



**SAINT LUCIA**



**World Bank**



## **Session IVa. Use of Technology for Data Capture**

# **Survey Solutions For Census 2020**

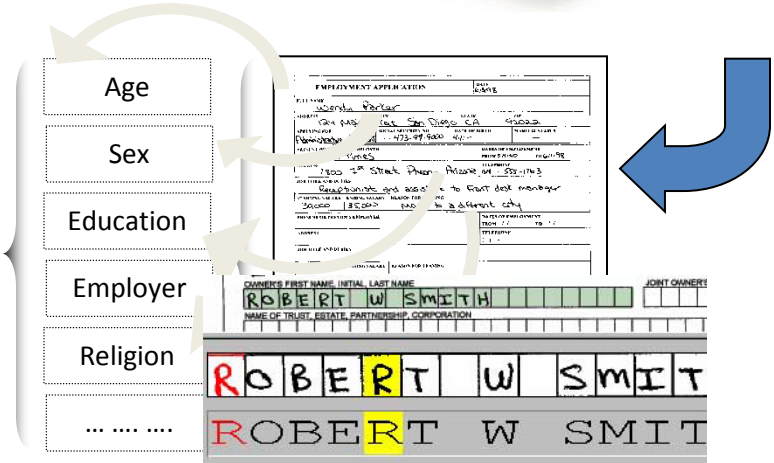
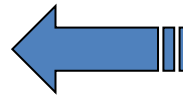
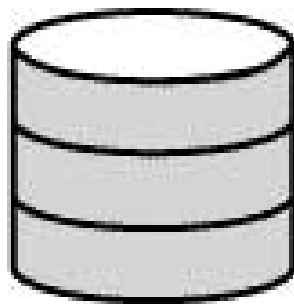
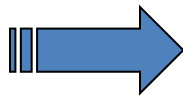
**COMPUTER-ASSISTED PERSONAL and WEB INTERVIEWING**

*Edwin St Catherine, Director of Statistics, Saint Lucia*



# FROM PAPER TO TABLET DATA COLLECTION

## Automated Paper Processing and Business Process

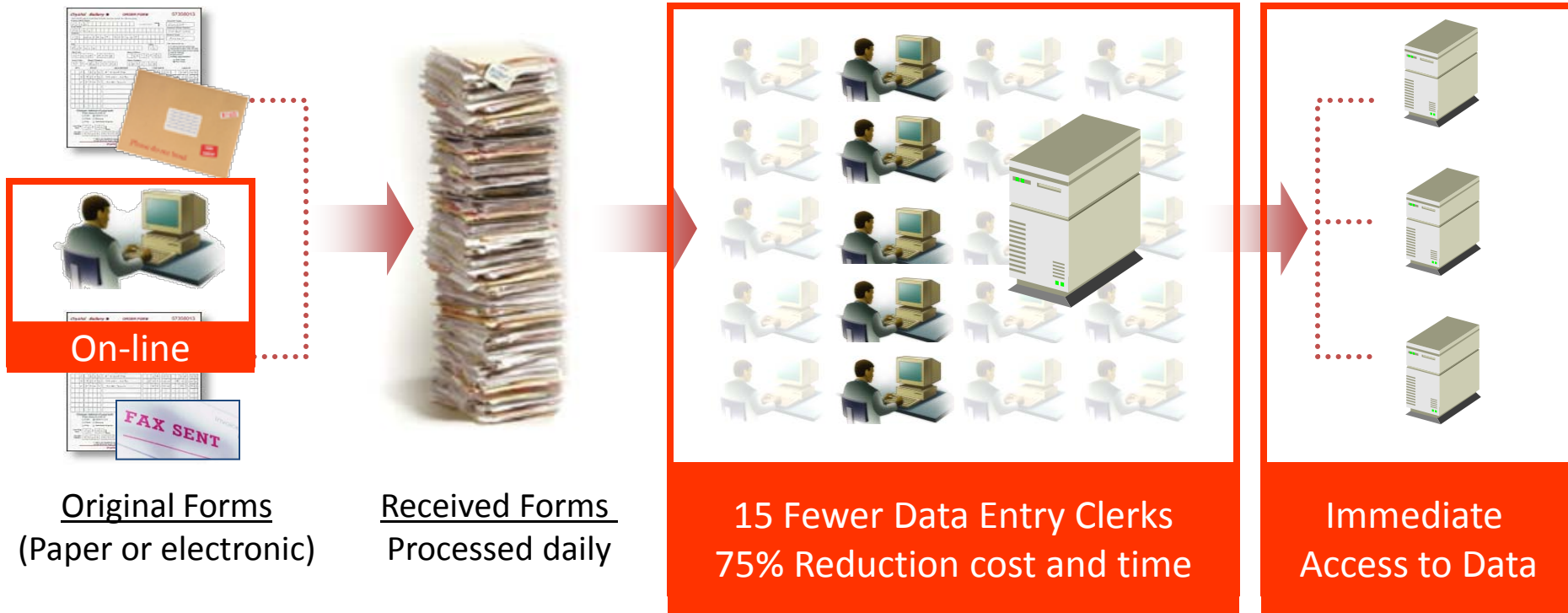




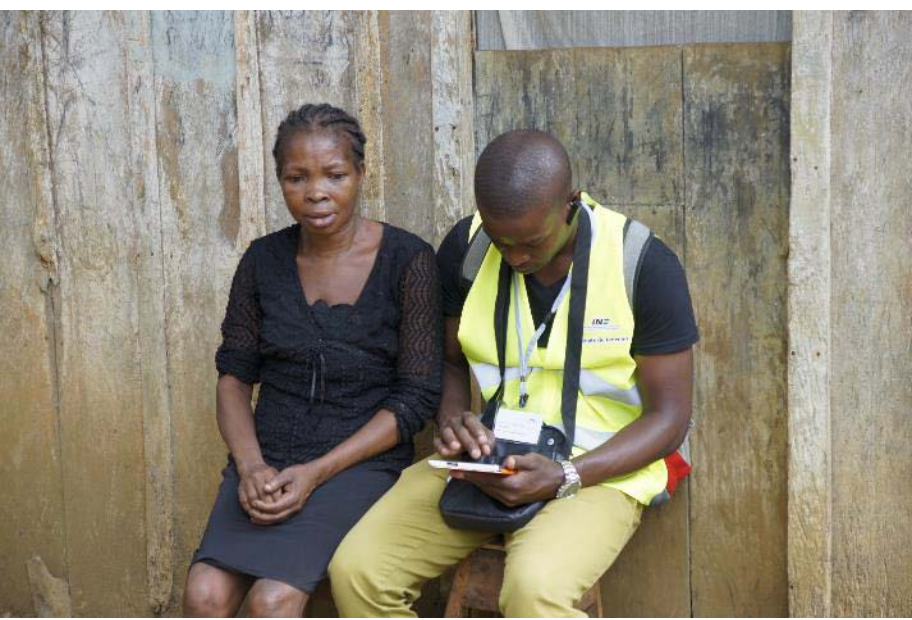
# Where's the Improvement?

