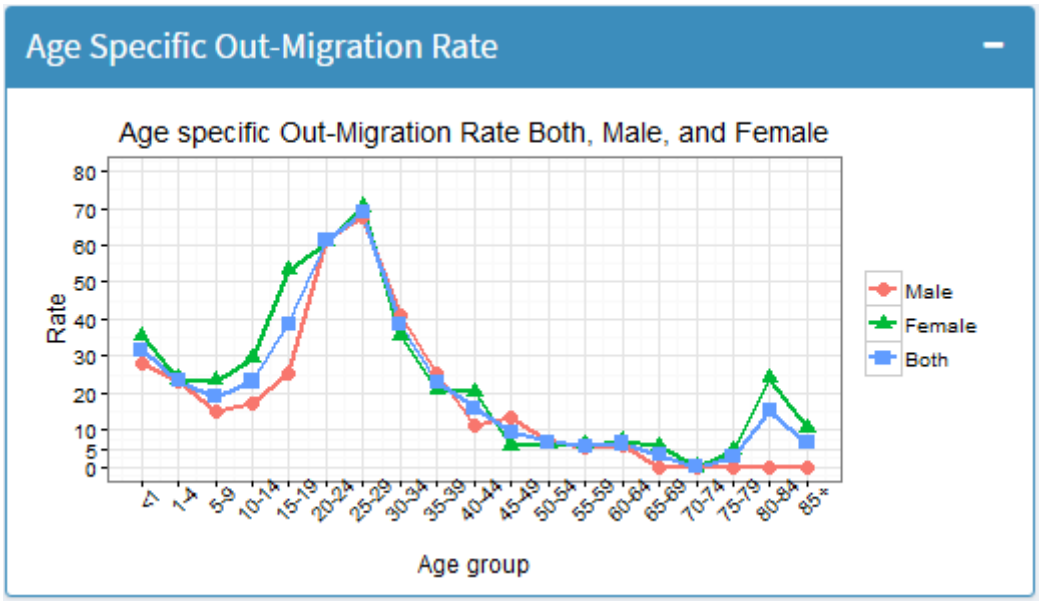

Figure 20: Out-Migration Rate by Age Group - Niakhar, 2014

the core HDSS but are often conducted depending on the available funding. To share the findings of those studies, a module was implemented to import dataset into the infographic platform and perform adequate visualisation and insight. This module allows researchers to have a quick descriptive analysis of any dataset collected from the HDSS.

The cohort study plugin consists of importing and visualising data from the longitudinal studies to the infographic platform. The principle was to import a dataset in a comma de-

