# Research Terms of Reference 2020 Multi-Sector Needs Assessment NGA2002 Nigeria

01/06/2020 V1

#### REACH Informing more effective humanitarian action

# 1. Executive Summary

Country of	Niger	Nigeria				
intervention		I				
Type of Emergency		Natural disaster	X	Con	oflict	
Type of Crisis		Sudden onset		Slov	v onset	X Protracted
Mandating Body/	OCH	OCHA Nigeria Inter-Sector Working Group (ISWG)				
Agency						
Project Code	35iAN	IB/35EAB				
Overall Research						
I imetrame (from	01/04	/2020 to 31/12/2020				
research design to final						
Research Timeframe	1 Sta	art collect data: 29/06/2020			5 Preliminary pre	sentation: mid-October
Add planned deadlines	2 Da	ta collected: 23/08/2020			6 Outputs sent fo	r validation: 30/09/2020
(for first cycle if more than	2. Du	ta analysed: 16/09/2020			7 Outputs publish	ned: 16/10/2020
1)	4 Da	ta sent for validation: 04/09/	2020	)	8 Final presentati	ion: mid-November
Number of	1. Du	Y. Data serie for validation. 04/03/2020     0.1 mail presentation. mid-November       X. Single assessment (one cycle)				
assessments		Multi assessment (more th	an c	ne cv	vcle)	
	IDescribe here the frequency of the cycle					
Humanitarian	Miles	tone	loy c		Deadline	
Humanitarian	Miles	tone		+	Deadline Mid-October	
Humanitarian milestones Specify what will the	Miles X	tone HNO Sectoral Analysis Su	ppor	t	Deadline Mid-October	
Humanitarian milestones Specify what will the assessment inform and	Miles X X	tone HNO Sectoral Analysis Su HNO Joint Analysis Works	ppoi hop	t	Deadline Mid-October Mid-October	
Humanitarian milestones Specify what will the assessment inform and when	Miles X X X	tone HNO Sectoral Analysis Su HNO Joint Analysis Works Donor plan/strategy	ppoi hop	t	DeadlineMid-OctoberMid-OctoberOct 2020	
Humanitarian milestones Specify what will the assessment inform and when e.g. The shelter cluster	Miles X X X X X	tone HNO Sectoral Analysis Su HNO Joint Analysis Works Donor plan/strategy Inter-cluster plan/strategy	ppoi hop	t	DeadlineMid-OctoberMid-OctoberOct 2020Oct 2020	
Humanitarian milestones Specify what will the assessment inform and when e.g. The shelter cluster will use this data to draft its Revised Elash Appeal:	Miles X X X X X X	tone HNO Sectoral Analysis Su HNO Joint Analysis Works Donor plan/strategy Inter-cluster plan/strategy Cluster plan/strategy	ppoi	t	DeadlineMid-OctoberMid-OctoberOct 2020Oct 2020Oct 2020Oct 2020	
Humanitarian milestones Specify what will the assessment inform and when e.g. The shelter cluster will use this data to draft its Revised Flash Appeal;	Miles X X X X X X	tone HNO Sectoral Analysis Su HNO Joint Analysis Works Donor plan/strategy Inter-cluster plan/strategy Cluster plan/strategy NGO platform plan/strateg	ppor hop	t	Deadline           Mid-October           Mid-October           Oct 2020           Oct 2020           Oct 2020           Oct 2020	
Humanitarian milestones Specify what will the assessment inform and when e.g. The shelter cluster will use this data to draft its Revised Flash Appeal; Audience Type &	Miles X X X X X X Audie	tone HNO Sectoral Analysis Su HNO Joint Analysis Works Donor plan/strategy Inter-cluster plan/strategy Cluster plan/strategy NGO platform plan/strateg	ppoi hop	t	Deadline           Mid-October           Mid-October           Oct 2020           Oct 2020           Oct 2020           Oct 2020           Dissemination	
Humanitarian milestones Specify what will the assessment inform and when e.g. The shelter cluster will use this data to draft its Revised Flash Appeal; Audience Type & Dissemination Specify	Miles X X X X X X Audie X Stra	tone HNO Sectoral Analysis Su HNO Joint Analysis Works Donor plan/strategy Inter-cluster plan/strategy Cluster plan/strategy NGO platform plan/strateg ence type ategic	ppor hop	t	Deadline         Mid-October         Mid-October         Oct 2020         Oct 2020         Oct 2020         Dissemination         X General Product	t Mailing (e.g. mail to NGO
Humanitarian milestones Specify what will the assessment inform and when e.g. The shelter cluster will use this data to draft its Revised Flash Appeal; Audience Type & Dissemination Specify who will the assessment	Miles X X X X X X Audie X Stra X Pro	tone HNO Sectoral Analysis Su HNO Joint Analysis Works Donor plan/strategy Inter-cluster plan/strategy Cluster plan/strategy NGO platform plan/strateg ence type ategic	ppor hop	t	Deadline         Mid-October         Mid-October         Oct 2020         Oct 2020         Oct 2020         Oct 2020         Semination         X General Product consortium; HCT p	t Mailing (e.g. mail to NGO participants; Donors)
Humanitarian milestones Specify what will the assessment inform and when e.g. The shelter cluster will use this data to draft its Revised Flash Appeal; Audience Type & Dissemination Specify who will the assessment inform and how you will	Miles X X X X X X Audie X Stra X Pro	tone HNO Sectoral Analysis Su HNO Joint Analysis Works Donor plan/strategy Inter-cluster plan/strategy Cluster plan/strategy NGO platform plan/strateg ence type ategic ogrammatic erational	ppoi hop	t	Deadline         Mid-October         Mid-October         Oct 2020         Oct 2020         Oct 2020         Oct 2020         Semination         X General Product         consortium; HCT p         Cluster Mailing (E         and presentation of	t Mailing (e.g. mail to NGO participants; Donors) Education, Shelter and WASH)
Humanitarian milestones Specify what will the assessment inform and when e.g. The shelter cluster will use this data to draft its Revised Flash Appeal; Audience Type & Dissemination Specify who will the assessment inform and how you will disseminate to inform the audience	Miles X X X X X Audie Audie X Stra X Pro Ope [0]	tone HNO Sectoral Analysis Su HNO Joint Analysis Works Donor plan/strategy Inter-cluster plan/strategy Cluster plan/strategy NGO platform plan/strateg ence type ategic ogrammatic erational her, Specify]	ppor hop y	t	Deadline Mid-October Mid-October Oct 2020 Oct 2020 Oct 2020  Dissemination X General Product consortium; HCT µ □ Cluster Mailing (E and presentation of meeting	t <b>Mailing (e.g. mail to NGO</b> participants; Donors) Education, Shelter and WASH) findings at next cluster
Humanitarian milestones Specify what will the assessment inform and when e.g. The shelter cluster will use this data to draft its Revised Flash Appeal; Audience Type & Dissemination Specify who will the assessment inform and how you will disseminate to inform the audience	Miles X X X X X Audie X Stra X Pro Ope [Ott]	tone HNO Sectoral Analysis Su HNO Joint Analysis Works Donor plan/strategy Inter-cluster plan/strategy Cluster plan/strategy NGO platform plan/strateg ence type ategic ogrammatic erational her, Specify]	ppor hop	t	Deadline Mid-October Mid-October Oct 2020 Oct 2020 Oct 2020  Dissemination X General Product consortium; HCT p □ Cluster Mailing (E and presentation of m meeting; Cluster n	t Mailing (e.g. mail to NGO participants; Donors) Education, Shelter and WASH) findings at next cluster findings (e.g. at HCT neeting)

