ACKNOWLEDGMENTS

This field handbook was prepared by RIA senior research fellows, Dr Christoph Stork and Steve Esselaar with inputs from the field from RIA training associate, Mariama Deen-Swarray; South African survey delivery partner iKapadata; senior researcher Mothobi Onkokame and RIA research manager, Dr Enrico Calandro. The efforts of all the local partners in Nigeria (Fola Odufuwa), Rwanda (Albert Nsengiyumva) Kenya (Dr Margaret Nyambura) Mozambique (Francisco Mabila) South Africa (Jan Skenk) and Tanzania (Dr Godiel Moshi) in getting the surveys done during 2017 under often challenging circumstances, are gratefully acknowledged. For policy papers and indicators arising from these surveys see the Research ICT Africa website www.researchICTafrica.net. The 2017 questionnaire was developed as part of a Global South ‘After Access’ 18-country #afteraccess survey undertaken by RIA together with LIRNEasia and DIRSI. The ‘After Access’ ICT access and use surveys are made possible with the support of the Canadian International Development Research Centre (IDRC).

Alison Gillwald
Executive Director
2017
# Table of Contents

ACKNOWLEDGMENTS 2  
Project Summary 1  
Country Mission 3  
Check list: What to bring 3  
Arrival in Country - at the airport 3  
Enumerators 3  
Preparing the data collection devices 5  
Questionnaire Training 6  
Listing 7  
**Sampling** 11  
Sample Size 12  
Weighting 12  
**GIS Maps** 14  
Hard copy maps 14  
Shape files 14  
**ODK tips and tricks** 15  
**Example South Africa** 16  
Sampling 16  
Listing 18  
**Definitions and Resources** 20  
References 21
# Project Summary

<table>
<thead>
<tr>
<th>Project Name</th>
<th>ICT indicators4Africa #afteraccess</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Concept</strong></td>
<td>To pool public, private and donor resources to undertake resource-intensive national representative household surveys to produce the information communication technology (ICT) indicators and analysis required for evidence-based policy and regulation on the African continent and to thereby avoid duplication, proprietary research that could serve the public interest and form part of a wider information and policy commons for the continent.</td>
</tr>
<tr>
<td><strong>Host Organisation</strong></td>
<td>Research ICT Africa (RIA), University of Cape Town (ZA non-profit 2009/017831/08)</td>
</tr>
<tr>
<td><strong>Institutional history</strong></td>
<td>With the support of the Canadian <em>International Development Research Centre</em> (IDRC) nationally representative surveys have been run three times in countries across the continent over the last decade. RIA data is the only demand-side data systematically collected in Sub-Saharan Africa that is capable of providing actual affordability data and allowing estimation of unmet demand (non-users’ willingness to pay). As such, it is referenced in evidence-based public policy debates and by multilateral organisations such as the International Telecommunications Union, UNCTAD and the World Bank. Examples include: ITU: Chapter 5 - Increasing Internet use: the role of education, income, gender, age, and location, in <em>Measuring the Information Society 2011</em>, International Telecommunications Union. <a href="http://www.itu.int/net/pressoffice/backgrounders/general/pdf/5.pdf">http://www.itu.int/net/pressoffice/backgrounders/general/pdf/5.pdf</a> UNCTAD: ICTs, Enterprises and Poverty Alleviation in <em>Information Economy Report 2010</em> <a href="http://unctad.org/en/pages/PublicationArchive.aspx?publicationid=1575">http://unctad.org/en/pages/PublicationArchive.aspx?publicationid=1575</a></td>
</tr>
<tr>
<td><strong>Problem</strong></td>
<td>Advocates of improving research for development policy make a mistake when they take for granted the availability of hard data as the foundation of policy advice. Most African countries suffer a severe shortage of basic ICT statistical data and analysis that is fundamental to correctly identifying points for policy and regulatory intervention. National statistics offices and regulators generally do not collect the demand-side data needed to measure ICT access and use to determine current policy and regulatory outcomes and thereby identify points of policy intervention to meet public interest objectives. They are therefore unable to report to international and multilateral agencies (ITU, WB, etc.) for them to accurately reflect on the position of Africa in global indices. While the ITU collects supply-side (subscribers/pricing) statistics, which are also drawn on by WB, OECD and WEF for comparative evidence-based telecom policy discussions, no organisation collects corresponding demand-side (usage/spending) statistics across Sub-Saharan Africa. For this reason, one can talk about service prices but not actual affordability of services, except in broad terms (estimating affordability by fraction of income – where the fraction considered affordable is extrapolated from OECD survey numbers). Similarly, service uptake statistics (internet use, mobile subscriptions and non-usage rates, etc.) are very dated, unreliable, unsystematic and extremely inaccurate. Disaggregation of data by income level, gender, age and urban-rural divide is not possible with supply-side data.</td>
</tr>
<tr>
<td><strong>Objectives</strong></td>
<td>To collect a range of household and individual ICT indicators (including access to and use of fixed, mobile and internet services, ICT spend, non-users’ willingness to pay) that meet the threshold compliance of the WSIS-initiated Partnership for Measuring ICT for Development by running nationally representative household and individual surveys (with NSO or NRA where possible) in order to build a national and regional evidence base to inform policy and regulation.</td>
</tr>
<tr>
<td><strong>Project Description</strong></td>
<td>Gather, analyse, and publicise demand-side (non-users and users) data in Africa in 2016-2017 through nationally representative sampling (census sample frame) enabling disaggregation of data to understand ICT access and use by urban poor, rural inhabitants, youth, women, bottom of the pyramid and other potentially marginalised groups.</td>
</tr>
<tr>
<td><strong>Methodology</strong></td>
<td>Nationally representative surveys based on national statistical office (NSO) census sample frames of households and individuals aged 15yrs+; (nationally indicative informal business survey in residential areas if undertaken); focus groups and/or ethnographic research to support the quantitative data analysis. Deployment of appropriate modeling techniques (fixed effect and instrumental variable models) to measure impacts of policy and regulatory interventions.</td>
</tr>
<tr>
<td><strong>Expected Project Outcomes</strong></td>
<td>Delivery of demand-side indicators and analysis essential to inform evidence-based policy that is more accurate and timely, and which will create the data (time-series and cross-sectional) to enable in-depth analysis of policy outcomes and points of intervention. The data will allow regulators and policy makers to measure the impact of policies and regulation. International organisations (ITU, UNCTAD) would receive the only demand-side data in the public domain to verify supply-side data and enable the inclusion of Africa into comparative demand-side studies and analyses. The data would be made publicly available for further use by governments, research institutions, NGOs, industry and trade unions to enable more informed participation in public policy processes and to complement, and be informed by, other pricing, quality of services and institutional analysis research projects that collectively can provide a better evidence base.</td>
</tr>
</tbody>
</table>
Country Mission

