



ICT HH and Business Survey Field Manual

Prepared By

Research ICT Africa

April 2018

Contact: Dr O. Mothobi - omothobi@researchictafrica.net or
Dr A. Gillwald - agillwald@researchictafrica.net

Acknowledgements

This field handbook was prepared by past 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 Dr. 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 Schenk) 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 an Global South #afteraccess 18-country #afteraccess survey undertaken by RIA together with LIRNEasia and DIRSI available on www.afteraccess.net. The 'After Access' ICT access and use surveys are made possible with the support of the Canadian International Development Research Centre (IDRC).

Table of Contents

Acknowledgements	2
Country Mission	4
Check list: What to bring	4
Arrival in Country - at the airport	4
Enumerators	4
Letter of Introduction	6
Preparing the data collection devices	6
Questionnaire Training	9
Quality Control and Data Management	10
Quality Control	12
Data Management	14
Sampling	15
Sample Size	16
Weighting	16
GIS Maps	18
Hard copy maps	18
Shape files	18
ODK tips and tricks	19
Example South Africa	20
Example Uganda	24
Past Field Experience	25
Definitions and Resources	27
References	28

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

The Enumerator database can be found here: [Dropbox/Research ICT Africa/13-Latest HH survey data/Enumerator hh 2017.fmp12](#).



Name	Position	ID	Survey	Survey Period	Mobile	Languages	Driver's Lic...	Region	Country
First and Lastname	Enumerator	E001	Beyond Access Survey 2017	Jun-Aug 2017					South Africa

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.

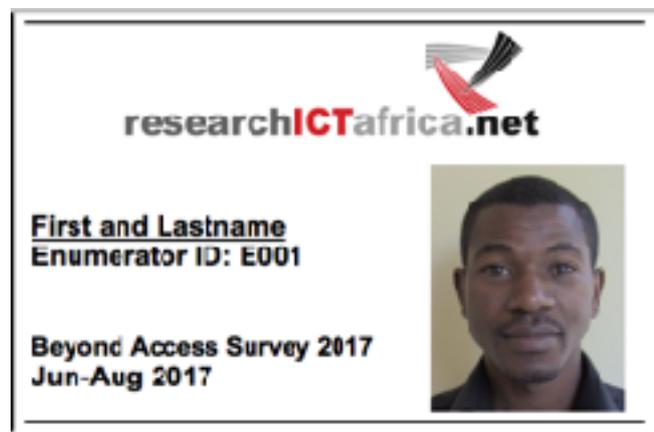


Figure 2: Badge with only RIA logo

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.

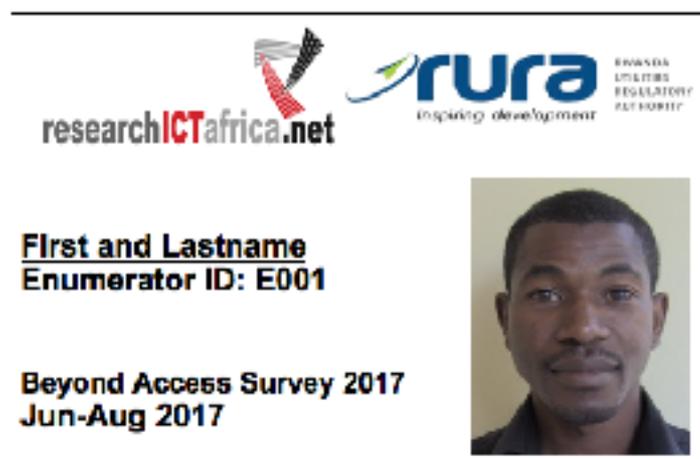


Figure 3: Batch with logos from RIA and Partner organisation

- 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

Data will be collected using surveyCTO platform, which allows researchers to collect data using mobile phones, tablets or computers. The platform is based on an Open Data Kit (ODK). The following devices are critical to the data collection processes and should be checked throughly:

- Tablets/mobile phone
- Extra battery
- Power bank

SurveyCTO Collect

SurveyCTO collect is one of the two options for collecting data on the SurveyCTO platform, the other is webform. It is a mobile application with an Android application system. SurveyCTO collect offers full offline data collection operations if required.

Installing SurveyCTO Collect

To download the app onto a device or fill out a form using your web browser, go to the following web page:

ria.surveycto.com/collect



Figure 4: SurveyCTO Download link screenshot

tap on the download link and launch the installer from the notification, work through the prompt and open the app. Configure some survey setting from general settings.



Figure 5: SurveyCTO General setting tab screenshot

Add a server name, username and password and you have now downloaded surveyCTO collect.