Detailed	X	Yes 🗆 No				
dissemination plan						
required						
General Objective	Provi	de strong, evidence-based information on multi-sectoral humanitarian needs of				
•	affect	ed populations in accessible areas across Borno. Adamawa and Yobe (BAY) States				
	of No	rth East Nigeria and inform multi-sector humanitarian programming for 2021				
Specific Objective(s)	Provi	de a comprehensive overview of sectoral and inter-sectoral needs, coping capacity				
-p	and o	verall well-being among conflict-affected population in the BAY States including				
	cina c	Internally Displaced Persons (IDPs) residing either in formal/informal camp				
		settings or within host communities:				
		Returnees including returning IDPs and returning Nigerian refugees:				
		<ul> <li>Non-displaced populations:</li> </ul>				
	Undo	retand how these people and their severity year based on geographical leastion by				
	Unde	rstand now these needs and their seventy vary based on geographical location - by				
	local	government area (LGA) - and based on population group.				
Research Questions	•	What is the severity of humanitarian needs of the crisis-affected population, and				
		how do these vary between geographical locations, population groups and				
		household profiles? [inter-sectoral analysis]				
	•	What are the underlying characteristics and vulnerabilities of the households?				
	•	What are the sector specific humanitarian needs of the crisis affected population,				
		including:				
		<ul> <li>What are the nutrition challenges experienced by children under 5 and</li> </ul>				
		mothers?				
		<ul> <li>What are the challenges of accessing health services for households</li> </ul>				
		and what is the vaccination status of children under 15?				
		• What are the issues related to water quantity and use, hygiene practices				
		and access to sanitation for households?				
		<ul> <li>What are the issues related to shelter and NFIs for households?</li> </ul>				
		• What are the issues related to food quantity and access, availability of				
		firewood/fuel and practices of agriculture for households?				
		• What are the issues related to household livelihoods and access to				
		essential services and infrastructure?				
		<ul> <li>What is the level of access to education in the bousehold?</li> </ul>				
		What is the level of access to education in the household?				
		and movement restrictions for households?				
		What are the issues related to access to assistance?				
		What are the states is a dasted by branched is associated?				
	•	what are the strategies adopted by nouseholds in reponse to the current COVID-				
		19 nealth emergency?				
Geographic Coverage	-	All LGAs in Borno, Adamawa and Yobe states, North East Nigeria, that can be				
		assessed remotely;				
	-	All physically accessible areas within LGAs that cannot be assessed remotely.				
Secondary data	Рори	lation estimates:				
sources	•	Vaccination Tracking System (VTS) population dataset (March 2020);				
	•	IOM DTM datasets (April 2020):				
		<ul> <li>Return data round XXXI</li> </ul>				
		<ul> <li>IDP Site assessment dataset round XXXI</li> </ul>				
		<ul> <li>IDP Location assessment dataset round XXXI</li> </ul>				