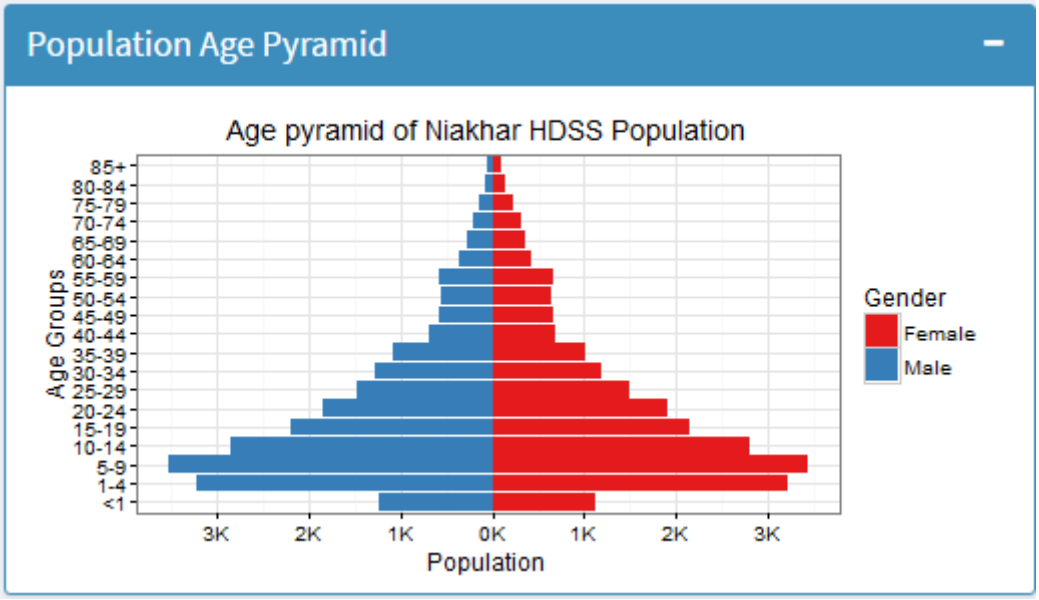

Figure 21: Population Age Pyramid

limited format (CSV) using the *Import CSV* control. After importing the CSV file, the user can choose the coordinate to be plotted into a graph. For the X-Axis, only one variable can be chosen but multiple variables can be plotted in the Y-Axis for the purpose of comparison (see Figure 22).

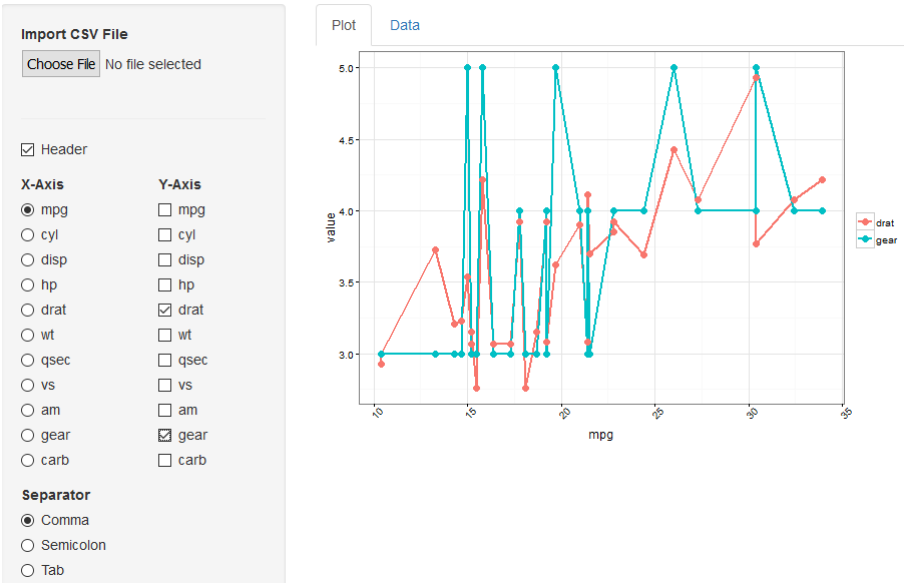

Figure 22: Cohort Study Plugin

The cohort plugin was developed to assist health researchers in determining relationship between two phenomena. For example, one can think about getting the link between mortal-

ity rate and malaria incidence or total fertility rate and use of contraceptives in the HDSS area over time using the same cohort and can draw a conclusion on causes of death related to malaria. Figure 23 illustrate malaria incidence rate from 2000 to 2016 combined with the mortality rate for the same cohort. This can determine whether the increase of malaria cases can increase death rate in the same area.