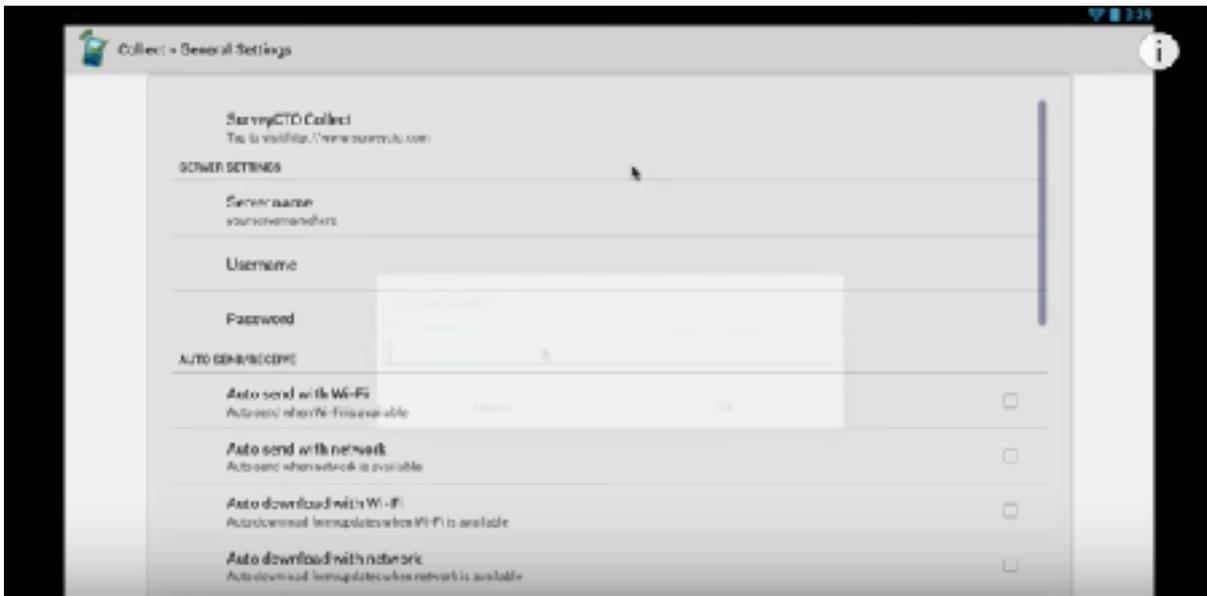


Figure 6: SurveyCTO General setting screenshot

Then get a blank form to access the household & individual and business online questionnaires

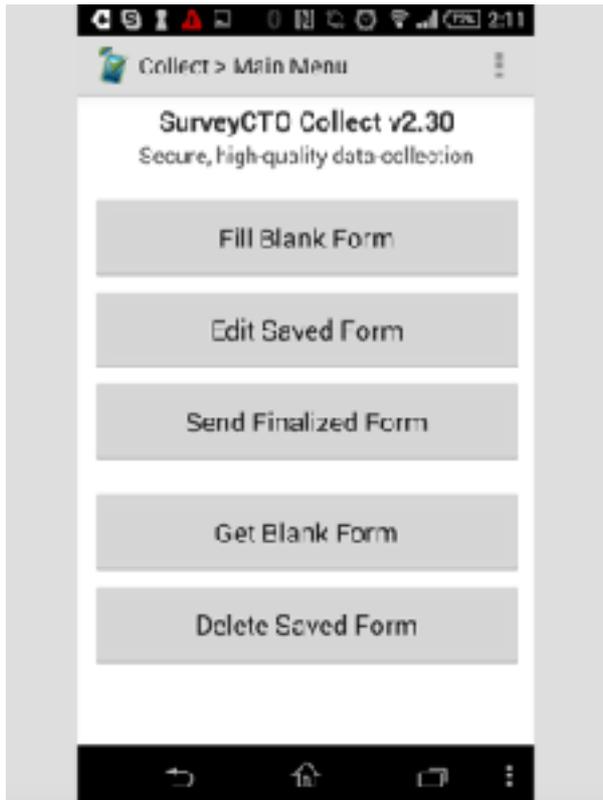


Figure 6: SurveyCTO Forms menu screenshot

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@researchictsolutions.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>