<sup>&</sup>lt;sup>1</sup> Including the key Humanitarian Project Cycle (HPC) milestones for 2020 (Humanitarian Needs Overview and update of the multi-year Humanitarian Response Plan)

	Huma	Humanitarian needs:						
	•	<ul> <li>Previous REACH Nigeria assessments, including MSNA 2019</li> </ul>						
	•	REACH "monitoring of ha	ard-	to-rea	ch	areas" results (c	cor	ntextual and secondary
		<u>information</u> on populations residing in Hard-to-Reach areas not covered by the MSNA)						
	•	• Sectoral-specific assessments gathered through the secondary data review						
		Any other relevant huma	nita	rian. i	ouri	nalistic, academ	nic	study on humanitarian
		needs in North East Nige	ria	as ne	ede	ed		
Population(s)	X	IDPs in camp			X	IDPs in inform	na	l sites
Select all that apply	X	IDPs in host communities	s			IDPs [Other, Sp	pe	cify]
		Refugees in camp				Refugees in in	nfo	rmal sites
		Refugees in host communi	ties			Refugees [Oth	er,	Specify]
	X	Host communities			Х	Returnees		
Stratification	X	Geographical #: 60	X	Grou	up ‡	ŧ: IDP,	[	[Other Specify] #:
Select type(s) and enter		LGAs		Retu	urn	ee, Non-		Population size per
number of strata		Population size per strata		disp	olac	ed		strata is known?
		is known? X Yes □ No		Рор	ulat	tion size per		
				strat	ta is	s known?		
Data collection tool(s)	Y	Structured (Quantitative)		X Y	es 🗆	∃ N0 Somi structure	d d	(Qualitative)
	Samr	Jing method			Semi-structured (Qualitative)			
Structured data	oump							
collection tool # 1	D Pur	posive			□ Key informant interview (Target #):			
Select sampling and data	🗆 Pro	bability / Simple random			□ Group discussion (Target #):			
collection method and	X Pro	obability / Stratified simple ra	ndo	m	X Household interview (Target #): 5748			
specify target # interviews	(Plan	A)			□ Individual interview (Target #):			
	🗆 Pro	bability / Cluster sampling			Direct observations (Target #):			s (Target #):
	X Pro	bability / Stratified cluster sa	mpl	ing	□ [Other, Specify] (Target #):			
	(Plan	B/ Contingency)						
	□ [Ot	her, Specify]						
Target level of	050/ 1				40			
precision if	95% 16	evel of confidence			10	I+/- % margin of e	erro	Ŋ
probability sampling								
Analytical Framework	X	JIAF (All components)				JIAF (Some co	om	ponents only)
used								
						[Specify which p	oilla	ars/ sub-pillars excluded]
<u> </u>		[Other, Specify]						
Lessons Learned	X	Documentation available	and	t t		No lessons lea	arr	ned documentation
incorporation from		consulted				available		
past MonAs	⊔ ▼	No MSNAs conducted in the past			[			
Data management	^	IMPACT				UNITER		
plationin(s)		[Other Specify]						
Expected ouput		Situation overview #	X	Ren	ort	#: 1 Report	Г	Profile #
type(s)				oni	nte	r-sectoral		
- <b>J</b>   <sup>2</sup> ( • )		findinas. 1						
				Exe	cut	ive summary		