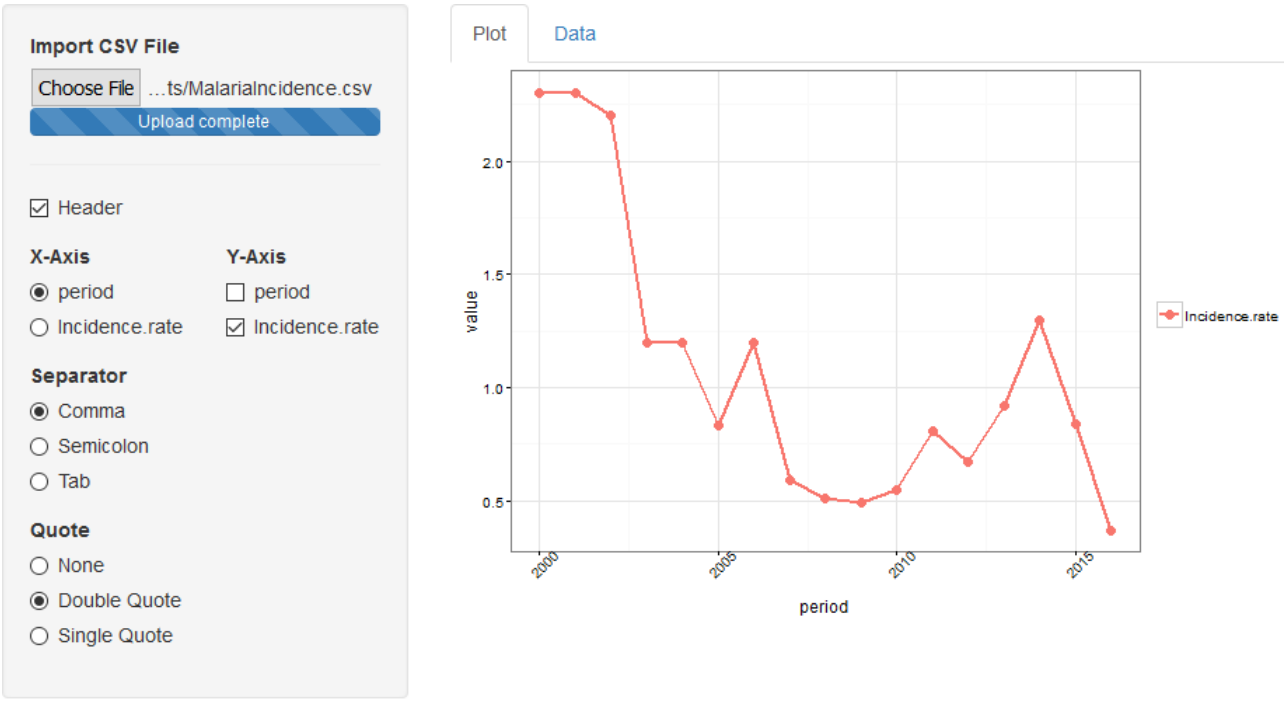

Figure 23: Malaria Incidence Rate from 2000-2016

4.4 COMBINING DISSEMINATION STRATEGIES

The dissemination of research evidence using visualisation has become crucial to heath research. It is however limited by a lot of factors such as internet connectivity, literacy level of the targeted audience etc. It is important to share research evidence with a variety of stakeholders to increase understanding and awareness and the use of web visualisation may not be adequate to reach some of the audiences.

The use of the web visualisation platform is mostly convenient for affluent decision-makers to influence and encourage changes in policies and practices. For greater impact, the infor-

mation found on fertility and mortality displays need to be shared using other dissemination channels such as mass media and print out poster converted into the local language. These results can be communicated using radio and/or television in the local language via talk shows or theatre to address to the targeted community. Sharing information in the language of the local community makes the findings understandable and can increase awareness and affect the behaviour and lifestyle changes in the communities.

The visuals generated by the system can be used in newsletters to inform the policy-makers and community leaders on the research evidence. These newsletters can be automatically shared via email to give the community leader and policy-makers real time updates on the ongoing research.

5

IMPLEMENTATION AND DEPLOYMENT CONSIDERATIONS

In this chapter, we describe the guidelines and actions to take into consideration in the deployment of the platform for other HDSS sites. We also explain the specification required, in term of infrastructure and hosting needed to deploy the platform. The chapter will also give pointers on how the visualisation platform will be used with others dissemination channels to enhance community engagement activities into research.