## Same Data Entry Scenario – With Scanning Software

Automatically converts census forms to digital information, stores and index the form in dynamically created folders by Parish and ED, and sends the data to enterprise systems









# Benefits of CAPI



**Improve timeliness of data collection**



**Ensure data quality and comparability**



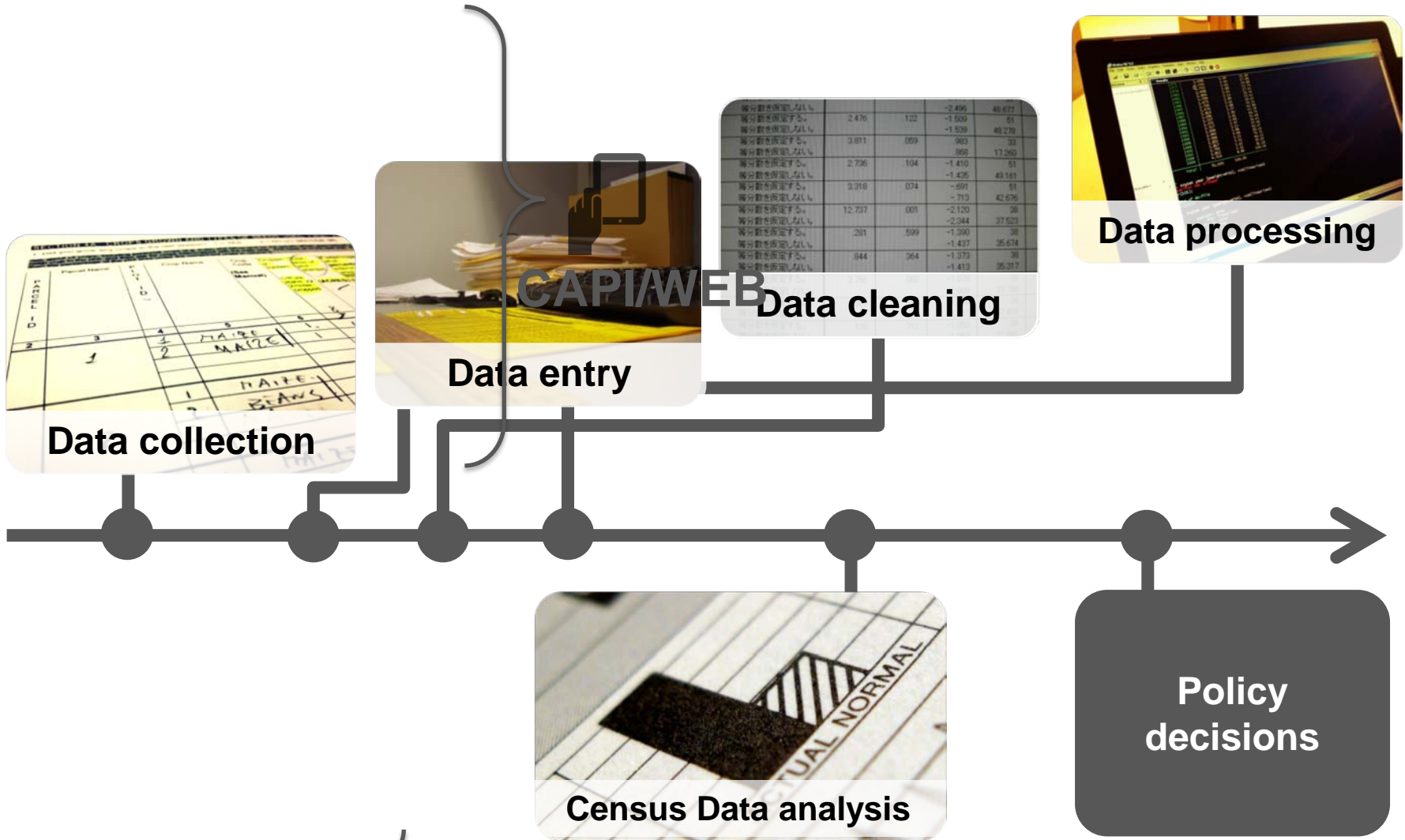
**Allow collection of new types of information/data**



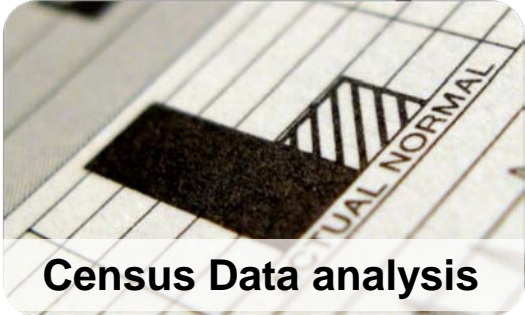
**Cost-effective, sustainable solution for NSOs**



# Results:



等分數を指定しない			-2,496	48,877
等分數を指定する	2,476	122	-1,529	81
等分數を指定しない			-1,529	48,279
等分數を指定する	3,811	859	983	33
等分數を指定しない			983	17,260
等分數を指定する	2,736	104	-1,410	51
等分數を指定しない			-1,410	49,161
等分數を指定する	3,318	874	-991	81
等分數を指定しない			-713	42,676
等分數を指定する	12,737	881	-2,120	38
等分數を指定しない			-2,344	37,523
等分數を指定する	201	599	-1,290	38
等分數を指定しない			-1,437	35,674
等分數を指定する	944	364	-1,373	38
等分數を指定しない			-1,413	35,217



**From data collection to policy: 4-5 months**

# CAPI System Requirements

- Simple yet flexible system for the non-expert users. Typical clients – National Stat Offices
- Functionality for
  - **data capturing**: entering data on a tablet
  - **census management**: managing teams of enumerators
  - **data management**: data aggregation, versioning, reporting
- Tablet-based with ability to display and navigate through multi-level large questionnaires.
- Support of complex validation algorithms.
- Cost effective system that can be used and supported by NSOs without external TA.

# Survey Solutions: hybrid approach

- Sustainable, low-cost system for NSO
- Simple, flexible interface for questionnaire development and testing.
- Tablet interface allows easy navigation through complex questionnaires.
- Standardized survey management protocol based on the best practices of data collection
- Intuitive, informative survey status reporting, survey maps.
- Yet, powerful language for data validation and control of questionnaire flow.



# Main differences from other systems

- **Out-of-the-box solution for survey data collection:** data capturing, *data management, and survey management*. No software on the market provides such a package. All other system focus mostly on data capturing.
- **Minimum TA; Focus on Capacity Building:** Survey Solutions is designed to minimize the TA. Lowest learning curve. Other systems require constant and significant TA. Expert versus User centered approach.
- **Designed for large surveys:** Survey Solutions is specifically designed for LSMS and HBS-types of surveys: Nested Rosters, Cascading and Linked questions, Roster-specific validations. Online collaboration. Web Interviewing for difficult to reach and more sophisticated respondents.
- **Data security:** Survey Solutions allows storing data on the local servers of NSO thus complying with the local data privacy and anonymity laws.

# Survey Solutions: data capturing

- Large questionnaires: several thousand questions
- All standard types of questions:
  - text, numeric, date; multi-choice; dynamic lists;
  - Linked questions: Whose cow is it? User can select from the list of household members.
  - GPS location; Time; Barcode; Binary files (pictures)
  - **HQ/Supervisor-filled questions**
- Rosters: can be generated from:
  - Fixed lists; dynamic lists; numeric; multi-choice questions.
  - **Nested rosters with unlimited degree of nesting**
- Interviewer Comments on a question and interview
- Question Instructions

# Para Data: Adaptive Survey Design

- **Improve data quality by correcting survey process during the field operation.**
- System records all events with timestamps that happen on a tablet: Data entry, Data correction, Responsibility changes, etc.
- Analysis of time per interview, time per question, section.
- Changes in productivity over time, for different interviewers, teams
- Quality control, monitoring and evaluation