Check list: What to bring

- 2 Mobile phones, one for home SIM card, one for country SIM card.
- Plug adapters
- Box of badges
- Pre-drafted intro letter for enumerators, laminated (a copy for each enumerator)
- Large scissors
- Box of chalks
- Sunhat
- Mosquito repellent
- Small Umbrella for rainy countries
- Cheap day bag (Woolworth shopping bag) for field training
- Digital camera or reasonable smart phone to take profile pictures of enumerators.

Arrival in Country - at the airport

When arriving in a country, it is important to get a SIM card and have local currency at hand before leaving the airport. It is more difficult to get hold of a SIM card later, in particular if registration is required and exchanging money in hotels is generally more expensive than at the airport. Suggested sequencing is:

1. Exchange USD 100 in local currency, either withdrawing with credit card or cash.
2. Get a local SIM card (choose one of the larger operators when leaving the capital).
3. Load 1 GB data. If hot spotting the phone for computer use budget 2G per week minimum.
4. Load about USD 2 additional airtime for voice calls.
5. Log into WhatsApp and confirm that you are using your home number. The home number needs to be in a phone as well since they may send you a confirmation SMS.
6. Look for person that picks you up or organise local transport.

Enumerators

Enumerator IDs and Passwords

Enumerator IDs can be generic. For example, for Rwanda it could be:E-Rwanda_1, E-Rwanda_2, E-Rwanda_3….The passwords need to be easy to remember but still a bit safe. Password for E-Rwanda_1 could be E-Rwanda_1d. The last letter should be arbitrarily chosen. Use these Enumerator ID and Passwords to register each enumerator as a user on ONA. Help on the ONA registration and first steps can be found here.
Enumerator Database and Badges


![Image of Enumerator Database](image1)

**Figure 1: Table entry of enumerator details**

The Filemaker database allows you to capture data from all enumerators and allocate IDs. The database can easily be modified to include further information such as ONA user names and passwords, bank details etc.