	X Presentation (Preliminary fin #: 1	X dings)	Presentation (Final) #: 1	X	Factsheet #: 3 State- level sectoral factsheets, 60 LGA-level factsheets		
	Interactive dash	ooard #:_ □	Webmap #:		Map #:		
	X Dataset # : 1	·					
Access	X Public (availabl platforms)	Public (available on REACH resource center and other humanitarian platforms)					
	<ul> <li>Restricted (bilate publication on R</li> </ul>	Restricted (bilateral dissemination only upon agreed dissemination list, no publication on REACH or other platforms)					
Visibility Specify which	REACH						
logos should be on	Donor: ECHO, NHF (OCHA)						
outputs	<b>Coordination Framew</b>	rdination Framework: ISWG					
	Partners: All partners supporting data collection (TBC)						
Access Visibility Specify which logos should be on outputs	<ul> <li>Interactive dashi</li> <li>Dataset # : 1</li> <li>Public (available platforms)</li> <li>Restricted (bilate publication on R</li> <li>REACH</li> <li>Donor: ECHO, NHF (C)</li> <li>Coordination Framew</li> <li>Partners: All partners s</li> </ul>	e on REACH eral dissemina EACH or othe CHA) ork: ISWG supporting data	Webmap #: resource center and oth tion only upon agreed dis r platforms) a collection (TBC)	ner ser	humanitarian		

## 2. Rationale

## 2.1. Rationale

North East Nigeria continues to experience significant humanitarian needs after over 10 years of conflict affecting the Lake Chad region. The 2020 Humanitarian Needs Overview (HNO) identified 7.9 million individuals in the three states of Borno, Adamawa and Yobe (collectively, the "BAY" states) to be in need of humanitarian assistance.<sup>2</sup> Partners have faced increasing access restrictions to these areas since summer 2019,<sup>3</sup> a situation exacerbated further by the COVID-19 pandemic in early 2020; this has, in turn, decreased capacities to identify the scale of the needs in a comprehensive way.<sup>4</sup> Providing an updated evidence-based overview of the needs of populations in the BAY states is therefore vital to inform partners' strategic planning.

A Multi-Sector Needs Assessment (MSNA) will be implemented in 2020. This follows the 2018 and 2019 MSNAs, facilitated by REACH with the support of the United Nations Office for the Coordination of Humanitarian Affairs (OCHA) and the Inter-Sector Working Group (ISWG). This assessment aims to identify and compare needs per sector and across sectors across population groups affected by the protracted crisis in all accessible areas in the BAY states; as such, it aims to support evidence-based planning and decision-making by humanitarian actors in Nigeria and, more specifically, inform the 2021 HNO and the update of the current multi-year Humanitarian Response Plan (HRP).

# 3. Methodology

## 3.1. Methodology overview

REACH will facilitate a multi-sectoral household survey, in coordination with the Assessment and Analysis Working Group (AAWG) and the ISWG, covering Adamawa, Borno and Yobe states in North East Nigeria. All relevant sectors were asked to engage in a secondary data review exercise, to ensure that existing information gaps were identified. Partners were also engaged for feedback on the overall methodology. While the 2018 and 2019 MSNA exercises were conducted mainly through in-person household surveys, the 2020 assessment methodology was instead adapted to the current COVID-19 health emergency, following a strict 'do no harm' approach.<sup>5</sup>

A mixed methodology was proposed, to avoid in person data collection wherever possible and limit staff and vulnerable populations' potential exposure to COVID-19:

<sup>&</sup>lt;sup>2</sup> OCHA. Nigeria: 2020 Humanitarian Needs Overview. April 2020

<sup>&</sup>lt;sup>3</sup> OCHA. Nigeria – Borno State, North-East Flash Update No. 2 - Displacement to Damboa LGA, 30 May 2019

<sup>&</sup>lt;sup>4</sup> United Nations News. UN prepares for potentially devastating COVID-19 outbreak in conflict-ravaged northeast Nigeria. 30 April 2020

<sup>5</sup> REACH. Data Collection SOPs During COVID - 19

 LGAs that can be assessed remotely through a fully remote methodology (47 LGAs, Table 1), are identified; for these LGAs, a random sampling strategy will be applied based on existing phone number directories made available by the Nigerian Communications Commissions (NCC).<sup>6</sup> The target precision will be 95% confidence with 10% margin of error at the LGA-level for the overall population.