# Audit Trail in Each Data Case

- Accurate Tracking of Data on each household

## Household Data



000d9f007a73492ead8df3451d267ede.tab

## Person Data



00ad1581885b404a875483dc6ac838b4.tab



**Id**

enumeration_dis	ref_week	month	year	address	community	town	district	household_no	GeoLocation__Latitude	GeoLocation__Longitude	GeoLocation__Accuracy	GeoLocation__Altitude	GeoLocation__Timestamp
001c105eb3054f3c84e0baea1484f9a5	2304	5/13/2017	5	2017Bagatelle	Bagatelle	Castries	2	16	14.00038	-60.9833	20	-10.609916	2017-05-15:15:31
009c9795edfb4947b3fdca61927335d2	9801	5/6/2017	5	2017jetrine saltibus	jetrine	laborie	Laborie	51	13.78323	-61.0063	9	103.30	2017-06-18:22:00
00ad1581885b404a875483dc6ac838b4	113	5/2/2017	5	2017Cantier	Canter	Cantier	1	115	13.75567	-60.9776	4	24.10	2017-05-23:09:40
00c5476754934bcb8460af606f4d0140	12109	7/8/2017	7	2017viger	viger	micoud	10	33	13.78704	-60.9325	6	67.25	2017-07-21:00:34
00e6c037d7c5408eb8198ae96a5c71c9	12101	5/6/2017	5	2017desruisseaux	desruisseaux	micoud	10	99	13.80612	-60.9462	48	141.4383	2017-05-27:13:59:18
0188c26c0ca24467bb344e214a4c0eba	501	7/29/2017	7	2017Faux Au Faux	Castries	Castries	1	40	14.00877	-60.9953	20	84.71567	2017-08-27:19:48:16
0189e6bc56ab44b794fd43725e44f549	1301	9/30/2017	9	2017Morne Du Don	Morne Du Don	Castries	2	75	14.01151	-60.9801	7	67.04	2017-11-16:12:55

# Households

Relations_to_head_of_HH	sex	Resident_of_SL	ParentId1
0	1	1	1001c105eb3054f3c84e0baea1484f9a5
1	3	2	1001c105eb3054f3c84e0baea1484f9a5
2	7	1	1001c105eb3054f3c84e0baea1484f9a5
3	5	2	1001c105eb3054f3c84e0baea1484f9a5
0	1	1	1009c9795edfb4947b3fdca61927335d2
1	2	2	1009c9795edfb4947b3fdca61927335d2
2	3	1	1009c9795edfb4947b3fdca61927335d2
0	1	1	100ad1581885b404a875483dc6ac838b4
1	2	2	100ad1581885b404a875483dc6ac838b4
2	3	1	100ad1581885b404a875483dc6ac838b4
3	3	2	100ad1581885b404a875483dc6ac838b4
0	1	1	100c5476754934bcb8460af606f4d0140
0	1	2	100e6c037d7c5408eb8198ae96a5c71c9
1	5	2	100e6c037d7c5408eb8198ae96a5c71c9
0	1	1	10188c26c0ca24467bb344e214a4c0eba
1	2	2	10188c26c0ca24467bb344e214a4c0eba
2	3	2	10188c26c0ca24467bb344e214a4c0eba
0	1	2	10189e6bc56ab44b794fd43725e44f549
1	3	1	10189e6bc56ab44b794fd43725e44f549
2	5	2	10189e6bc56ab44b794fd43725e44f549

# Persons

**Automatic Management of links between records is a powerful management feature which avoids the creation of duplicates and manages the unique tracking of records**



# Household Questionnaire Capture Log on Tablet

• AnswerSet	05/05/2017	13:36:51	district	9
• AnswerSet	05/05/2017	13:36:51	enumeration_district	11204
• AnswerSet	05/05/2017	13:36:51	subsample	5042b
• AnswerSet	05/05/2017	13:36:51	ref_week	5/20/2017
• AnswerSet	05/05/2017	13:36:51	district	Vieux-Fort
• AnswerSet	05/05/2017	13:36:51	household_no	98
• QuestionDisabled				Males_Left_SL_to_live_abroad
• QuestionDisabled				Females_left_SL_to_live_abroad
• QuestionDisabled				Food_Insecure_NoofTimes
• QuestionDisabled				indicate_if_makeshift
• QuestionDisabled				other_type_ofdwelling
• QuestionDisabled				future_monthly_rent
• QuestionDisabled				monthly_rent
• SupervisorAssigned	05/05/2017	13:36:51		
• InterviewerAssigned	05/18/2017	13:13:02		
• ReceivedByInterviewer	05/18/2017	21:19:32		
• AnswerSet	05/23/2017	20:56:45	GeoLocation	13.745046,-60.9645087[24.1809997558594]0
• AnswerSet	05/24/2017	13:24:57	address	Cedars Heiht
• AnswerSet	05/24/2017	13:25:06	community	"Cedars "
• AnswerSet	05/24/2017	13:25:12	address	Cedars Height
• AnswerSet	05/24/2017	13:25:15	community	Cedars Height
• AnswerSet	05/24/2017	13:25:20	town	Vieux fort
• Completed	05/24/2017	13:25:33	No contact	
• ReceivedBySupervisor		05/24/2017	13:33:52	
• ApproveBySupervisor		05/24/2017	14:44:48	no contact
• ApproveByHeadquarter		06/16/2017	17:56:43	

