



Smartphone Data Collection

Kaduna State Bureau of Statistics

Jonathan Phillips, Harvard University



Objectives



- Introduce New Data Collection Methodologies
- Consider Rapid Sampling Techniques
- Discuss Logistics and Procedures
- Prepare Field Staff for Forthcoming Surveys
- Give you the skills to train others



Kaduna Data Revolution



- A Data Revolution based on **Useful, Local and Open** Data
 - **Useful:** Data must be rapidly collected, in line with policy needs, and accurate
 - **Local:** Not just state averages, but disaggregated data to specific communities, villages, wards and LGAs
 - **Open:** Data should not be hidden in hard copy files, but publicly available for verification



Provisional Plans



- Initiate the Kaduna State Statistical Masterplan
 - Plug in to existing data collection efforts: DHIS, EMIS, BATMIS, HRIS etc.
- Maximize existing data
 - Access and analyse existing datasources from other agencies, satellite and spatial sources
- Develop capabilities and skills
 - Including an online databank of information
- Generate new data to fill the data gap
 - Using new electronic survey methods



Recent Example

- Nigeria's MDG Information System
nmis.mdgs.gov.ng





Sampling



- The Population: Everybody we want to know about
- The Sample: The group of people within the population we actually get data from
- Goals for a Sample:
 - Representative of the population
 - Cost-effective
- In general, this means we need a ***random*** sample



Sample Size Calculations



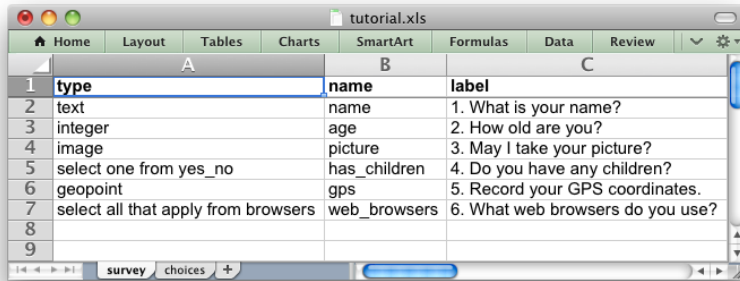
- Samples always have uncertainty in the data
 - Our goal is to minimize that uncertainty
- Law of Large Numbers
 - After gathering data on a few hundred people, our data does *not* get much more accurate
 - Larger sample sizes only needed if we want to disaggregate by gender, region etc.



Sampling Techniques



- List Sampling
 - We already have a list of the population and randomly pick a sample from it
- Spatial Sampling
 - We pick a random starting point from the map and then a random household
- Stratified Sampling
 - Make sure we have similar number of interviews in each Local Government so we have enough data
- Cluster Sampling
 - Ask Interviews of people close together, to minimize cos



	A	B	C
1	type	name	label
2	text	name	1. What is your name?
3	integer	age	2. How old are you?
4	image	picture	3. May I take your picture?
5	select one from yes_no	has_children	4. Do you have any children?
6	geopoint	gps	5. Record your GPS coordinates.
7	select all that apply from browsers	web_browsers	6. What web browsers do you use?
8			
9			

Prepare surveys

Smart Survey Data Collection Framework



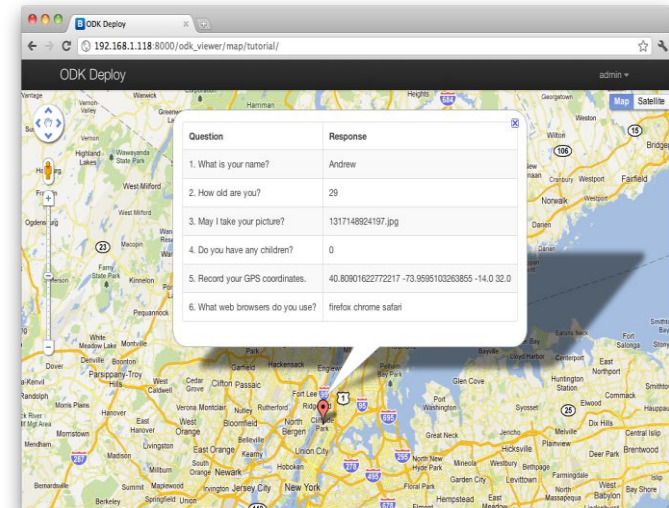
ODK Collect > tutorial

1. What is your name?

Andrew



Collect data on smartphone

Question	Response
1. What is your name?	Andrew
2. How old are you?	29
3. May I take your picture?	1317148924197.jpg
4. Do you have any children?	0
5. Record your GPS coordinates.	40.8090162272217 -73.9995103263855 -14.0 32.0
6. What web browsers do you use?	firefox chrome safari