In the remaining 18 LGAs (Table 1), where a fully remote approach would not be feasible due to poor network coverage or low phone ownership rates among households,

- 2a) All physically inaccessible areas, due to security reasons or COVID-19-related restrictions, are identified and excluded from the sampling frame; among these, Abadam, Guzamala, Kukawa, Marte and Nganzai LGAs, Borno state, are entirely inaccessible and will be thus excluded from the assessment;
- 2b) A two-stage random sampling strategy is applied, based on existing population estimates, aiming for 95% confidence with 10% margin of error at the LGA-level (physically accessible areas only) for the overall population;
- 2c) Partners already present and active in these LGAs will facilitate or conduct data collection in the target locations.<sup>7</sup>

In LGAs where the above strategies were not found to be meeting the sampling targets, and as a last resort contingency option,

3) A non-probability household quota sampling will be used: a target number of households to interview (minimum 30 households per LGA, and further adjusted as relevant based on population size) will be defined for each of these LGAs; phone contacts of households residing in these areas will be identified through existing key informant networks and snowball sampling. REACH enumerators will then conduct phone interviews. Failure to meet the sampling targets could be due to changes of the security and health emergency situation in the field, low response rates to remote phone surveys.

Details on the above approaches are provided in section 3.4.

Results will be aggregated at the state level, where possible, and will be representative for each population group of interest (non-displaced, IDPs and returnees) with a target precision of 95% confidence and 10% margin of error. Aggregation of results at the state level may be negatively affected if the last resort option previously outlined were to be adopted; specifically, if different sampling methodologies were to be used in a given state, aggregation of results for that state would be affected.

	IDPs	;	Return	ees	Non-displa	aced
State	Estimated number of households (accessible areas only)	Target sample	Estimated number of households (accessible areas only)	Target sample	Estimated number of households (accessible areas only)	Target sample
Adamawa	36,845	96	13,5215	96	60,1658	96
Borno	292,681	96	14,0751	96	640,444	96
Yobe	20,251	96	7,925	95	407,908	96

The level of representativeness achieved per LGA and per state will be indicated in the final outputs. Further, due to the different sampling approach used in 2020, a complete, robust comparison with 2019 MSNA results will not be possible.

<sup>&</sup>lt;sup>6</sup> Should the NCC not be able to share a phone directory, the main phone companies operating in the BAY states will be approached by REACH to the same end.
<sup>7</sup> Only partners with ongoing presence and activities in the target LGA will be mobilized, to avoid increased circulation between LGAs; partners will provide "full" data collection resources (enumerators, cars, technical and health equipment), with REACH covering phone credit expenses and providing the necessary training on tool, household selection in the field and relevant COVID-19 standard operating procedures (SOPs).

In case no directory were shared with REACH by 29 June 2020 by the NCC or Telecommunications companies operating in North East Nigeria, a **Contingency Plan** (section **3.5**) will be implemented which consists of the following:

• For LGAs that cannot be assessed remotely (see point 2 above) and priority LGAs, a two-stage cluster sampling approach will be used across areas that are considered accessible (both in terms of security and COVID-19 risks), followed by data collection in the field; this will allow comparison of findings with last year's MSNA when the same sampling strategy was adopted. Partners already present and active in these LGAs will facilitate or conduct data collection in target locations. The list of priority LGAs was drafted in coordination with the AAWG, to identify key areas that are considered as severely affected by the North East Nigeria crisis;

• For all **remaining LGAs (non-priority LGAs that can be assessed remotely)**, non-probability household quota sampling will be used: a target number of households to interview (minimum 30 households per LGA, and further adjusted as relevant based on population size) will be defined for each of these LGAs. Phone contacts of households residing in these areas will be identified through existing key informant networks and snowball sampling. REACH enumerators will then conduct phone interviews. This approach will also be the last resort option used in case sampling targets were not met for priority LGAs, for example due to changes in access conditions in the field.

Aggregation of results at the state level will yield indicative results only in this contingency scenario. Although findings would be indicative only, they also would be more accurate than community-level estimates based on key informant interviews. Further, although not generalizable with a known level of statistical precision, non-probability sampling can still generate indicative findings with some level of representation if participant selection is done well.

Details on this contingency approach are provided in section 3.5.

## 3.2 Population of interest

Three distinct population groups are targeted by the primary data collection within the MSNA research cycle:

- Non-displaced populations those who have never been displaced due to the crisis;
- Returnees those that have been displaced but have since returned to their areas of origin (not including regular movements such as seasonal migrations);
- Internally displaced people those who have been displaced by the crisis and are still displaced in their primary or in a secondary displacement location.

The MSNA will target households residing in accessible LGAs (remotely or in person) across the BAY states (Table 1).