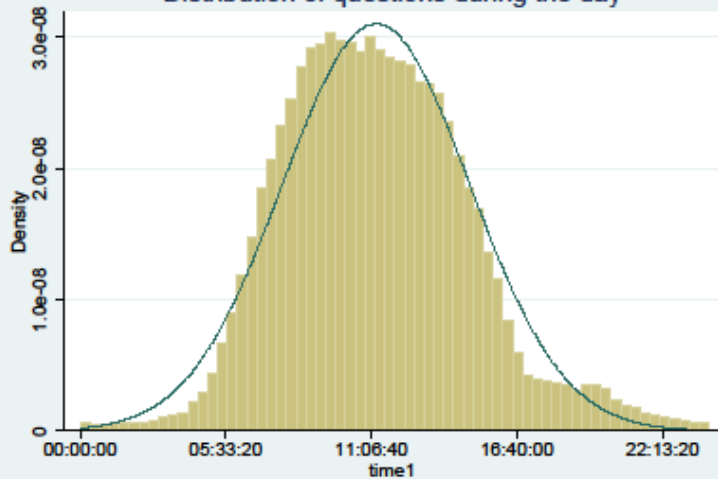


# Person Questionnaire Capture Log on Tablet

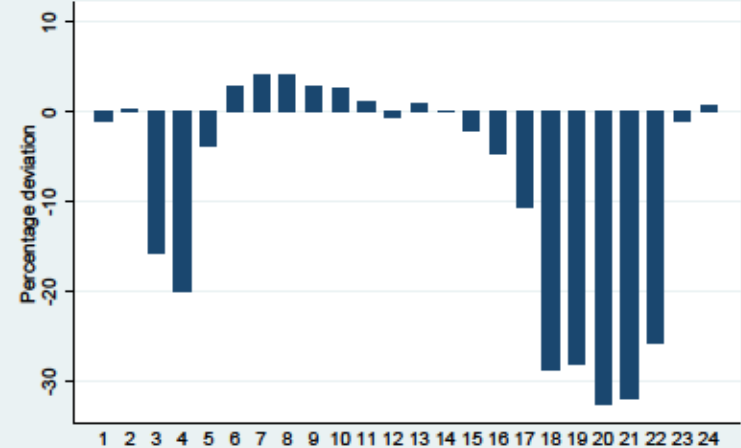
• AnswerSet	05/05/2017	13:36:50	district	9			
• AnswerSet	05/05/2017	13:36:50	enumeration_district	11305			
• AnswerSet	05/05/2017	13:36:50	subsample	5041b			
• AnswerSet	05/05/2017	13:36:50	ref_week	5/6/2017			
• AnswerSet	05/05/2017	13:36:50	district	Vieux-Fort			
• AnswerSet	05/05/2017	13:36:50	household_no	115			
• QuestionDisabled					Males_Left_SL_to_live_abroad		
• QuestionDisabled					Females_left_SL_to_live_abroad		
• QuestionDisabled					Food_Insecure_NoofTimes		
• QuestionDisabled					indicate_if_makeshift		
• QuestionDisabled					other_type_ofdwelling		
• QuestionDisabled					future_monthly_rent		
• QuestionDisabled					monthly_rent		
• SupervisorAssigned	05/05/2017	13:36:50					
• InterviewerAssigned	05/05/2017	14:03:00					
• ReceivedByInterviewer	05/05/2017	17:23:54					
• AnswerSet	05/10/2017	23:02:01	GeoLocation	13.75566509,-60.97762324[4]24			
• AnswerSet	05/10/2017	23:03:27	phone				
• AnswerSet	05/10/2017	23:03:34	contact				
• AnswerSet	05/10/2017	23:03:44	Number_of_ppl	4			
• AnswerSet	05/10/2017	23:03:46	Qty_persons_live_abroad	0			
• AnswerSet	05/10/2017	23:03:47	ChildUnder5Death	2			
• AnswerSet	05/10/2017	23:03:48	Food_Insecurity	2			
• AnswerSet	05/10/2017	23:04:02	name		Fake Name	0	
• AnswerSet	05/10/2017	23:04:06	Relations_to_head_of_HH	1		0	
• AnswerSet	05/10/2017	23:04:08	sex	1		0	
• AnswerSet	05/10/2017	23:04:09	Resident_of_SL	1		0	
• AnswerSet	05/10/2017	23:04:16	level_of_education	2		0	
• AnswerSet	05/10/2017	23:04:18	health	1		0	
• AnswerSet	05/10/2017	23:04:19	health	2		0	
• AnswerSet	05/10/2017	23:04:25	age	52		0	
• AnswerSet	05/10/2017	23:04:29	Highest_level_of_exam	1		0	
• AnswerSet	05/10/2017	23:04:29	Highest_level_of_exam	2		0	
• AnswerSet	05/10/2017	23:04:33	Highest_level_of_exam	1		0	
• AnswerSet	05/10/2017	23:04:38	training_received	1		0	
• AnswerSet	05/10/2017	23:04:39	Method_of_training	2		0	
• AnswerSet	05/10/2017	23:04:44	Working_history_for_past	52		0	
• AnswerSet	05/10/2017	23:04:47	economic_activity	1		0	
• AnswerSet	05/10/2017	23:04:51	multiple_job_holding	2		0	
• AnswerSet	05/10/2017	23:04:56	usual_working_time	40		0	
• AnswerSet	05/10/2017	23:04:58	actual_hrs	40		0	

# Paradata: Improving data quality

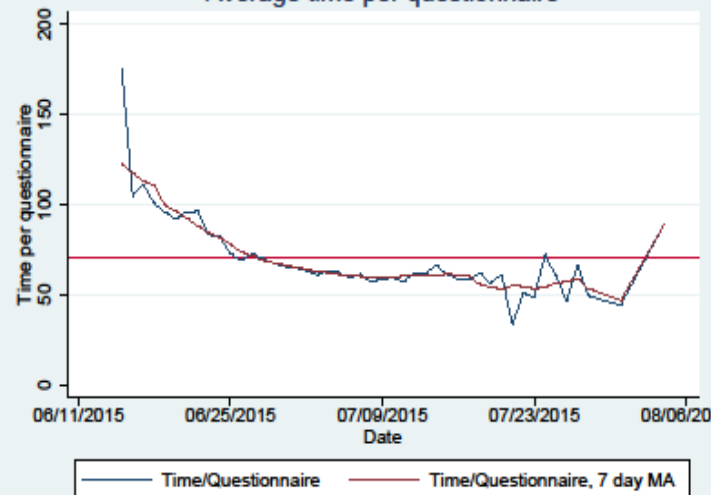
Distribution of questions during the day



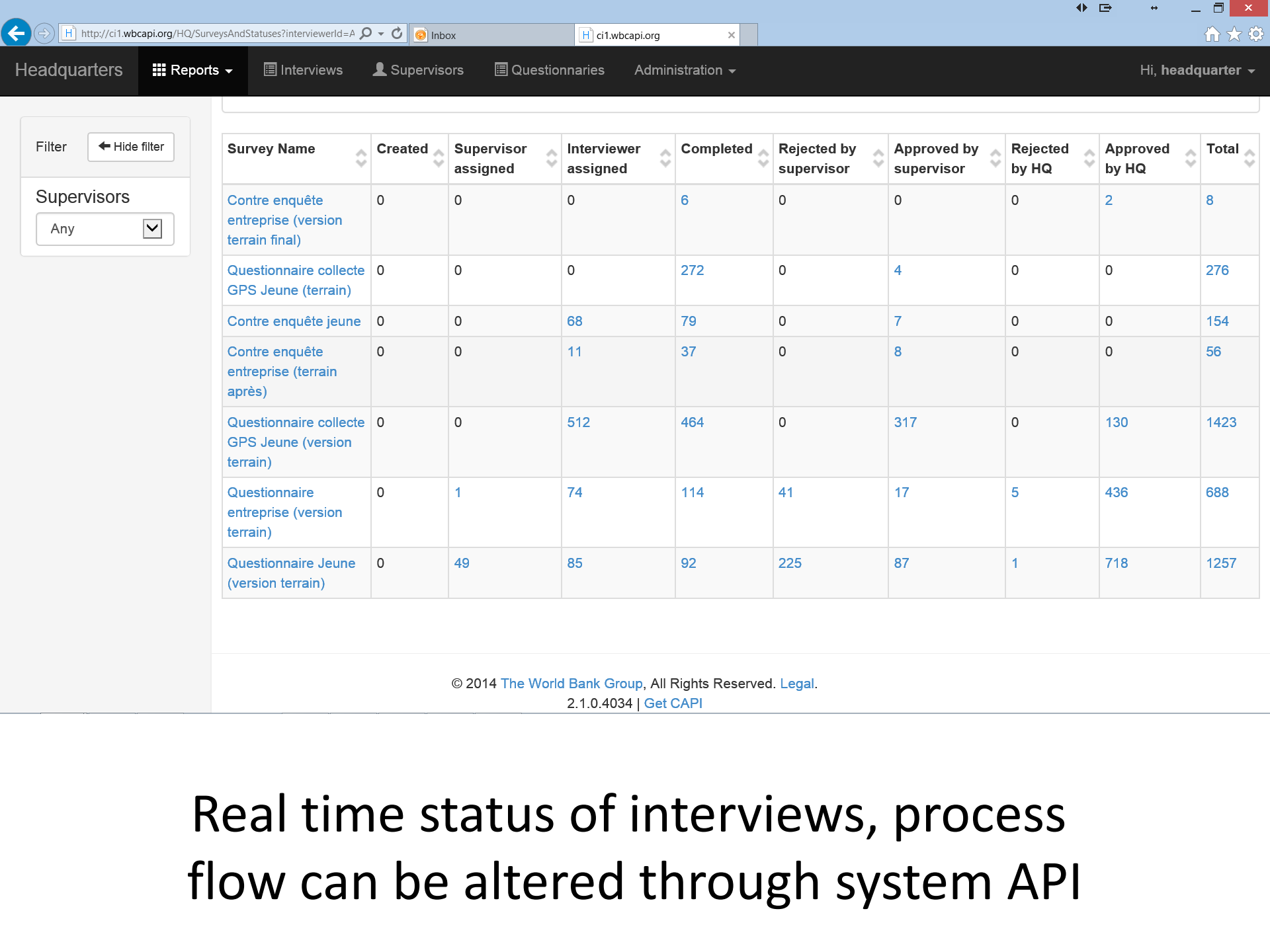
Deviations from question specific mean by period of the day (in 1 hour steps)