![Image of Badge with only RIA logo](image2)

**Figure 2: Badge with only RIA logo**

The Database has 2 default label (badge) views, one with RIA logo only and one with RIA logo and a logo from a partner organisation. Size and layout can easily be adjusted by anyone with basic Filemaker knowledge.

![Image of Batch with logos from RIA and Partner organisation](image3)

**Figure 3: Batch with logos from RIA and Partner organisation**
Enter name and contact details in database and print labels as PDF. The badges will be printed as PDF and can be printed in the Hotel or in a print shop.

Bring badges from home. They may be difficult to purchase in-country and purchasing it will waste time.

Letter of Introduction

The letter needs to contain what the survey is about and that all necessary documentation, such as a research visa (Rwanda) has been granted.

The letter should also contain the name, contact details and signature of the local partner.

It is best to print the letters and laminate before you arrive in the country.

You will need at least one introductory letter per enumerator.

Preparing the data collection devices

Software that needs to be installed on the Android Smartphone includes:

- ONA Collect
- NetRadar

ONA Collect

- Enter Enumerator ID and Password generated for enumerator in previous chapter.
- Once signed in, click on Get Blank Forms. This will allow the user to download all questionnaires that the user is authorised for. It is important that all enumerators are registered as users on ONA.
- Once forms are downloaded, the enumerator can click on “Fill Blank Form” to start an interview.
NetRadar

NetRadar allows Quality of Service (QoS) testing in the field by the field teams. This QoS data for each EA can be linked to ICT adoption models. We have set up a single email address and password for a NetRadar account for that purpose, which is:

**Username:** netradar@researchictolutions.com

**Password:** ICThh2017

NetRadar should be set to test at least 20 times a day. Our contact at Netradar will provide us country level data for Netradar in addition to the data which is accessible through the account given above:

Jukka MJ Manner, Professor, PhD.
Department of Communications and Networking (Comnet) / Aalto University
**Mobile:** +358+(0)50+5112973 / **Fax:** +358+(0)9+470 22474 / **E-mail:** jukka.manner@aalto.fi

http://www.netradar.org

---

**Figure 4: ONA Connect sign in screen and follow-up screens**

**Figure 5: Netradar example output Main, preference screens and example output**

---

**Questionnaire Training**

The questionnaire training encompasses various levels of knowledge transfer. Enumerators need to understand the subject matter and the various ICTs included in the survey need to be explained. Enumerators must be able to operate the electronic ODK
version and must be able to ask and record questions in multiple languages and dialects. The format of the training process is as follows:

- **Step 1:** Questionnaire training: Connect ODK device to a projector or use webform. This is better than using the MS Word or PDF hard copies. This way the enumerators get directly exposed to what they will be using.
- **Step 2:** Enumerators interview each other in English.
- **Step 3:** Enumerators interview each other is another relevant language.
- **Step 4:** Enumerators interview the trainer. Choose those that seem to struggle during the one-on-one tests. Sometimes enumerators need to be excluded from the survey if they are not catching on quick enough.
- **Step 5:** Enumerators interview real households and businesses during a pilot. This is done at the same time as the listing training. The main purpose here is to get feedback on how the questionnaire flows and what additional language changes need to be made.

The training may reveal additional changes that need to be made to the local language version. A taxi can be called many things in Africa, for example. These will be discussed at the wrap-up session on the last day of the five day enumerator training.

### Listing

Part of the pilot is the listing training. Listing is first discussed in the class room during the training and then practised in the pilot EA. All structures in an EA need to be listed and given a structure ID. All households and businesses within an EA structure need to be listed. Households can be identified by household head name and businesses by name of the business owner. The EA questionnaire consists of two forms: the Cover page displayed in Table 1 and the listing form displayed in Table 2.