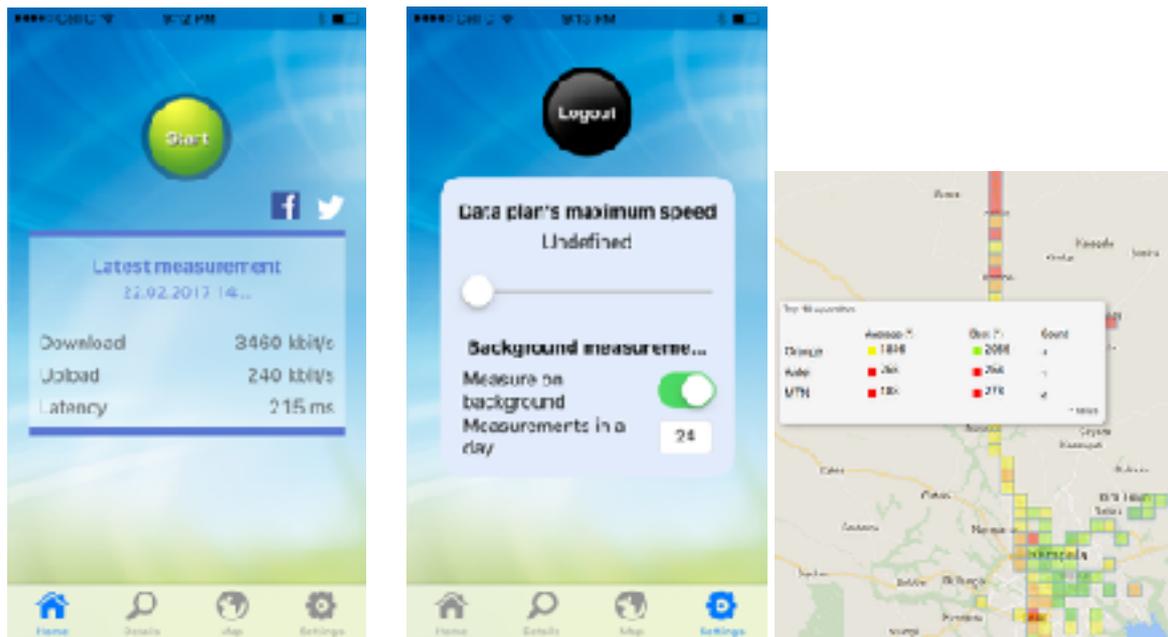


Figure 5: Netradar example output Main, preference screens and example output

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 in 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.

Quality Control and Data Management

Listing

In the 2017 surveys, RIA introduced a digital listing solution to replace the previous paper-based listing method. The digital listing involves the capturing of every household and business in an EA using a mobile form. The fieldworkers create a new form instance for every structure and then identify every household and business in that structure by using names, house numbers, visual cues and the GPS location of the structure.

The use of GPS for the listings is crucial as team leaders are notified on Slack (a company-wide messaging app) by a backend system. If a fieldworker did a listing outside an EA, the fieldworkers are warned directly on the device if they tried to do a listing outside the EA.



Figure 1: Listing outside EA

The listings are submitted in real time and visualized almost instantly on an online map accessible to the in-office quality control staff as well as the team leaders on the ground. Once a team leader felt that they were done with the listing of an EA, she would contact the quality control staff via Slack. After checking the map, quality control would either ask the team to go back and complete the listing by pointing out gaps on the map, or confirm that the listing was completed.

Once the listing is completed a Stata command developed for RIA's sampling requirements will be run to produce two printable lists of households and businesses with addresses, names and descriptions, two CSV files with the same information and GPS coordinates (for uploading to Google Maps) and another CSV file with additional variables (for uploading to SurveyCTO as a preload). QC staff will then send a printable lists to the team leader via Slack, to assign households and businesses to the individual fieldworkers. The fieldworkers had access to the Google Map with the listing on their devices and could tap on a household on the map to view the household/business information and receive directions to the location of the household/business.