Average time per questionnaire



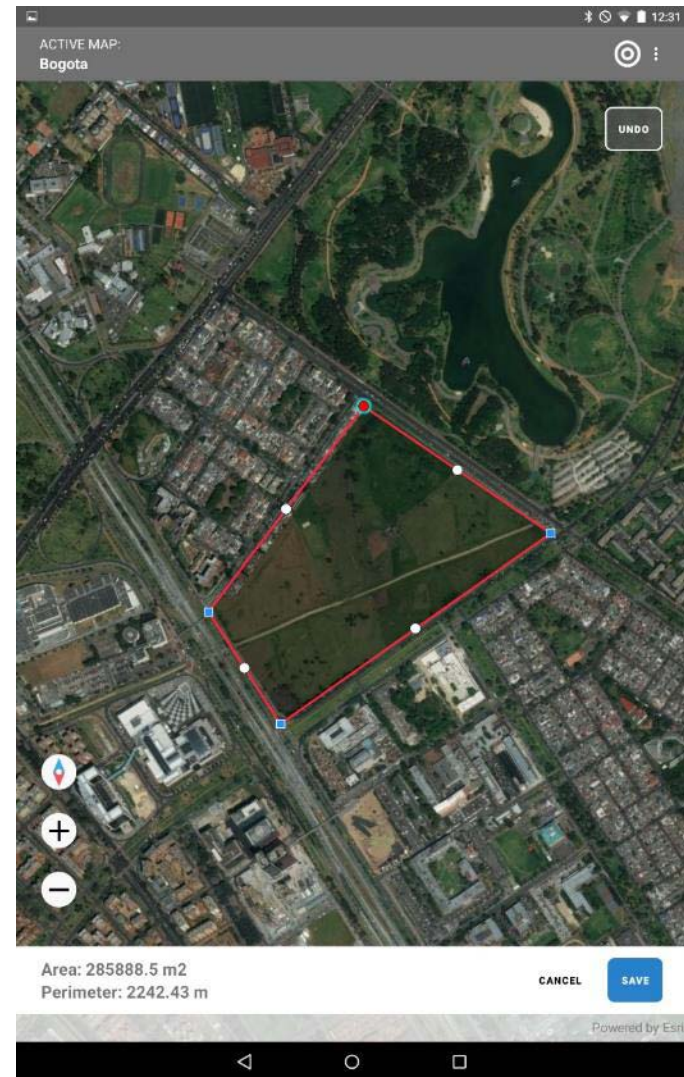




Real time status of interviews, process flow can be altered through system API

# ESRI Map Integration on GIS For Sub-Meter Accuracy for GPS Location

The screenshot shows a mobile application interface with a green header bar containing a menu icon and the text "New Section". Below the header, there are two sections for marking a plot. The first section, titled "Mark your plot", displays "Area: 168490.810743947" and "Length: 1673.5900266637" with a close icon. Below this is a button labeled "Tap to edit area". The second section, titled "Area", displays "Area: 285888.496087831" and "Length: 2242.42645896673" with a close icon. Below this is another button labeled "Tap to edit area". At the bottom of the form is a green button labeled "TO COMPLETE SCREEN".







Headquarters Reports Interviews Supervisors Questionnaires Administration Hi, karol

Questionnaire  
First Workers Follow-Up - BRAC

Version  
1

Variables  
q0\_15

Add markers on map

Currently on map

- First Workers Follow-Up - BRAC Training Project (ver. 1, var. q0\_15)

Interviewer: Immaculate  
Supervisor: Annav  
Status: Completed  
Completion date: 9/3/2014  
[View interview content: details](#)

Masaka - Kampala Rd

Masaka - Kampala Rd

Clear all markers

Map data ©2014 Google Imagery © 2014 The World Bank Group, All Rights Reserved. Legal.

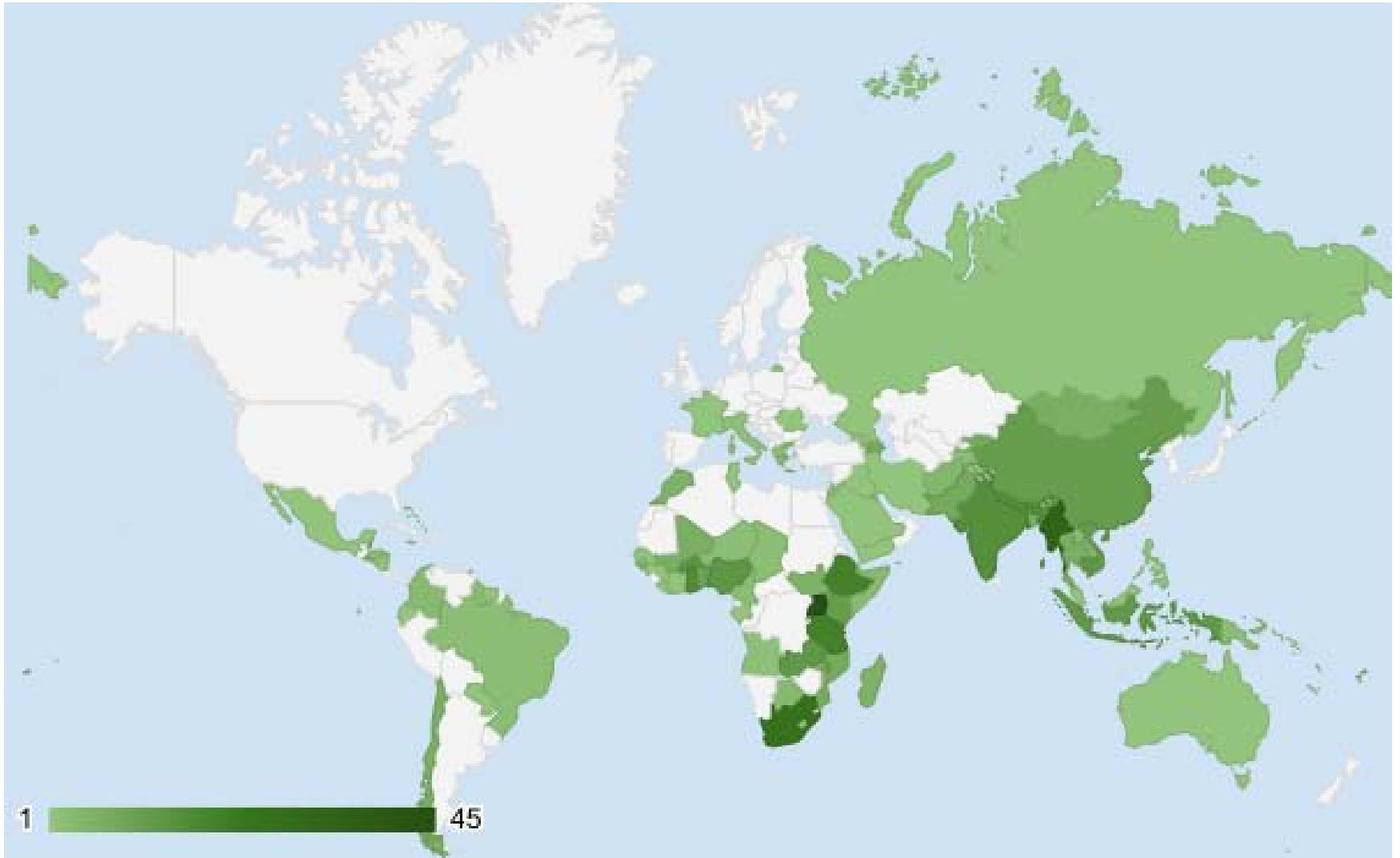
Monitor the survey by checking the GPS location of where and when the interview took place.