<table>
<thead>
<tr>
<th>Table 1 EA Questionnaire and Listing Form Part 1</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong></td>
<td></td>
</tr>
<tr>
<td>Q.1 EA_ID (12 digit)</td>
<td></td>
</tr>
<tr>
<td>Q.2 Province</td>
<td></td>
</tr>
<tr>
<td>Q.3 District</td>
<td></td>
</tr>
<tr>
<td>Q.4 County Name</td>
<td></td>
</tr>
<tr>
<td>Q.5 Sub county Name</td>
<td></td>
</tr>
<tr>
<td>Q.6 Parish Name</td>
<td></td>
</tr>
<tr>
<td>Q.7 Number of households in last Census</td>
<td></td>
</tr>
<tr>
<td>Q.8 Type: 1= Urban 0= Rural</td>
<td></td>
</tr>
<tr>
<td>Q.9 Mobile Phone Coverage in EA: [0] No [1] Yes</td>
<td></td>
</tr>
<tr>
<td>Q.10 Time of Listing start (HH:MM)</td>
<td></td>
</tr>
<tr>
<td>Q.11 Date of Listing</td>
<td></td>
</tr>
<tr>
<td>Q.12 Name of Supervisor</td>
<td></td>
</tr>
</tbody>
</table>
Table 1 EA Questionnaire and Listing Form Part 1 | Example
--- | ---
**Household** |  
HH.1 | Total Number Household in EA | 200  
HH.2 | HH Sampling target for EA | 20  
HH.3 | Sampling Interval = Total number of hh in EA / sampling target for EA | 10  
HH.4 | Random Starting Point (add 4 digits of the starting time = H+H+M+M) | 17  
**Businesses** |  
B.1 | Total Number Businesses in EA | 50  
B.2 | Business Sampling target for EA | 10  
B.3 | Sampling Interval = Total number of businesses in EA / sampling target for EA | 5  
B.4 | Random Starting Point (add 4 digits of the starting time = H+H+M+M) | 7

The Cover page (Table 1) is completed last and summarises the listing. The Listing form (table 2) may be 10 - 20 pages and each enumerator may complete several pages on his or her own during the listing process. Routes demarcated by letters can be laid through the EA. Figure 6 provides an example. One enumerator can be sent from A to B to apply the Right Hand Method (RHM) to list structures, households and businesses. Another can be sent from B to C to D and so forth.

![Figure 6: EA Map drawn from Google Earth - Michell's Plain, South Africa](image)

The Listing process is as follows:

- Arrive at EA and determine the boundary by walking around it and checking the various landmarks.
- Mark letters in chalk on the road or elsewhere easily visible.
Allocate enumerators walking routes.

Enumerators each fill their own listing form (Table 2).

The structure IDs used by enumerators must correspond with the route: someone walking from A to B would use structure IDs AB001, AB002 etc. so that there will be no duplication of structure IDs.

Enumerators write the Structure ID in easily visible chalk on the gate or house wall and enter the household and business information in Form 2 after consultations with owners and residents.

Once listing is complete the field manager walks through EA to make sure no structures were overlooked. Every structure must have a structure ID.

The field manager takes all the listing forms for the EA and staples them together with the cover page (Table 1) on top. The order of the listing form does not play any role. **But once stapled together the order must remain.**

The field manager then allocates serial numbers to each household and each business on the listing form. Households starting with HH1, HH2, etc and business with B1, B2, etc.

The last serial number indicates the number of households in the EA and the number of businesses in the EA. These two figures are recorded on the cover page (Table 1).

The field manager then determines the sampling interval by dividing the number of households in the EA by the target number of households to be interviewed.

The interval is always rounded down.

The field manager then notes down the time in hours and minutes and adds the digits together. 13:47 = 1+3+4+7=15, for example. This number is the random starting point. The first selected HH is HH15. The next HH15+ the sampling interval and so forth.

The interval can be continued from the start if the end of the listing form is reached and the required number of households have not yet been interviewed.

Random replacements are selected by continuing with the sample interval. For example, if there are 257 households in the EA, the last one selected household is HH253, and the sample interval is 10, then the next selected household is HH6 and after HH6 would be HH16 etc.

Should the same number coincidentally be selected again, then generate a new random starting point.