- An LGA that can be assessed remotely is here defined as an area where Nigerian phone network coverage is good, based on information provided by field staff and their network of Key Informants (KIs), and where at least 60% of households reportedly own a mobile phone, based on 2019 MSNA results.
- For LGAs that cannot be assessed remotely and where the presence of partner enumerators in the field is needed, a physically accessible area is here defined as an area that is accessible both in terms of security and COVID-19 risk considerations. Inaccessible areas due to security reasons or COVID-19 risks/ movement restrictions will be identified on a map by field officers, as well as based on Access Working Group and other security-related documents, and excluded from the sample. The feasibility of field visits will be revisited on a daily basis during data collection.

State	LGAs that can be assessed remotely	LGAs that cannot be assessed remotely (partial physical access)	LGAs that cannot be assessed remotely (no physical access)*
Adamawa	Demsa, Fufore, Ganye, Gombi, Girei, Guyuk, Hong, Jada, Lamurde, Madagali, Maiha, Mayo-Belwa, Michika, Mubi North, Mubi South, Numan, Shelleng, Song, Yola North, Yola South	Toungo	1
Borno	Askira/Uba, Bayo, Biu, Chibok, Hawul, Jere, Kaga, Konduga, Kwaya Kusar, Maiduguri, Shani	Bama, Damboa, Dikwa, Gubio, Gwoza, Kala/Balge, Mafa, Magumeri, Mobbar, Monguno, Ngala	Abadam, Guzamala, Kukawa, Marte, Nganzai
Yobe	Bade, Bursari, Damaturu, Fika, Fune, Gujba, Gulani, Jakusko, Karasuwa, Machina, Nangere, Nguru, Potiskum, Tarmua, Yunusari, Yusufari	Geidam	1

Table 1. Categorisation of LGAs, based on the applicability of a fully remote methodology and based on physical access considerations.

\*These LGAs have been identified as both remotely and physically inaccessible due to security constraints, COVID-19 related restrictions, and poor mobile network infrastructure; as such they will not be covered by primary data collection.

As much as possible, the MSNA final analysis and outputs will also try to incorporate contextual and secondary data regarding populations residing in hard-to-reach areas excluded from primary data collection, as this population group is also a core group included in the Humanitarian Programme Cycle milestone documents. This information will stem from another REACH assessment which monitors multi-sectoral needs in hard-to-reach areas.

## 3.3. Secondary data review

A comprehensive secondary data review is currently underway, led by OCHA with the support of the IMWG and REACH. Relevant assessments will be compiled and sorted by their relevance to pillars of the Joint Inter-Sectoral Framework. This secondary data review, along with bilateral consultations between REACH and sectors, is used to inform the MSNA team on where information gaps lie and how to best address them; the SDR will also be feeding into the HNO and HRP narratives. Other secondary sources used for this research cycle include:

For population estimates

- Vaccination Tracking System (VTS) population dataset (March 2020);
- IOM Displacement Tracking Matrix (DTM) datasets (April 2020): Return data round XXXI, IDP Site assessment dataset round XXXI, IDP Location assessment dataset round XXXI.

For contextual information on humanitarian needs

- Previous REACH Nigeria assessments, including MSNA 2019;
- The 2019 Humanitarian Needs Overview;
- REACH "monitoring of hard-to-reach areas" results (contextual and secondary information on populations residing in Hard-to-Reach areas not covered by the MSNA).
- Any other relevant humanitarian, journalistic, academic study on humanitarian needs in North East Nigeria as needed.

## 3.4 Primary Data Collection

Data collection is planned to start on 29 June and last eight weeks. The household survey used, developed in consultation with all sectors and the AAWG and based on standardised MSNA and draft indicators proposed for the Joint Inter-Sectoral Analysis Framework, will be coded using Kobo and will be filled in by REACH or partner enumerators on their work

smartphones during data collection. A different methodology will be used in different LGAs, based on the applicability of a fully remote methodology and based on physical access considerations. The total target sample consists of 5,745 surveys, across 60 LGAs.

#### LGAs that can be assessed remotely

A random directory sampling strategy with remote phone data collection will be used in LGAs where network coverage is good, based on information provided by field staff and their network of Key Informants (KIs), and where at least 60% of households reportedly own a mobile phone, based on 2019 MSNA results. An anonymised phone number directory (phone number, State and LGA of registration) will be required for the three BAY states from the NCC. Each SIM card purchased in Nigeria is registered with the mobile phone network, associated with the owner's residence address and then passed on to the NCC. Through OCHA, REACH wishes to obtain such a directory as a basis for random directory sampling. The target samples are defined aiming for findings to be representative at the LGA level for the overall population, with a confidence level of 95% and 10% margin of error.