# Collaboration with AWS and Facebook to Reduce Costs of Data For Census/Surveys

- Sampling based on population estimates using high resolution satellite images and ML
- Anthropometric measurements using image recognition
- Verification of individuals in panel surveys using face recognition and face aging algorithms.
- Experimentation in determining the reliability of answers based on facial expressions.
- ML for non standard units
- Sign language recognition

**SURVEY PIPELINE: 700+ NATIONAL SURVEYS IN 109 COUNTRIES**  
**7,500,000+ FACE TO FACE INTERVIEWS**



# SURVEY/CENSUS SCOPE

- **Census: South Pacific Countries, Fiji etc Saint Lucia in Planning Phase**
- **Survey types:** LSMS, HBS, LFS, Enterprise Survey, EDU, Health
- **Large Clients:** India NSSO; Stats SA; Thailand NSO; FAO; IFAD; OECS; AfDB; IADB; OECS; IFS, LSTD, ..., BRAC; OPM; Mathematica
- **Largest survey:** SA – 2,600,000 households, 12,000 enumerators, 500 supervisors, 60K questionnaires per day
- **Questionnaire:** Malawi – 3,000+ questions, 94 rosters
- **Other large surveys:**
  - Indonesia Sakarnas 200,000
  - Ethiopia, Uganda, Malawi LSMS
  - Nigeria LSMS
  - WAEMU Household surveys in 12 countries of West Africa
  - Many others



## PROCESS OF IMPLEMENTATION IN ST LUCIA

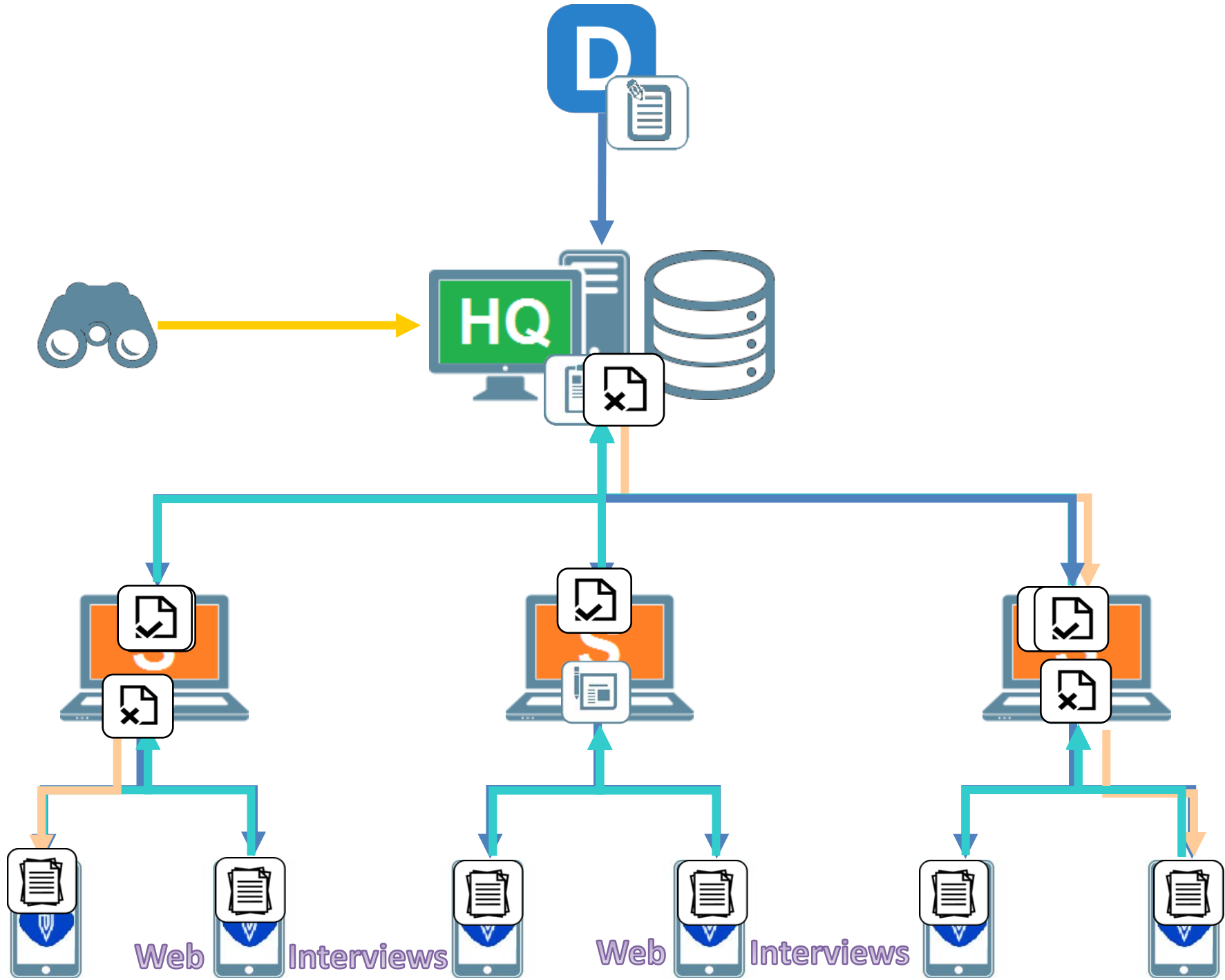
- From January 2013, full implementation of Survey Solutions v5.5
  - Quarterly Labour Force Survey
- The SLC/HBS 2016 is implemented using Survey Solutions
  - Completed a video in collaboration with the FAO/GIS to promote the use of Survey Solutions





## STARTUP OF CAPI IMPLEMENTATION PROCESS

- CSO Trinidad expected to commence in April 2016.
- Suriname: Request made for Training to the World Bank Computational Tools Division
- St Vincent and the Grenadines/Grenada/St Kitts and Nevis/Antigua and Barbuda: Use of Survey Solutions in the LFS upcoming.
- Belize – LFS and Pilot Testing for HBS