**Table 2: EA Questionnaire and Listing Form Part 2**

<table>
<thead>
<tr>
<th>Field Manager</th>
<th>Listers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Business Serial</strong></td>
<td><strong>HH Serial</strong></td>
</tr>
<tr>
<td>B1</td>
<td>AB001</td>
</tr>
<tr>
<td>HH1</td>
<td>AB001</td>
</tr>
<tr>
<td>HH2</td>
<td>AB002</td>
</tr>
<tr>
<td>Business Serial</td>
<td>HH Serial</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------</td>
</tr>
<tr>
<td>B2</td>
<td>CD007</td>
</tr>
<tr>
<td>HH3</td>
<td>AB003</td>
</tr>
<tr>
<td></td>
<td>CD023</td>
</tr>
<tr>
<td>HH4</td>
<td>CD024</td>
</tr>
</tbody>
</table>
Sampling

The random sampling for households, individuals and businesses is based on Census sample frames. A Census divides a country in Enumerator Areas (EAs) which roughly have a household density of 200.

- Step 1: The national census sample frames was split into urban and rural Enumerator areas (EAs).
- Step 2: EAs were sampled for each stratum using probability proportional to size (PPS).
- Step 3: For each EA two listings were compiled, one for households and one for businesses. The listings served as sample frame for the simple random sections of households and businesses.
- Step 4: \(X\) Households and \(Y\) businesses were sampled using simple random sample for each selected EA.
- Step 5: From all household members 15 years or older or visitor staying the night at the house one was randomly selected based on simple random sampling.

![Diagram of sampling steps](image)

*Figure 7: Sampling steps*
The number of households per EA, or the extent of clustering, should be between 10 and 20. NSOs typically sample 15 households per EA.

When moving with a team, 20 is more convenient (5 questionnaires per enumerator for a team of 4, and can still fit in a regular car) and also reduces the number of EAs to be listed at a given sample size. Pushing it higher than 20 increases geographic/social risk, i.e., sampled EA’s might not reflect the diversity of a country.

**Sample Size**

The desired level of accuracy for the survey was set to a confidence level of 95% and a margin of error of 5%, which yields a minimum sample size per tabulation group of 385.

<table>
<thead>
<tr>
<th>Margin of error</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>+1</td>
<td>9.604</td>
<td>16.577</td>
</tr>
<tr>
<td>+2</td>
<td>2.401</td>
<td>4.145</td>
</tr>
<tr>
<td>+3</td>
<td>1.068</td>
<td>1.842</td>
</tr>
<tr>
<td>+4</td>
<td>601</td>
<td>1.037</td>
</tr>
<tr>
<td>+5</td>
<td>385</td>
<td>664</td>
</tr>
<tr>
<td>+6</td>
<td>267</td>
<td>461</td>
</tr>
<tr>
<td>+7</td>
<td>196</td>
<td>339</td>
</tr>
<tr>
<td>+8</td>
<td>151</td>
<td>260</td>
</tr>
<tr>
<td>+9</td>
<td>119</td>
<td>205</td>
</tr>
<tr>
<td>+10</td>
<td>97</td>
<td>166</td>
</tr>
</tbody>
</table>

Source: Rea and Parker (2014)

**Weighting**

Three weights need to be constructed: for households, individuals and informal businesses. The weights are based on the inverse selection probabilities and gross up the data to national level when applied.

Household weight:  
\[ HH_w = DW \frac{1}{P_{HH} \times P_{EA}} \]

Individual weight:  
\[ IND_w = DW \frac{1}{P_{HH} \times P_{EA} \times P_l} \]

Business Weight:  
\[ Bus_w = DW \frac{1}{P_{Bus} \times P_{EA}} \]

See UNSD (2005) page 119 for a detailed discussion on sampling weights.
Household Selection Probability: \[ P_{HH} = \frac{n}{HH_{EA}} \]

EA Selection Probability: \[ P_{EA} = \frac{HH_{EA}}{HH_{STRATA}} \]

Individual selection Probability: \[ P_i = \frac{1}{HH_{m15+}} \]

Business Selection Probability: \[ P_{BUS} = \frac{q}{BUS_{EA}} \]

\( DW \) = design weight compensation for over-sampling of urban EAs and under-sampling of rural EAs;

\( HH_{EA} \) = number of households in selected EA based on information of last census or updated listing by field team;

\( HH_{STRATA} \) = number of households in strata (urban, rural);

\( HH_{m15+} \) = number of household members or visitors 15 years or older;

\( m \) = target number of EAs for each strata, (urban, rural);

\( n \) = target number of households in EA;

\( q \) = target number of businesses in EA;

The target number of households in each EA varied from country to country. Usually 20 households are selected from each EA and 10 businesses.
GIS Maps

Maps may be available from the census office. There are two formats:

- Hard copy from the census office;
- Shape files (.shp) that are compatible with Geographic Information System software.