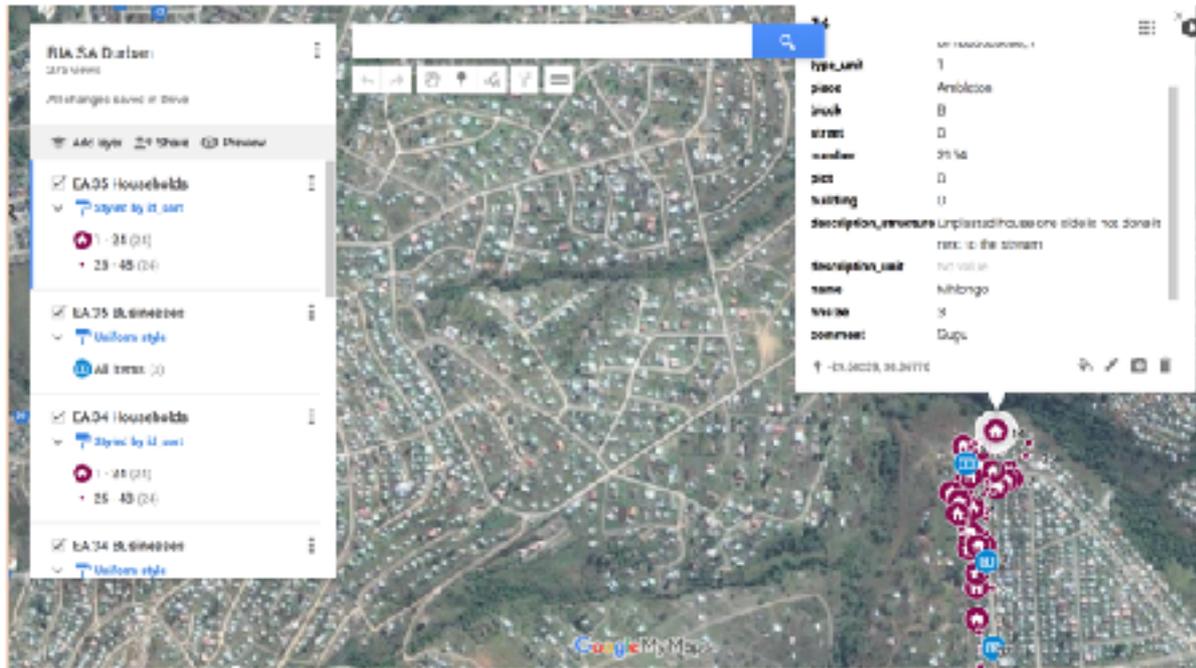


Figure 2: Google Map for Fieldworkers

At the household/business a fieldworker can compare the information they received from the respondent with the information captured during the listing straight in the household/business mobile form. The form also warns if the location is more than 20m away from the location where the listing of the household/business took place.

Quality Control

Quality control consists of the employment of automatised data checks by the backend combined with automatised messaging, regular checks by dedicated quality control using the online dashboard and live dataset as well as back-check calls.

The backend flags “suspicious” submissions and notifies team leaders and QC via Slack based on a series of criteria, including duration, GPS accuracy, distance from listed household/business, duplicate ID/contact number as well sampling inaccuracies.

Additionally, every hour the team leaders receives a progress report for the day, including quality control indicator for every team member.