5.1 DEPLOYMENT CONSIDERATIONS IN OTHER HDSS SITES

The INDEPTH Network HDSS sites use different database structure to store surveillance dataset. Each site uses an in-house data model to stores data from the routine longitudinal surveillance. However, they all utilise the same core fundamental demographic models during each cycle of the surveillance. These are primarily Birth, Death, and Migration dataset (Figure 24). The entity structures and schemes may differ significantly from one site to another but HDSS uses the standard INDEPTH data model to share their data in the data repository [18]. This makes the synchronisation of the core HDSS dataset from other sites to our platform daunting but doable. INDEPTH iShare is a project funded by Welcome Trust to enhance the research data management capacity of INDEPTH member centres and enable them to develop, document, extract, harmonise and create analytical datasets from their operational databases [1], allowing the HDSS site to make data available to the scientific community using the INDEPTH Data Repository platform. The Infographic platform was implemented in accordance to this iShare data dictionary, described in Table 4, in order

1 http://www.indepth-ishare.org/index.php/about

Variables	Data type	Description
CountryId	Number	Country Code of the HDSS
CentreId	Number	INDEPTH Identifier of the HDSS
IndividualId	Text	Identifier of the individual
Sex	Text	Sex of the Individual (Value: 1=Male, 2=Female)
DoB	Date	Date of Birth
EventCount	Number	Number of occurred even during the follow up
EventNr	Number	Sequence Number of the Event
EventCode	Text	Code of the Event
EventDate	Date	Date that the event occurred
ObservationDate	Date	Date of observation correspond of the date of collect
LocationId	Number	Identifier of the Village
MotherId	Number	Identifier of the Mother
DeliveryId	Number	Identifier of the Delivery in case of Pregnant woman

Table 3: INDEPTH - iSHARE Data Dictionary

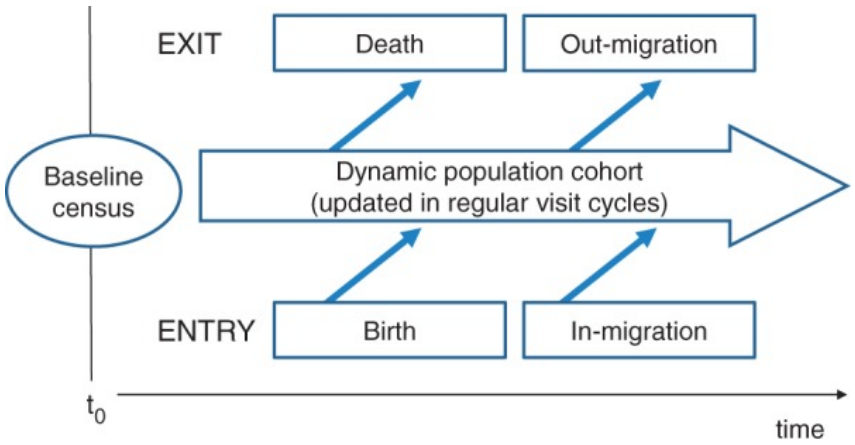

Figure 24: INDEPTH HDSS from Dynamic Cohort Model

to allow different HDSS site to deploy the developed system by extracting the INDEPTH core dataset from their databases using and Extract Transform and Load (ETL) process. The Event Code and Event Date represent the entry and exit events and the date of their occurrence as illustrated in Figure 24. These variables are sorted chronologically according to a changing sequence event number created.

In running this platform, the deployment requires installation of R engine and the processing of all application logic requires shiny web server to handle the user request. To effectively use this platform, HDSS are required to extract data from their longitudinal database using

SQL query or Pentaho Data Integration using the standard variable format in Table 4. The ETL was developed according to the data model of the HDSS.

5.2 INFRASTRUCTURE CONSIDERATION

The developed infographic platform requires hosting on an in-house server or on a cloud hosting service with the following specifications: 8GB or more of RAM for the production instance and minimum of 100GB disk space to install the dataset and host the code on the platform. However, the disk space specification will depend on the size of the HDSS datasets. The platform was developed for a cross-platform environment and can be hosted using Widows and Unix operating system. It requires R and the shiny package installed. For HDSS sites that have opted for cloud hosting service, they need to use the shiny cloud server through the link http://www.shinyapps.io/. Shiny could server offers secure hosting and scalability as the applications and users increase.