Hard copy maps

If the census office will only provide hard copies of maps, they should ideally be digitized. This can be done by scanning the hard copy map and importing into QGIS. This tutorial explains how to digitize a scanned map in more detail. Once the map has been digitized, it can be imported into Google Earth. This tutorial explains how to do that.

Shape files

If the census office provides maps in shape file format, then the shape files need to be imported into QGIS and saved using the correct projection.

When receiving the shape files from the census office, it’s critical to get the projection that the census office has used to create the maps. The projection data will save you a lot of time in the future!

Once the maps have been imported into QGIS, they can be saved as .kml files, the required format for Google Earth or Google maps. This tutorial explains how to convert shape files into .kml files. Remember that the projection for .kml files is WGS84, so when converting the maps to .kml, you also need to change the projection.

When the maps have been imported into Google Earth or Google Maps (Google Earth is preferred because it is easier to manipulate the maps and save the files in your “My Places” folder), then you can zoom down to the selected EA and print the appropriate map. This tutorial explains how to create a polygon in Google Earth, save it to your folder and print it out.

---

2 http://www.igismap.com/convert-shapefile-kml-qgis/
ODK tips and tricks

XLSForm is a standard created to simplify uploading of questionnaires, using Excel, to a data collection platform such as Open Data Kit (ODK). A basic explanation of XLSForm can be found [here](#). Below is a list of shortcuts that make formatting the questionnaires easier:

- Open and close groups immediately
- Name groups with letters and numbers so it is easy to spot open groups, e.g., a, b, c or a1, a2 a3, (numbers aren’t accepted by themselves)
- No spaces in skip formulas
- In the “appearance” column, using the term “field-list” puts all the questions in the group on one page. However, this doesn’t work for sub-groups (i.e., groups within groups). Therefore, try to limit the number of sub-groups.
- Date questions: In the “type” column, if you use the “date” term combined with “year” in the “appearance” column, XLSForm will translate that as a dropdown list. This is convenient for limiting the answer to a year automatically and preventing answers that are nonsensical (such as 8890).
- “relevant” column: formatting as follows: ${name}="1" rather than ../name=1 is more consistent
- Dropdown lists: In the “appearance” column, using the term “minimal” produces dropdown lists, saving space.
- Test the ODK questionnaire: print the hard copy and have it lying next to the mobile while going from question to question.
Example South Africa

Sampling

Two sample frames were available for South Africa one based the 2011 census: one for Enumerator Areas (EAs) and one for Small Area Layers (SAL). RIA has shape files for both. One complication is that EAs are classified into what can be converted into Urban and rural but does not include the number of households per EA.

- The SAL data set can be used to sample SAL based on PPS while for the EA data set only SRS can be used.
- The advantage of EA's is that the listing should be quicker since EAs are mostly smaller in geographic size compared to SALs. There are 103,576 EAs compared to 84,908 SALs.
- Figure 8 shows that EAs which are not covered by SALs. EAs provide thus a complete sample frame

<table>
<thead>
<tr>
<th>Table 4: 2011 Census shape files</th>
<th>SALs</th>
<th>%</th>
<th>EAs</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parks and recreation</td>
<td>266</td>
<td>0,3%</td>
<td>539</td>
<td>0,5%</td>
</tr>
<tr>
<td>Industrial</td>
<td>775</td>
<td>0,9%</td>
<td>1.676</td>
<td>1,6%</td>
</tr>
<tr>
<td>Commercial</td>
<td>789</td>
<td>0,9%</td>
<td>1.481</td>
<td>1,4%</td>
</tr>
<tr>
<td>Collective living quarters</td>
<td>975</td>
<td>1,1%</td>
<td>1.637</td>
<td>1,6%</td>
</tr>
<tr>
<td>Smallholdings</td>
<td>1.039</td>
<td>1,2%</td>
<td>1.216</td>
<td>1,2%</td>
</tr>
<tr>
<td>Vacant</td>
<td>1.719</td>
<td>2,0%</td>
<td>7.818</td>
<td>7,5%</td>
</tr>
<tr>
<td>Farms</td>
<td>3.480</td>
<td>4,1%</td>
<td>5.227</td>
<td>5,0%</td>
</tr>
<tr>
<td>Informal residential</td>
<td>4.768</td>
<td>5,6%</td>
<td>5.313</td>
<td>5,1%</td>
</tr>
<tr>
<td>Traditional residential</td>
<td>28.962</td>
<td>34,1%</td>
<td>30.929</td>
<td>29,9%</td>
</tr>
<tr>
<td>Formal residential</td>
<td>42.134</td>
<td>49,6%</td>
<td>47.740</td>
<td>46,1%</td>
</tr>
<tr>
<td>Total</td>
<td>84.908</td>
<td>100%</td>
<td>103.576</td>
<td>100%</td>
</tr>
</tbody>
</table>