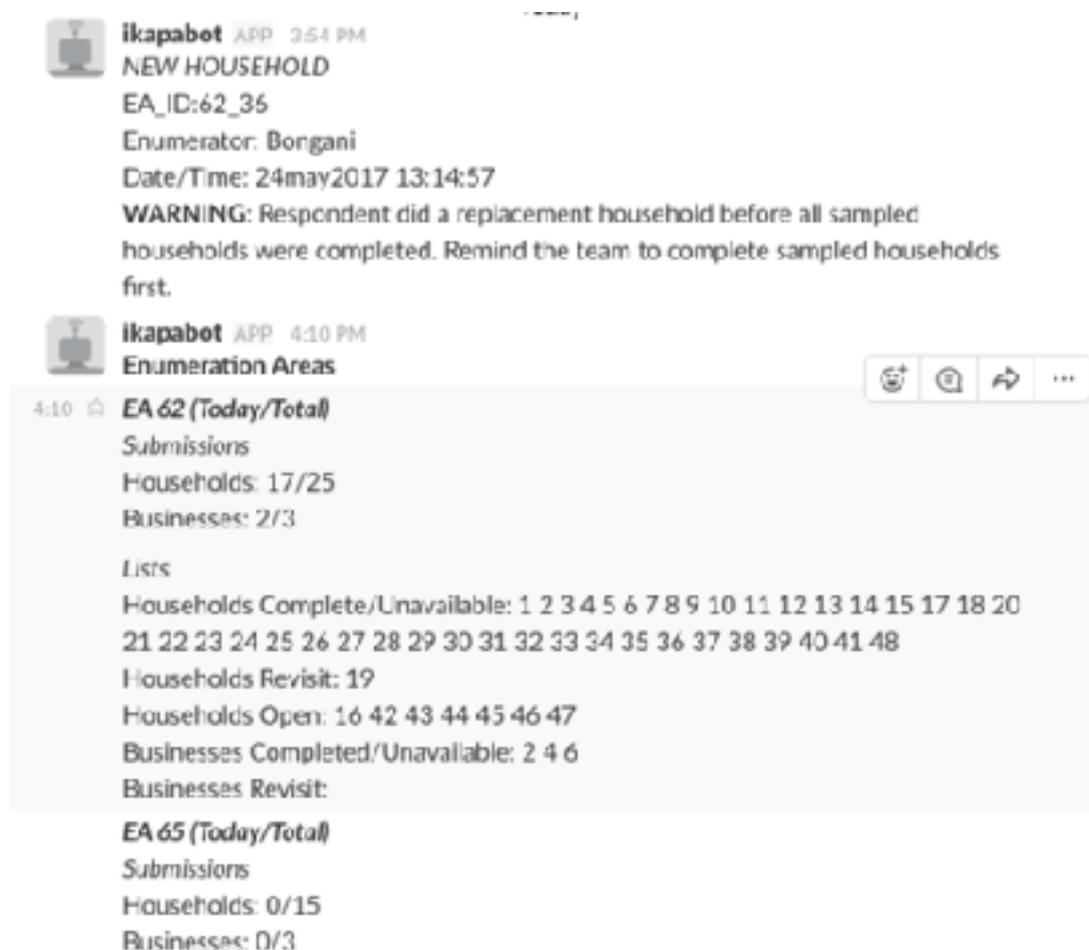


Figure 3: Progress Report for Team Leader

The online dashboard gave all relevant stakeholders access to the main dashboard showing an online map with all interviewed households, graphs depicting survey progress and indicators of data quality. The team leaders also had access to their team-specific dashboards where all indicators were broken down by team members.



Figure 4: Online Dashboard

If necessary, QC would contact the team leader to enquire about irregularities in the data, and the team leader would talk to the fieldworker in question and sometimes accompany them to an interview to ensure that they adhere to the fieldwork protocols.

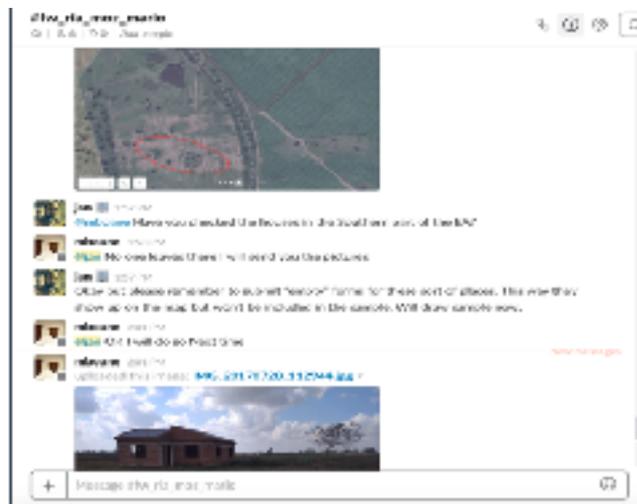


Figure 5: Field Communication via Slack

Data Management

All interview submissions are sent via the cellphone network after each interview (meaning that the data would still be available even if a tablet got lost or stolen) using Secure Sockets Layer (SSL) technology and stored on a server hosted by the provider of the mobile data collection platform, SurveyCTO. The data can be encrypted by a user-generated 2048-bit encryption key, meaning that not even SurveyCTO's engineers can see the data. The server itself provides the kind of data security that satisfies even the most sophisticated and demanding institutional review boards (IRBs).

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.