5.3 USE OF THE INFOGRAPHIC PLATFORM TO ENHANCE COMMUNITY ENGAGEMENT

Different stakeholders in health research need to be informed, learn, and discuss from research outputs. To stimulate this interaction between stakeholders, the community engagement departments of HDSS sites organise seminars or community meetings to facilitate community engagement for knowledge exchange or in decision-making process. Appropriate tools are needed to facilitate the process. For instance, the developed Infographic platform could be utilised by these HDSS community engagement section to encourage learning, dialogue, and discussion between the stakeholders by sharing relevant information of the research project to help improving understanding and behavioural change. The Infographic can enhance engagement and interaction of communities for knowledge exchanges. However,

visualisation cannot be the only channel for such engagements. Different channel could be combined with such visualisation depending on the location and the audience (print out report translated in a different language, communication via Radio and Television, etc.) to reach the targeted audience as discussed in Section 4.4.

6

CONCLUSION, RECOMMENDATIONS AND FUTURE DIRECTIONS

With decades of research in the Niakhar community, the population is still asking questions about the main interest of routine data collection and some are not engaged anymore by refusing to open their doors to the HDSS. Although the emerging of the Information Communication Technologies (ICT) is prevalent, Niakhar and other INDEPTH affiliated HDSS sites are still behind in terms of dissemination of research result to enhance community engagement. This may be related to the absent of community engagement section in most of the INDEPTH affiliated HDSS for creating community engagement activities for information sharing and learning to create a network with the community. Those sites that do have such community engagement units may not have the requisite tools to extract insightful information from the core HDSS longitudinal datasets.

To improve community engagement, we have implemented a web platform to help research institutions to disseminate research finding to the local community and the stakeholders throughout infographics. The developed platform allows stakeholders to have a visual representation of the finding with aggregated in a graphical format and help researcher and HDSS on dissemination of research findings to the local community. HDSSs need to create community engagement activities and involve the community leaders in the research studies. They must also use other methods of dissemination and channels to allows the community to understand the message and improve health outcome and behaviour change especially in rural area where there is limited internet access.

The following activities have been identified for future work to improve the platform and increase its usability:

- This platform can be very useful to the various stakeholders in the health institution and structures in the countries of the HDSSs. This is an extension of the platform that allows stakeholders and other institutions to access the platform. This extension has been an earmark for future work.

- The system can be used as an early warning system for the local communities. This will allow HDSSs to conduct active syndromic surveillance instead of the current practice of passive surveillance. This can be implemented in future.

- The system can be used to show how dissemination channels using infographic may affect the uptake of research findings. This can also be implemented in future. Usability test to evaluate the platform and its usage in the HDSS data process cycle is needed to allow ease of use in other HDSS sites.

A

PLAGIARISM DECLARATION

WITS School of Public Health

Plagiarism declaration for written work: Master of Science in Epidemiology

I Djibril DIONE (Student number: 981721) am a postgraduate student registered for degree Epidemiology/Research Data Management in the Wits School of Public Health.

I hereby declare the following:

- **I am aware that plagiarism is the use of someone else's work without their permission and/or without acknowledging the original source).**
- **I am aware that plagiarism is wrong.**
- **I confirm that the work submitted for assessment for the above course is my own unaided work except where I have explicitly indicated otherwise.**
- **I have followed the required conventions in referencing the thoughts and ideas of others.**
- **I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this in not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.**

Signature:  **Date: 13/02/2017**

Figure 25: Plagiarism Declaration

B

ETHICS CLEARANCE CERTIFICATE

R14/49 Mr Djibril Dione

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M151032

NAME: (Principal Investigator)	Mr Djibril Dione
DEPARTMENT:	School of Public Health Institut de Recherche pour le Développement (IRD) Dakar, Senegal
PROJECT TITLE:	Information Feedback to Local Communities of HDSS Sites Using Infographics: The Case of Niakhar HDSS Site
DATE CONSIDERED:	30/10/2015
DECISION:	Approved
CONDITIONS:	South African Human Research Ethics Committees (HRECs) have no standing outside South Africa.
SUPERVISOR:	Gideon Nimako
APPROVED BY:	Professor P Cleaton-Jones,Chairperson, HREC (Medical)
DATE OF APPROVAL:	02/03/2016

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary in Room 10004, 10th floor, Senate House/2nd Floor, Phillip Tobias Building, Parktown, University of the Witwatersrand. I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. **I agree to submit a yearly progress report**.