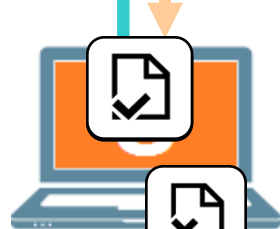
**OBSERVER**



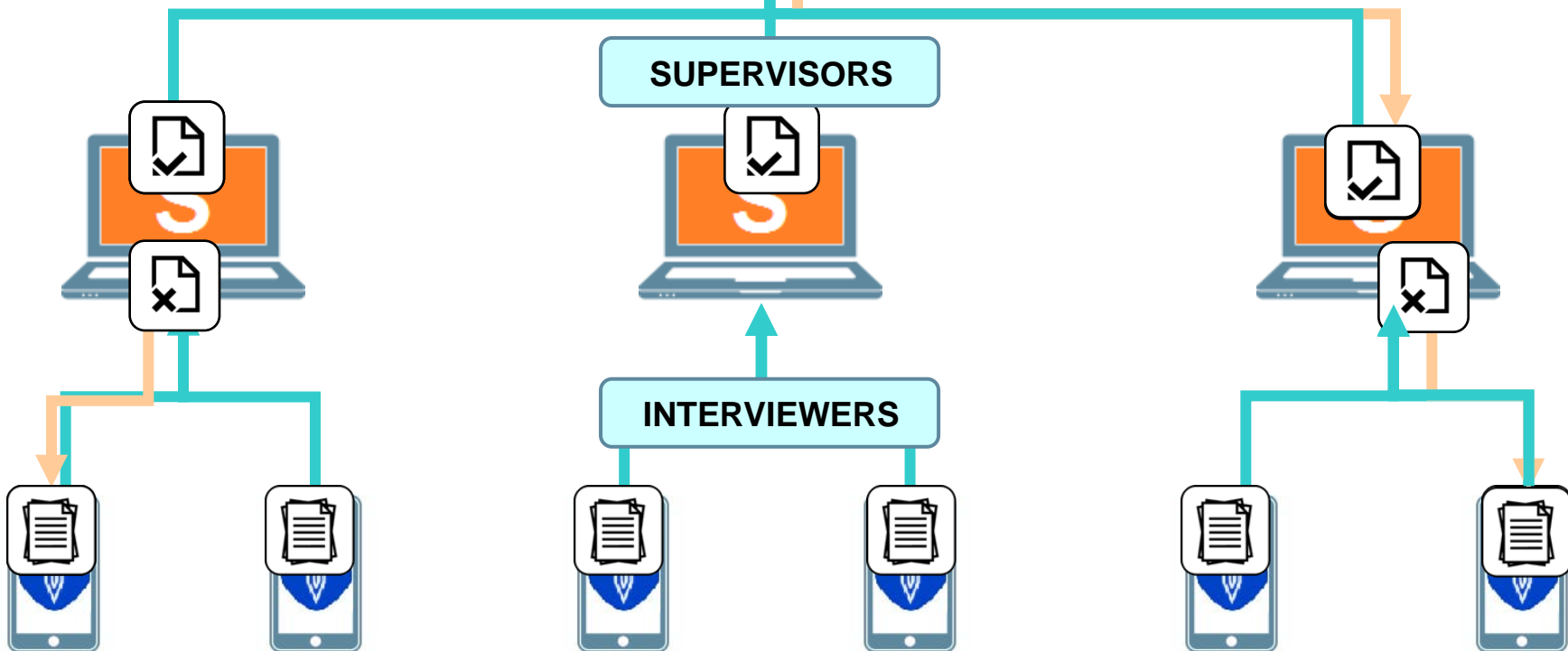
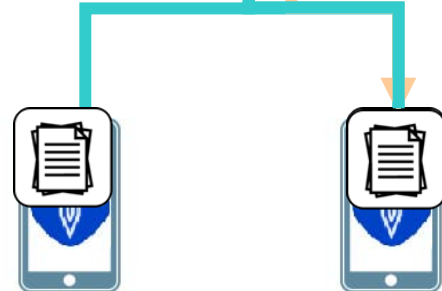
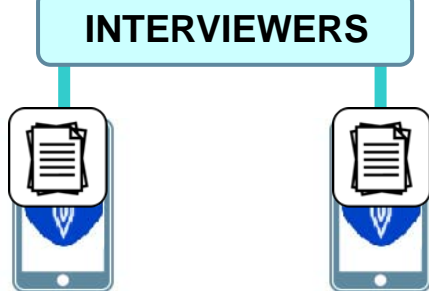
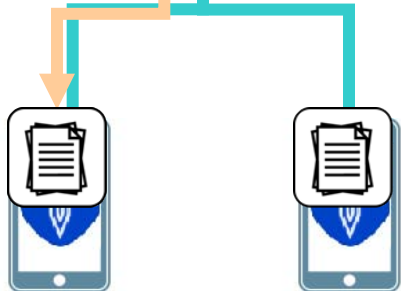
**HEADQUARTERS**

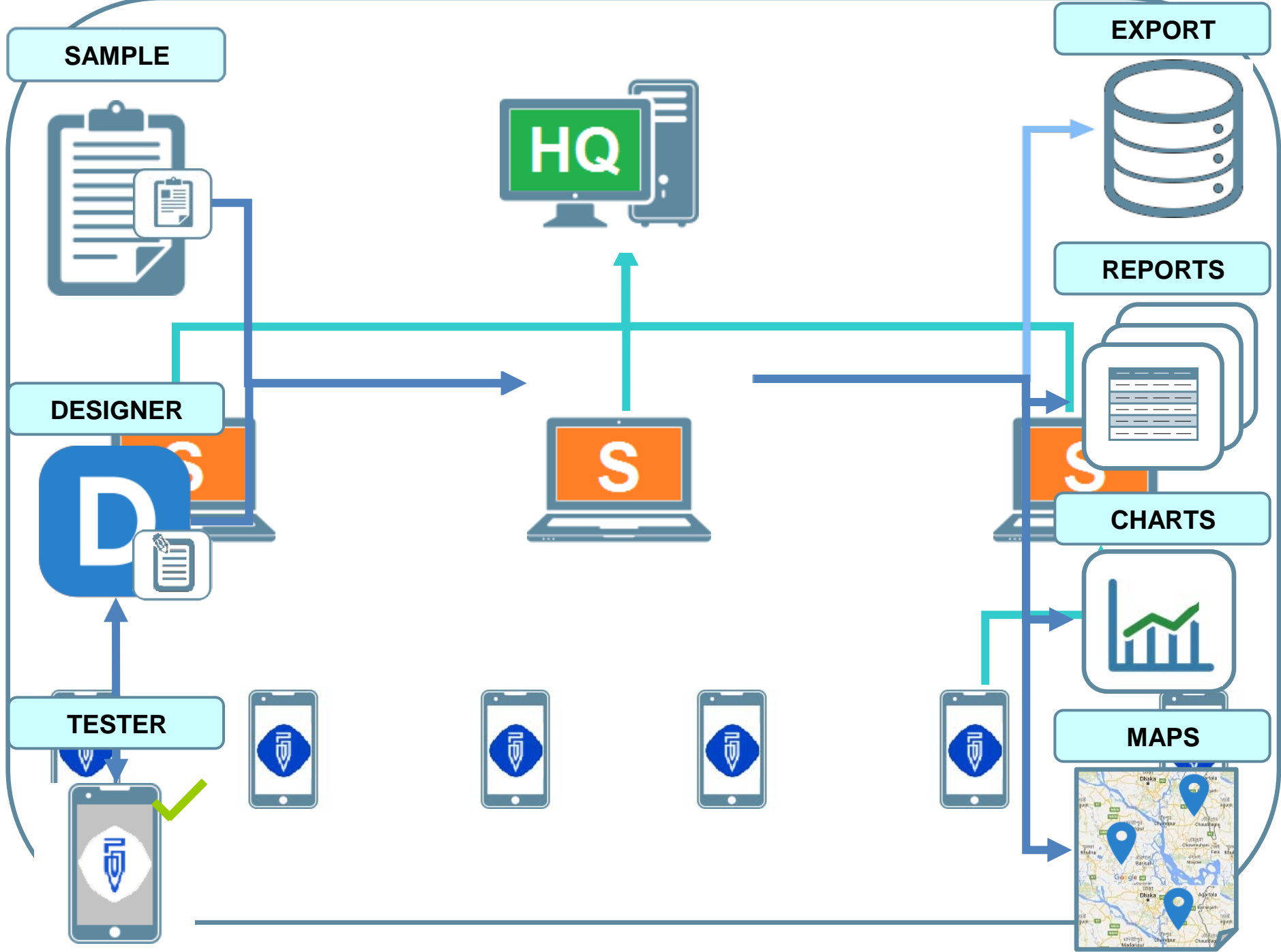


**SUPERVISORS**



**INTERVIEWERS**





# Main Issues

- Number of tablet devices to purchase for a de-jure census to be completed within four weeks
  - No of households to be surveyed
  - Final length of Questionnaire
  - Monitoring reports to be produced using data submitted on each EA
  - Elimination of control forms used in paper environment
- Items to preload on Headquarters Module and Tablet
  - Household Address list Census Questionnaire – Created Dynamically or preloaded
  - Pre-loaded showing routes to each address using Open GIS map.me or Google Maps, actions with missing addresses
  - Maps of shape files to constrain enumerators within enumeration boundaries
  - Census Questionnaire with full list of validations
  - 95% of validations can be pre-loaded on the tablets, 5% post census validation especially hot-decks type imputation validations
- Headquarter module of the survey requires use of desktop or laptop computer. Supervisor must demonstrate ability to use a browser and must own a laptop or desktop computer. May need to have some backups.
- Decide on rules for data submission to HQ
  - At Least once every other day to ensure supervisor can review and reject any forms