Adamawa					
	Total				
	estimated	Target			
LGA	number of	sample			
	households				
Demsa	29,229	96			
Fufore	41,367	96			
Ganye	32,212	96			
Gombi	34,446	96			
Girei	43,180	96			
Guyuk	24,261	96			
Hong	70,130	96			
Jada	32,935	96			
Lamurde	19,035	96			
Madagali	34,522	96			
Maiha	27,413	96			
Mayo-	33 130	06			
Belwa	55,159	90			
Michika	56,043	96			
Mubi North	47,699	96			
Mubi South	48,779	96			
Numan	21,632	96			
Shelleng	21,979	96			
Song	46,546	96			
Yola North	37,048	96			
Yola South	65,950	96			

Borno					
	Total				
IGA	estimated	Target			
LUA	number of	sample			
	households				
Askira/Uba	89,236	96			
Bayo	34,110	96			
Biu	64,869	96			
Chibok	31,324	96			
Hawul	58,774	96			
Jere	146,316	96			
Kaga	36,063	96			
Konduga	77,700	96			
Kwaya	27.055	96			
Kusar	21,000	30			
Maiduguri	186,670	96			
Shani	45,892	96			

Yobe					
	Total				
	estimated	Target			
LGA	number of	sample			
	households				
Bade	23,301	96			
Bursari	18,114	96			
Damaturu	37,975	96			
Fika	29,403	96			
Fune	35,991	96			
Gujba	34,111	96			
Gulani	26,647	96			
Jakusko	21,691	96			
Karasuwa	13,018	95			
Machina	13,547	95			
Nangere	25,525	96			
Nguru	21,276	96			
Potiskum	56,119	96			
Tarmua	16,657	95			
Yunusari	27,865	96			
Yusufari	22,002	96			

The overall sample for each LGA will be complemented by a 60% buffer, to account for invalid responses, phone numbers no longer active, respondents being unavailable or not belonging to the target population group. Data will be collected by three teams of about 50 enumerators each, each covering one of the three BAY states. Each team will be using a dedicated kobo form (three states, three teams, three forms) for easier tracking of data collection progress and smoother follow up. Each team, under the supervision of a Field Supervisor and a Field Assistant, will be subdivided into smaller sub-teams, each led by a team leader.

To avoid collecting data from the same households twice, REACH enumerators will enquire whether any member of the respondent's household has already been interviewed for the MSNA this year; additional checks will be carried out during

data cleaning, such as flagging households whose demographic composition matches perfectly. If all eligibility criteria are met by the respondent and consent is given, REACH enumerators will proceed with the interview, asking questions whilst filling in the survey on the Kobo platform. Forms will be submitted to a secure server at the end of every day of data collection, with the data being automatically deleted from the enumerator's phone.

Progress will be tracked on dedicated tracking sheets by Field Supervisors, on a daily basis, and the raw data will be anonymised by the Database Officer before any data cleaning can take place. Data cleaning and follow up will also take place daily (see Data Cleaning SOPs, Annex 1, for more details).

This methodology carries an inherent bias in that only households who own a mobile phone will be targeted. REACH will be able to use the MSNA 2019 results to compare the need profiles of households owning a phone versus households not owning one, to provide a better picture of the bias potentially introduced by this methodology. Results will be caveated accordingly.

In case a sufficient number of respondents was not reached by phone in a given LGA (LGA-level representativeness of results not achieved), REACH proposes to adopt non-probability quota sampling as a last resort option, as outlined at the end of this section.

#### LGAs where field presence is needed

A two-stage random sampling strategy will be applied in the remaining physically accessible LGAs. Physically inaccessible locations in these LGAs have been first identified, as detailed in section 3.2, and excluded from the sampling frame.

Population estimates in physically accessible locations are used to determine how many of the total surveys should be conducted in which location; this is based on VTS and IOM datasets (March and April 2020), which provide estimates of IDP and non-displaced population numbers in identified settlements, camps and informal sites, as well as estimated number of returnees at the LGA-level. Each primary sampling unit is considered to be a location for one specific population group: for example, the same community with both IDP and non-displaced populations would be considered as two separate locations. This to ensure that representativeness of results for each population group can be achieved at the state level upon aggregation.

There is an information gap on which locations have returnee presence, and this target population group may not be present in some of the selected locations. A 10% buffer will be added to each target sample to account for these instances, as well as for non-responses and invalid submissions. Further, teams will continue with data collection and plans to correct incomplete targets, if necessary, during wrap-up activities at the end of data collection.

The target samples are defined aiming for findings to be representative at the LGA level for the overall population, with a confidence level of 95% and 10% margin of error. The precise level of representativeness is determined post data cleaning and quality checks, and will be indicated on the final published outputs.