Principal Investigator Signature Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Figure 26: Ethics Clearance Certificate

C

SOURCE CODE

The source code of the functions used int this Project can be found in google drive following this link:

https://drive.google.com/drive/folders/0BxDLTClKpTQWQWE2TU9aTTI5d0E?usp=sharing

BIBLIOGRAPHY

[1] Adriana Arcia, Michael E Bales, Manuel C William Brown III, et al. "Method for the development of data visualizations for community members with varying levels of health literacy." In: *AMIA Annual Symposium Proceedings*. Vol. 2013. American Medical Informatics Association. 2013, p. 51.

[2] OC Ariyo, MO Ariyo, OE Okelola, OS Aasa, OG Awotide, AJ Aaron, and OB Oni. "Assessment of the role of mass media in the dissemination of agricultural technologies among farmers in Kaduna North Local Government Area of Kaduna State, Nigeria." In: (2013).

[3] Michael Bender, Ralf Klein, Andreas Disch, and Achim Ebert. "A functional framework for web-based information visualization systems." In: *IEEE Transactions on Visualization and Computer Graphics* 6.1 (2000), pp. 8–23.

[4] Ross C Brownson, Matthew W Kreuter, Barbara A Arrington, and William R True. "From the Schools of Public Health." In: *Public health reports* 121.1 (2006), pp. 97–103.

[5] CDC. *Program Evaluation Guide- Step 6 - CDC*. 2012. URL: http://www.cdc.gov/eval/guide/step6/ (visited on 11/12/2015).

[6] Pierre Cantrelle. "Etude démographique dans la région du Sine-Saloum (Sénégal): état civil et observation démographique 1963-1965." In: (1969).

[7] Lauren N Carroll, Alan P Au, Landon Todd Detwiler, Tsung-chieh Fu, Ian S Painter, and Neil F Abernethy. "Visualization and analytics tools for infectious disease epidemiology: a systematic review." In: *Journal of biomedical informatics* 51 (2014), pp. 287–298.

[8] Daniel Chandramohan, Kenji Shibuya, Philip Setel, Sandy Cairncross, Alan D Lopez, Christopher JL Murray, Basia Żaba, Robert W Snow, and Fred Binka. "Should data from demographic surveillance systems be made more widely available to researchers?" In: *PLoS Med* 5.2 (2008), e57.

[9] Hein De Haas. "Migration and development: a theoretical perspective1." In: *International migration review* 44.1 (2010), pp. 227–264.

[10] Valerie Delaunay, Laetitia Douillot, Aldiouma Diallo, Djibril Dione, Jean-François Trape, Oleg Medianikov, Didier Raoult, and Cheikh Sokhna. "Profile: the Niakhar health and demographic surveillance system." In: *International journal of epidemiology* 42.4 (2013), pp. 1002–1011.

[11] Maridelys Detres, Robert Lucio, and Judi Vitucci. "GIS as a community engagement tool: developing a plan to reduce infant mortality risk factors." In: *Maternal and child health journal* 18.5 (2014), pp. 1049–1055.

[12] Mary Duffy. "The Internet as a research and dissemination resource." In: *Health promotion international* 15.4 (2000), pp. 349–353.

[13] Robin Featherstone. "Visual research data: An infographics primer." In: (2014).

[14] Krista Francis, Michele Jacobsen, and Sharon Friesen. "The use of graphics to communicate findings of longitudinal data in design-based research." In: *Journal of Information Technology Education: Research* 13 (2014), pp. 233–255.

[15] M Michel Garenne. "Fertility changes in Sub-Saharan Africa." In: (2008).