# Tablet Application - Control

The screenshot shows a tablet application interface. At the top, there is a black status bar with various icons and the time 7:20 AM. Below this is a green header bar with a hamburger menu icon on the left, the text "HOUSEHOLD INFO", and a vertical ellipsis icon on the right. The main content area is divided into two sections. The first section is titled "Capture GPS Coordinates" and features a map of a coastal area with labels for "Bridgetown", "Four Roads", "Charnocks", and "Oistins". A blue location pin is placed on the map. Below the map, the coordinates "-59.596948, 13.0762477" are displayed. A button labeled "Tap to record GPS" is positioned below the coordinates. The second section is titled "Telephone Number" and contains a text input field with the value "453-7670" and a close button (X) on the right.

HOUSEHOLD INFO

Capture GPS Coordinates

Bridgetown Four Roads  
Charnocks  
Oistins

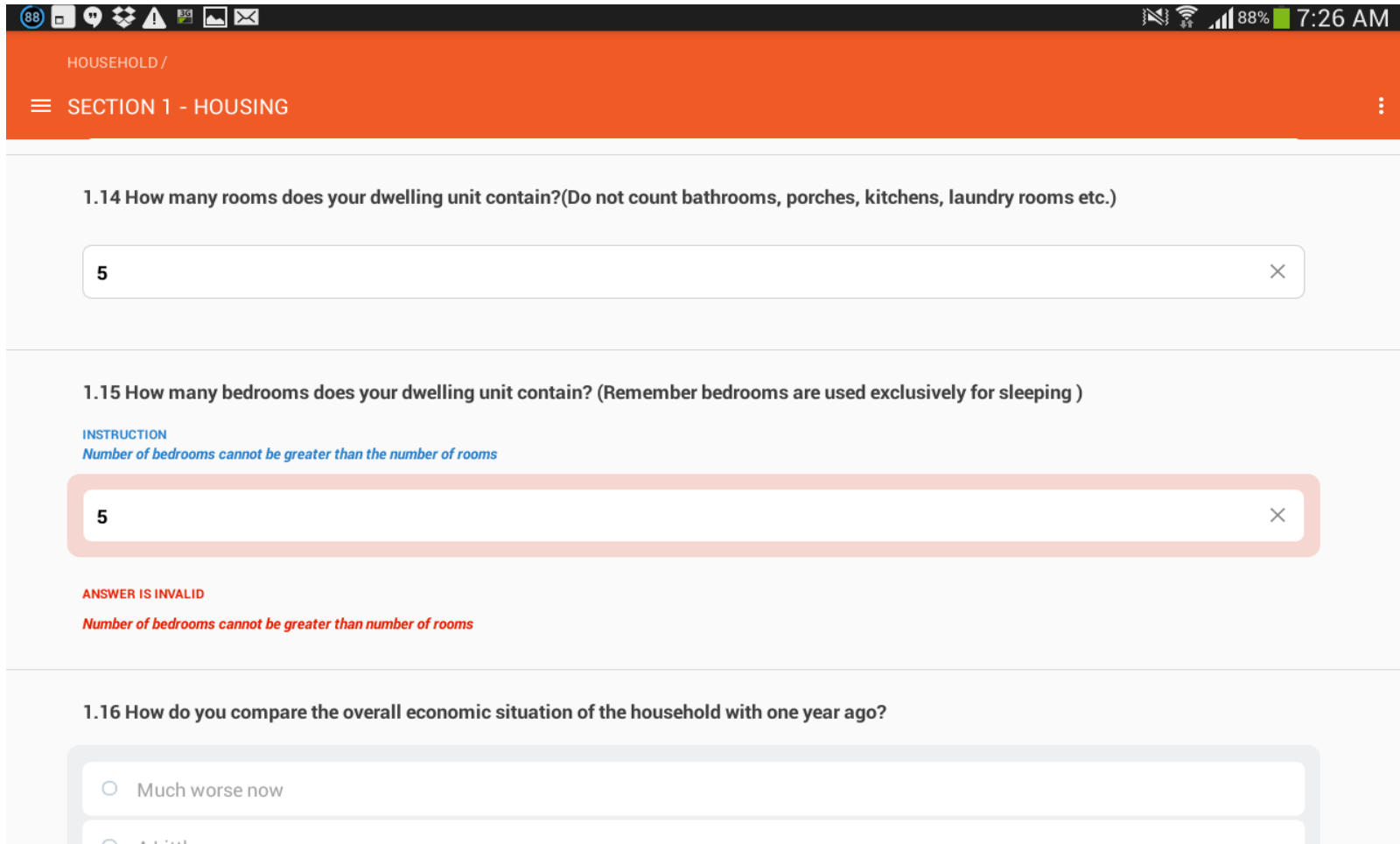
-59.596948, 13.0762477

Tap to record GPS

Telephone Number

453-7670

# Tablet Application - Housing



HOUSEHOLD /

SECTION 1 - HOUSING

1.14 How many rooms does your dwelling unit contain?(Do not count bathrooms, porches, kitchens, laundry rooms etc.)

5

1.15 How many bedrooms does your dwelling unit contain? (Remember bedrooms are used exclusively for sleeping )

INSTRUCTION  
*Number of bedrooms cannot be greater than the number of rooms*

5

ANSWER IS INVALID  
*Number of bedrooms cannot be greater than number of rooms*

1.16 How do you compare the overall economic situation of the household with one year ago?

Much worse now

A little worse now

# Tablet Application - Housing

90 [Icons] 89% 7:21 AM

☰ HOUSEHOLD ☰

**SECTION 1 - HOUSING**

Section started: 15 answers, 3 sub-sections, one error

**SECTION 2 - EXPENDITURE ON ACCOMMODATION**

Section not started

During the past 12 months have you incurred any expenditure for the purpose of repairing and maintaining the accommodation occupied by your household?

Yes

No

**SECTION 2 (2) - ROUTINE REPAIR AND MAINTENANCE OF DWELLING LAST 12 MONTHS**

Section completed: 0 answers, 2 sub-sections

# Tablet Application - Person

PERSONS / PERSON DETAILS - GERALDINE /

SECTION 1: CHARACTERISTICS

Male

Female

1.3 How old was Geraldine's on his/her last birthday?

25

1.3a If Geraldine refuses: Provide an estimate of age

1.4 To what ethnic, racial or national group do you think Geraldine's belongs?

African Descent/Negro/Black

Indigenous People

East Indian

# Tablet Application - Person



1.7 Do you/does Geraldine's have/use a cellphone?

- Yes
- No

1.8 What type of phone is Geraldine's main mobile phone?

1.9 How much did Geraldine's spend/was billed on cellular telephone (pre-paid and/or post-paid) in the last month (exclude balances from previous bills, include monthly package cost and additional cost for data, text messaging, caller ID, call waiting, etc)

1.10 Where did Geraldine's use the internet in the last 12 months? (all that apply)

- Home
- Work