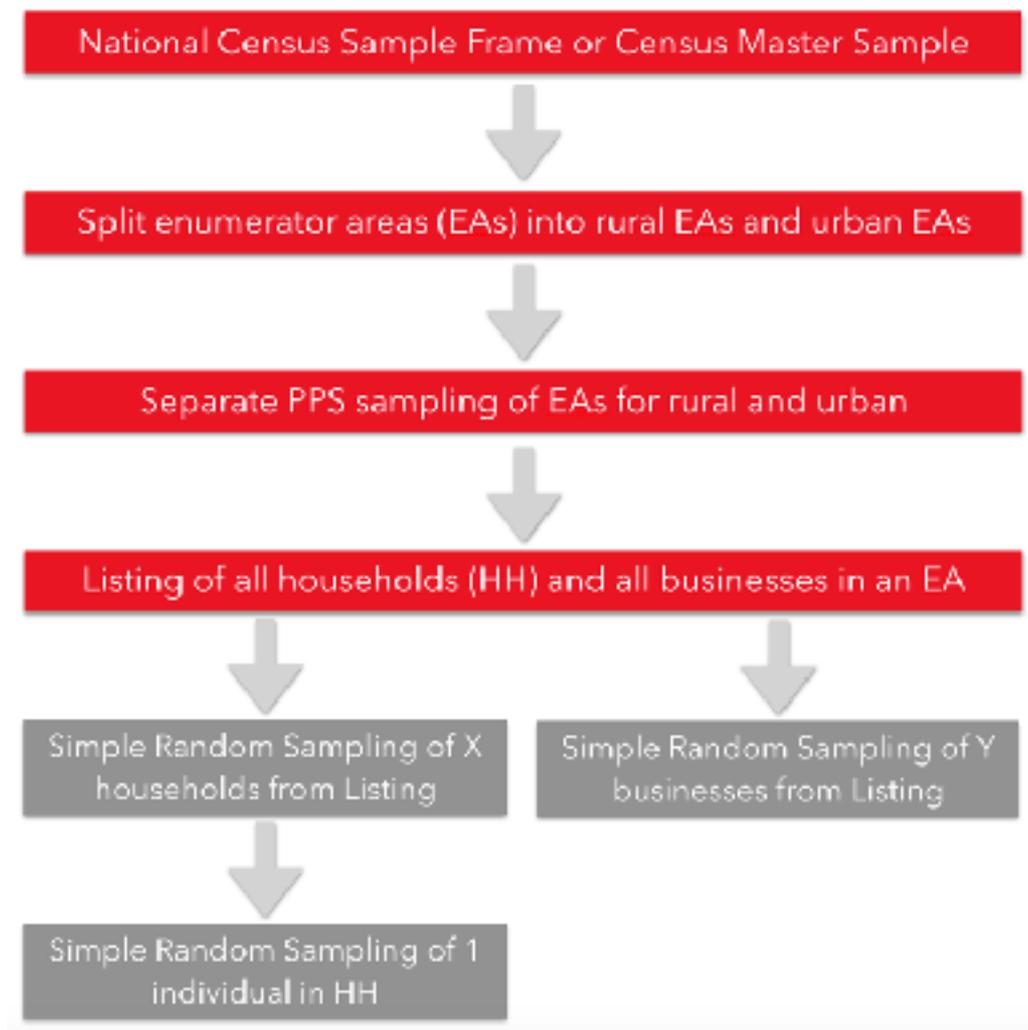


Figure 7: Sampling steps

- 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 3: Minimum sample size for variables expressed as proportions for large populations		
Margin of error	95%	99%
+1	9.604	16.577
+2	2.401	4.145
+3	1.068	1.842
+4	601	1.037
+5	385	664
+6	267	461
+7	196	339
+8	151	260
+9	119	205
+10	97	166
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 probabilities¹ and gross up the data to national level when applied.

$$\text{Household weight: } HH_w = DW \frac{1}{P_{HH} * P_{EA}}$$

$$\text{Individual weight: } IND_w = DW \frac{1}{P_{HH} * P_{EA} * P_I}$$

$$\text{Business Weight: } Bus_w = DW \frac{1}{P_{Bus} * P_{EA}}$$

¹ See UNSD (2005) page 119 for a detailed discussion on sampling weights.

Household Selection Probability: $P_{HH} = \frac{n}{HH_{EA}}$

EA Selection Probability: $P_{EA} = m \frac{HH_{EA}}{HH_{STRATA}}$

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](#)² 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](#)³ 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/>

³ https://ace.arsc.edu/system/files/Creating%20a%20polygon%20in%20Google%20Earth_0.pdf

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

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

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.

Table 5: SA EA sampling	Rural	Urban	Total
Sample of households	720	1.080	1.800
HHs per EA	24	24	24
Sample of EAs	30	45	75

Table 5: SA EA sampling	Rural	Urban	Total
Total EAs	37.372	53.053	90.425
Sampling Interval (Randbetween 1, N)	1.245,73	1.178,96	
Rounded down for extras	1245	1178	

- 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.