Based on these considerations we gone ahead and sampled EAs using SRS. The steps were as follows:

- Formal residential and informal residential EAs were classified as urban and farms, smallholdings and traditional residential EAs as rural.
- The target sample of 1800 was split 60% urban and 40% rural, yielding a target of 720 rural and 1080 urban households.
- Using a clustering of 24 households per EA translates then into 30 rural and 45 Urban EAs to be sampled.
- EAs were classified into urban and rural and then split into two groups. The sampling is done for each group separately.
EAs are sorted by province and SP number and then an numeric ID is given from 1 to N, with N being the number of EAs in group.

A random starting is determined for each group using the “randbetween” function in MS Excel for values between 1 and number of EAs in group. This random starting point is also the first EA randomly selected.

Next the sampling interval was determined by dividing the Number of EAs in the group by the target sample of EAs (30 rural and 45 urban).

Subsequent selected EAs were determined by adding the sampling interval on each previously selected EA numeric ID.

The random starting point is rounded down to safeguard that the sampling interval method does not overshoot random starting point.

---

<table>
<thead>
<tr>
<th>Table 5: SA EA sampling</th>
<th>Rural</th>
<th>Urban</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample of households</td>
<td>720</td>
<td>1.080</td>
<td>1.800</td>
</tr>
<tr>
<td>HHs per EA</td>
<td>24</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Sample of EAs</td>
<td>30</td>
<td>45</td>
<td>75</td>
</tr>
<tr>
<td>Total EAs</td>
<td>37.372</td>
<td>53.053</td>
<td>90.425</td>
</tr>
<tr>
<td>Sampling Interval (Randbetween 1, N)</td>
<td>1.245,73</td>
<td>1.178,96</td>
<td></td>
</tr>
<tr>
<td>Rounded down for extras</td>
<td>1245</td>
<td>1178</td>
<td></td>
</tr>
</tbody>
</table>

---

Figure 8: Red shows Has which are not covered by SALs
2 extra EAs for rural and 2 for urban areas were selected using same methodology in case is a need for a random replacement.

Using SRS instead of PPS has an impact on the construction of the weights. EA Selection Probability: \( P_{EA} = \frac{m}{EA_{STRATA}} \), \( m = \) target number of EAs for each strata.

**Listing**

First screenshot is a bird eye view from the Google Earth placing the EA in context to major landmarks, in this case Promenade Mall in Mitchell's Plain.

Next is a more detailed map of the EA. This screenshot comes from a Carto DB GIS Map provided by iKapaData. Given the structure of the EA we allocated blocks to listers marked by letters.
As an additional help we also included Google Street view pictures within the EA, always looking into the EA. The pictures are numbered to allow easy orientation.
## Definitions and Resources

<table>
<thead>
<tr>
<th><strong>Table X: Definitions and resources</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EAs</strong></td>
</tr>
<tr>
<td><strong>PPS</strong></td>
</tr>
<tr>
<td><strong>SRS</strong></td>
</tr>
<tr>
<td><strong>ODK</strong></td>
</tr>
<tr>
<td><strong>Census</strong></td>
</tr>
<tr>
<td><strong>ONA</strong></td>
</tr>
<tr>
<td><strong>XLSForm</strong></td>
</tr>
<tr>
<td><strong>ODK collect</strong></td>
</tr>
<tr>
<td><strong>ONA collect</strong></td>
</tr>
<tr>
<td><strong>Household</strong></td>
</tr>
<tr>
<td><strong>Head of household</strong></td>
</tr>
<tr>
<td><strong>Member of a household</strong></td>
</tr>
<tr>
<td><strong>Eligible Individual</strong></td>
</tr>
<tr>
<td><strong>Businesses</strong></td>
</tr>
</tbody>
</table>
References