View / download / analyze
data



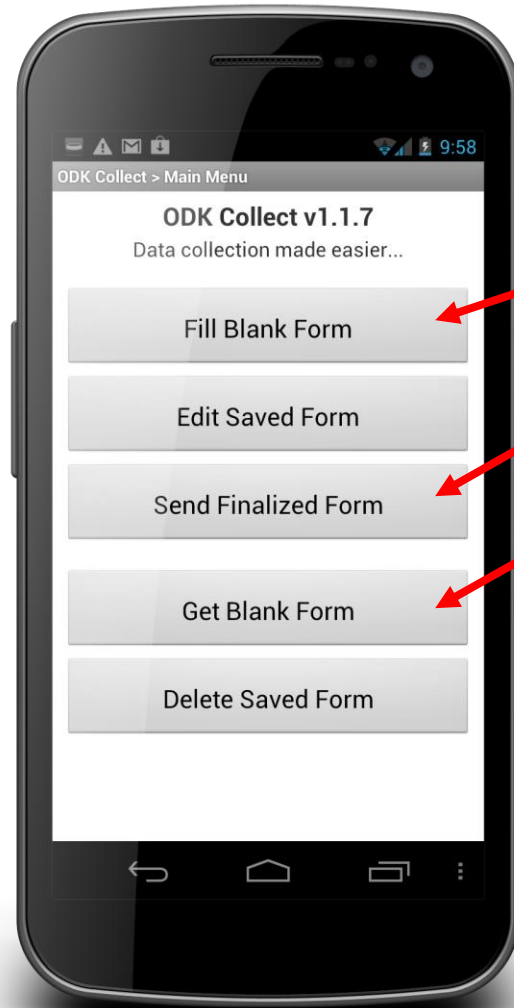
Smartphone Data Collection



1. **More reliable data collection** – errors are caught on entry (eg. age=200) and no need for manual data entry from paper copies
2. **More reliable data storage** – data immediately sent back to online server
3. **New functions** – quickly collect photos, GPS etc. with one device
4. **Real-time Analysis** – Patterns and challenges in the field can be immediately noticed
5. **Efficient supervision and support** – progress in surveys and location can be monitored remotely



ODK Main Menu



- Start a New Survey
- Send Surveys to Server
- Download new survey template (only need to do this once)



ODK Filling Surveys



1. What is your name?

Andrew



- ❑ Swipe left for the next question
- ❑ Enter the data using touch or keyboard



ODK Question Types



- Single Choice (select_one)
- Multiple Choice (select_multiple)
- Text Input (text)
- Number input (integer)
- Picture (image)
- Audio Recording (audio)
- GPS point (geopoint)
- GPS area (geoshape)
- Many others...



Sending Data



- Auto-send using mobile networks
 - Where they are available and you have the right SIM
- Manually send on mobile networks at the end of the day
 - Choose `Send Finalized Form`
 - `Toggle all`
 - `Send selected`
- Backup to a computer – transfer through USB
- Return the Smartphone and download from the memory card



Producing Surveys



- OPTION 1: Kobotoolbox
 - Create an account
 - Click `Add form`, `Start from Scratch`
 - Click `+`
 - Click `+Add Question`
 - Choose question type
 - Type the question name
 - Click `+` to add more questions



Producing Surveys



- OPTION 2: Write surveys in excel
 - One sheet for Questions - `survey`
 - One sheet for choices in multi-choice questions - `choices`
 - List each question as a new row
 - Define question type, unique name and how it will appear on the phone
 - Additional columns add advanced features
 - Upload completed excel sheet to kobotoolbox
 - Best to start from an existing survey template



Additional columns



- **Hint** – extra guidance on answering the question
- **Required** – Can't continue until you pick an answer
- **Relevant** – Only ask a question depending on a previous answer, like skip logic
- **Constraint** – Allow only certain types of answers
- **Label::language** – Alternative languages
- **Calculation** – perform a calculation and show it



Managing Surveys



- How do we get a survey from kobotoolbox to the smartphones?
 1. 'Deploy' your survey from kobo – find your survey and click 'Deploy form as survey project'
 2. Click 'How to collect data on mobile device?' – copy the link, eg. <https://kc.kobotoolbox.org/jonathanphillips>
 3. On the phone, open ODK Collect, go to 'settings', , and type this link into the phone
 4. In ODK Collect, click 'Get blank Survey', select your survey, and 'get selected'



Receiving Data



- All the surveys you conduct get sent back to a server (a computer)
- How do we collate and analyse this data?
- We can look at it online in kobotoolbox
 - It's a spreadsheet so every survey is a row
 - No data entry!
- We can also download it
 - And then map it!
 - In QGIS, google earth etc.