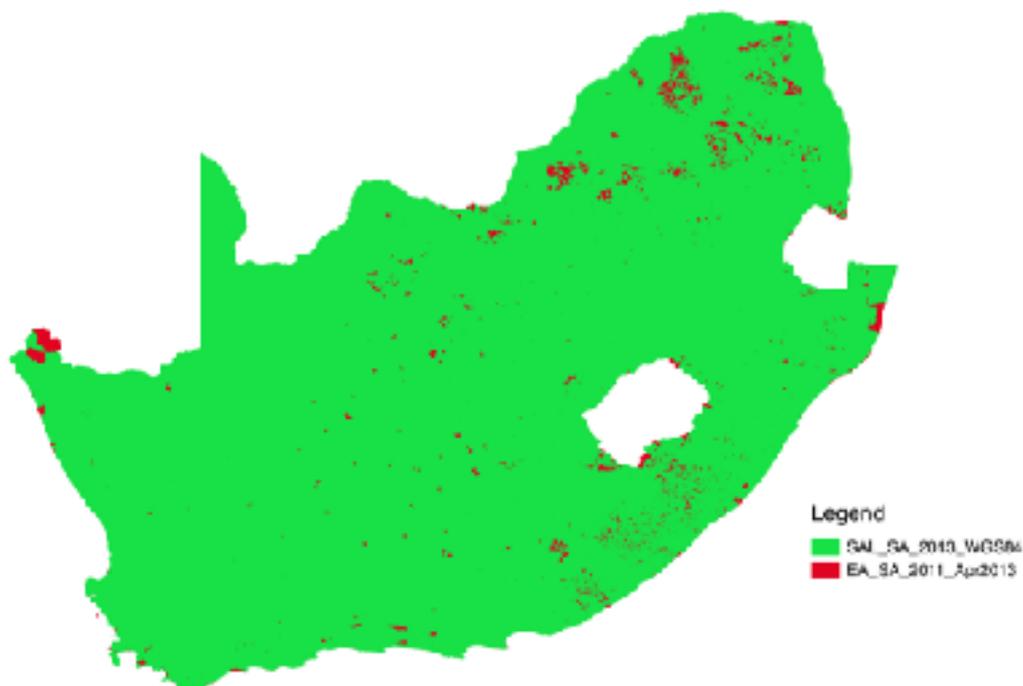


Figure 8: Red shows Has which are not covered by SALs

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

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.



Example Uganda

The Uganda #AfterAccess Survey will covers 29 districts out of 121 districts. A total of 64 EAs across these districts were selected based on the sampling frame. The sampling process was done by Uganda Bereau of Statistics (UBOS).

Table 6: Uganda EA sampling	Rural	Urban	Total
Sample of EAs	38	36	64
Hosehold per EA	24	24	1.776
Business per EA	10	10	640

The Uganda #AfterAccess Survey will covers 29 districts out of 121 districts. A total of 64 EAs across these districts were selected based on the sampling frame. The sampling process was done by Uganda Bereau of Statistics (UBOS).

Past Field Experience

In 2017 Research ICT Africa commissioned Ikapadata (Pty) Ltd to conduct #AfterAccess Surveys in Mozambique and South Africa. The South African survey ran from the 15th of May until 11th of August. Fieldwork for Mozambican survey was conducted between the 20th of July and the 6th of October 2017. For the Mozambican survey Ikapadata (Pty) Ltd only provided fieldwork management and quality control. The following section provides the challenges experienced in the field:

Fieldwork was scheduled to last 10 weeks but in the end took 14 weeks between the 15th of May and 21st of August 2017. The main reason for the delay was the longer-than-anticipated average time spent in each EA. The original fieldwork schedule and budget was based on the assumption that a team of four fieldworkers and one supervisor would do the full listing, 24 household interviews and approximately 6 business interviews in an EA over the course of two days. Using the timestamps associated with each listing and interview we now know that the average EA takes three to four days to complete, which has significant consequences for fieldwork planning and costs.

The other major challenge having an impact on the fieldwork schedule was the lack of businesses in the sampled EAs. Based on a response rate of 67%, the fieldworkers would have needed to list at least 10 businesses per EA, on average, in order to meet the target of 500 valid business interviews. In reality, the enumerators only counted 6.6 businesses per EA, with a few EAs exceeding the required minimum by far and many others falling short of it or not having any businesses at all. As a result, the teams had to revisit some of the EAs with businesses that were unavailable the first time around (e.g. because they were closed at the time of the first round of interviews).

Another factor that had a negative impact on survey progress was the amount of racism encountered by the fieldworkers. This is not a new phenomenon in predominantly white and wealthy neighborhoods and was taken into account during recruitment and assignment of EAs to teams, but it has now deteriorated to a level where fieldworker safety has become a real concern and response rates in better-off neighborhoods have fallen to extremely low levels. Future iterations of the survey should probably take this into account as one needs to consider using alternative survey modes to compensate for the lack of white, better-off respondents willing to participate in the survey.

Mozambique

In Mozambique the most significant issues stemmed from mistakes and omissions in the mobile forms. In South Africa, one question about the reason why a household would not have an internet connection did not show up due to an error in the skip logic until after half of the survey was done, and ikapadata had to call all respondents who had not been asked question although they should have. Another question about the respondent's education was erroneously dropped from the questionnaire before the survey because of a misunderstanding between the service provider and client. Other examples include instances of questions where the question type differs between the questionnaire version used in South Africa and Mozambique compared to the version used in other countries. Also, the variable and value labels differed between the two versions, and had to be changed afterwards.