[16] Malu AC Gatto. "Making Research Useful: Current Challenges and Good Practices in Data Visualisation." In: *Reuters Institute for the Study of Journalism with the support of the University of Oxford's ESRC Impact Acceleration Account in partnership with Nesta and the Alliance for Useful Evidence. Available at: https://reutersinstitute politics. ox. ac. uk/publication/making-research-useful (accessed March 2016)* (2015).

[17] Agincourt HDSS. *Learning, Information dissemination and Networking with the Community | Agincourt*. 2015. URL: http://www.agincourt.co.za/index.php/activities/linc/ (visited on 09/30/2015).

[18] Kobus Herbst, Sanjay Juvekar, Tathagata Bhattacharjee, Martin Bangha, Nidhi Patharia, Titus Tei, Brendan Gilbert, and Osman Sankoh. "The INDEPTH Data Repository An International Resource for Longitudinal Population and Health Data From Health and Demographic Surveillance Systems." In: *Journal of Empirical Research on Human Research Ethics* 10.3 (2015), pp. 324–333.

[19] INDEPTH. *INDEPTH Stat*. 2016. URL: http://www.indepth-ishare.org/indepthstats.

[20] Hudron K. Kari. "Availability and accessibility of ICT in the rural communities of Nigeria." In: *The Electronic Library* 25.3 (2007), pp. 363–372. DOI: 10.1108/02640470710754869. eprint: http://dx.doi.org/10.1108/02640470710754869. URL: http://dx.doi.org/10.1108/02640470710754869.

[21] Joy E Lawn, Anthony Costello, Charles Mwansambo, and David Osrin. "Countdown to 2015: will the Millennium Development Goal for child survival be met?" In: *Archives of disease in childhood* 92.6 (2007), pp. 551–556.

[22] R Craig Lefebvre. "The new technology: the consumer as participant rather than target audience." In: *Social Marketing Quarterly* 13.3 (2007), pp. 31–42.

[23] Nancy B Mueller, Ryan C Burke, Douglas A Luke, and Jenine K Harris. "Getting the word out: multiple methods for disseminating evaluation findings." In: *Journal of Public Health Management and Practice* 14.2 (2008), pp. 170–176.

[24] World Health Organization et al. *Implementation research toolkit*. WHO/TDR, World Health Organization, 2014.

[25] Randall M Rohrer and Edward Swing. "Web-based information visualization." In: *IEEE Computer Graphics and Applications* 17.4 (1997), pp. 52–59.

[26] Laurence Roulleau-Berger. *Migrer au féminin*. Presses universitaires de France, 2015.

[27] Osman Sankoh and Peter Byass. *The INDEPTH Network: filling vital gaps in global epidemiology*. 2012.

[28] Olivier Schmitz. "Les points d'articulation entre homéopathie et oncologie conventionnelle. Une enquête ethnographique auprès de praticiens et d'usagers de l'homéopathie." In: *Anthropologie & Santé. Revue internationale francophone d'anthropologie de la santé* 2 (2011).

[29] William J Schroeder and Kenneth M Martin. "The Visualization Toolkit-30." In: ().

[30] CJ Simiyu, V Naanyu, AA Obala, DO Odhiambo, P Ayuo, D Chelagat, et al. "Establishing Webuye Health and Demographic Surveillance Site in rural western Kenya: Challenges and lessons learned. Popualtion Association of America." In: *2013 Annual Meeting. Conference Paper*. 2013.

[31] Stacy E Sneeringer. "Fertility transition in sub-Saharan Africa: A comparative analysis of cohort trends in 30 countries." In: (2009).

[32] Smith Tom, Noble Stefan, Avenell David, and Lally Graham. *DataViz: Improving Data Visualisation for the Public Sector*. 2009. (Visited on 11/13/2015).

[33] Ondenge K McLellan-Lemal E Awuonda E Angira F Mills L et. al. "Disseminating results: community response and input on Kisumu breastfeeding study." In: *Translational behavioral medicine* 5.2 (2015), pp. 207–15.