Digital Listing

In Mozambique we discovered a grave case of erroneous fieldworker behavior during listing. Some fieldworkers marked households or business which were unavailable (i.e. nobody at home or business closed at time of listing) as “empty” which was an option reserved for structures that were truly empty (e.g. ruins or unfinished buildings). This happened although the use of the “empty” option was explicitly discussed during both rounds of training. Also, this sort of mistake did not occur during the pilot and only started to creep in gradually as the survey progressed, making it difficult to detect. We managed to contain the possible selection bias introduced by this behavior (basically households or businesses categorized as empty never appeared on the sample lists) by identifying the wrongly categorized households and businesses in the data and even revisiting many of them in Maputo. Nonetheless, the question must be asked to what extent digital listing enabled this sort of behavior, and why we included an empty category in the first place. The answer to the latter is that we simply copied the options from the paper version and it did make sense to have it digitally, as it allowed fieldworkers to GPS mark empty structures on the map so our quality control staff could see that they did not forget about a section of the EA as otherwise empty structures would have appeared as gaps on the map.

But given the experience we recommend dropping the “empty” option from the listing form and let team leaders and quality control staff communicate to explain gaps on the map as they occur.

Despite this issue, we strongly recommend using digital listing going forward for the following reasons:

It makes the process transparent and mistakes can be tracked and corrected (one could even argue that the same mistake documented above could also have been made on paper, but in that case it is unlikely that anyone would have ever noticed).

It makes it easier for the fieldworkers on the ground as they can be more flexible in the way they mark structures, or even do not mark them at all, as the use of GPS combined with useful descriptors are usually enough to identify households and businesses for later visits. Also, the randomisation to be done in-field, on paper, by the team leaders can be argued to be prone to mistakes and manipulation, whereas the randomisation done by the backend always results in a true simple random sample.

It eliminates the risk of teams listing EAs in the wrong location or incompletely. There have been many instances where we noticed that teams would start listing outside the designated EA, sometimes kilometres away from the actual EA, and only the automatised notifications from our backend or on their devices made them aware of it (usually followed up by an intervention of our QC staff to direct them to their actual EA). Also, teams would often claim that they finished the listing of an EA, but on further inspection of our QC staff it was clear that parts of the EA were still missing. This sort of control and transparency is simply impossible using pen and paper.

Definitions and Resources

Table 3: Definitions and resources	
EAs	An enumeration area (EA) is the smallest geographical unit during census enumeration. It is the geographic area and cover all the territory of a country. EAs are used for census data collection and are not political boundaries. The size of an EAs typically varies between 100 and 250 households but varies with terrain and other topological conditions and population density.
PPS	Probability proportional to size (PPS) sampling
SRS	Simple random sampling: the most widely-used probability sampling method, probably because it is easy to implement and easy to analyse.
ODK	Open Data Kit (ODK) is a free and open-source set of tools which help organisations author, field, and manage mobile data collection solutions on the Android platform: https://opendatakit.org also https://youtu.be/HqqUdfz9Uyc
Census	A population census is the process of counting the number of people, at a given point in time in a country, and collecting information about their demographic, social and economic characteristics.
ONA	ONA is a service provider that uses ODK: https://ona.io/features.html
XLSForm	Excel sheet (.xls not .xlsx) that can be uploaded to ONA and translates to a questionnaire on Android phone for be used with ODK collect or ONA collect.
ODK collect	Android app that requires username, password and URL if using ONA
ONA collect	Android app that requires username and password if using ONA. Otherwise the same as ODK collect.
Household	Constitutes a person or group of persons, irrespective of whether related or not, who normally live together in the same housing unit or group of housing units and have common cooking arrangements.
Head of household	A head of a household is a person who economically supports or manages the household or, for reasons of age or respect, is considered as head by members of the household, or declares himself as head of a household. The head of a household could be male or female.
Member of a household	All persons who lived and ate with the household for at least six months including those who were not within the household at the time of the survey and were expected to be absent from the household for less than six months. All guests and visitors who ate and stayed with the household for six months and more. Housemaids, guards, baby-sitters, etc. who lived and ate with the household even for less than six months.
Eligible Individual	Any household member 15 years or older that will sleep the night at the house and also any visitor 15 years or older that is going to spend the night at the house.
Businesses	Any business with a physical presence in the EA with the intent to make profit

References

Rea, L and Parker, R. (2014), *Designing and Conducting Survey Research: A Comprehensive Guide*.

Lwanga, S. and Lemeshow, S. (1991), *Sample Size Determination in Health Studies – A Practical Manual*, World Health Organisation, Geneva.

Thompson, S. (2002):, *Sampling, Second Edition*, Wiley Series in Probability and Statistics.

UNSD (2005), *Designing Household Surveys Samples: Practical Guidelines*, United Nations, New York. <https://unstats.un.org/unsd/demographic/sources/surveys/Handbook23June05.pdf>.

UNSD (2008), *Designing Household Surveys Samples: Practical Guidelines*, United Nations, New York. https://unstats.un.org/unsd/demographic/sources/surveys/Series_F98en.pdf.