State	LGA	Total estimated number of households (accessible areas only)	Target sample
Adamawa	Toungo	6,354	95
	Bama	23,193	96
	Damboa	21,622	96
	Dikwa	19,283	96
	Gubio	3,238	93
Borno	Gwoza	44,666	96
	Kala/Balge	12,380	95
	Mafa	4,057	95
	Magumeri	2,189	92
	Mobbar	5,970	95

	Monguno	50,371	96
	Ngala	12,363	96
Yobe	Geidam	7,440	95

Once the number of interviews per location has been determined, two different methods for household selection will be applied depending on the type of location:

- For big urban agglomerations (BUA, as defined in the VTS dataset), and formal camps (as defined in the IOM sites dataset), whose footprint/area is already known, random GPS points are created across the location's footprint, weighted based on population density should this vary across the target area; partner staff in the field are provided with KMZ files to use with maps.me, and printed maps, to navigate to the target locations the household located nearest to the GPS point is then interviewed\*
- For other types of location (hamlets, small settlements, informal camps and sites), whose footprint is not known, systematic random selection on site will be implemented. Partner staff will be provided with KMZ files to use with maps.me, and printed maps, to reach the target locations; once they arrive at the location, they will be using the pen spinning method, with the aid of a fixed parameter *n* pre-determined using the following calculation:

#### $n = (a \bmod b) + 4$

*a* is the total number of households at the location; *b* is the number of households to be interviewed They will stand in the centre of the settlement/camp, spin a pen, then follow the pen's direction walking to the outer limit of the settled area, then turn around and target every *n*<sup>th</sup> household encountered while walking back towards their starting point\*

\*If the household is present at the target location, but refuses to give consent, the household will be counted as part of the non-responses; if the household is not present or no adult representative is available to participate in the interview, the enumerators should target the closest household to the right; if the household is present but not from the target population group, the enumerators should target the closest household from the relevant population group.

Once the household has been identified, partner staff in the field will either

- Call a dedicated REACH enumerator and hand the phone to the respondent for the interview to be conducted remotely – this will be the case in LGAs where network coverage is good, based on information provided by field officers and KIs, but mobile phone ownership rates are low or unknown, based 2019 MSNA data. These LGAs are: Toungo, Geidam, Bama, Gwoza and Kala/Balge – Assisted data collection
- Conduct the interview in person, filling a Kobo form on their smartphones this will be the case in LGAs where network coverage is poor, based on information provided by field officers and KIs. These LGAs are: Damboa, Dikwa, Gubio, Mafa, Magumeri, Mobbar, Monguno, Ngala – Face to face data collection

Partners present and already active in target LGAs will be mobilised; they will be providing "full" data collection resources (enumerators, cars, technical and health equipment), with REACH covering phone credit expenses and providing the necessary training on tool, household selection methods and relevant health and safety SOPs. Partners will be working only in LGAs where they are already present and have on-going activities; this to minimise inter-LGA travel as a preventive measure in the context of the current COVID-19 pandemic.

REACH focal points will be designated, who will be responsible for coordinating with partners in the field and conduct daily briefs and debriefs. Partner enumerators will be required to submit either a copy of the preliminary consent form to the REACH focal point (Assisted data collection) or the completed survey to the secure Kobo server (Face to face data collection) on a daily basis. Progress will be tracked on dedicated tracking sheets by REACH Field Supervisors, on a daily basis, and the raw data will be anonymised by the Database Officer before any further data cleaning can take place. Data cleaning and follow up will also take place daily (see Data Cleaning SOPs, Annex 1, for more details).

## Last resort option: non-probability quota sampling

This methodology will be used as a last resort solution whenever the strategies described above were not successful in meeting target samples, due to an inability to obtain phone directories for random sampling, low response rates to remote interviews or changes in security and COVID-19-related restrictions.

A target number of households to interview (minimum 30 households per LGA, and further adjusted as relevant based on population size) will be defined as follows:

- Total LGA population < 50,000 households: 30 interviews
- Total LGA population >= 50,000 households & Total LGA population < 100,000 households: 60 interviews
- Total LGA population >= 100,000 households: 90 interviews

REACH will aim to interview an equal number of male and female headed households for each population group of interest; for example, it the target number of interviews is 30, 10 interviews per population group will be conducted, half of which with female headed households. Phone contacts of households residing in these areas will be identified through existing key informant networks and snowball sampling. REACH enumerators will then conduct phone interviews at the household level, and fill in the Kobo survey, until the target number of households is reached.

Although findings would be indicative only, they also would be more accurate than community-level estimates based on key informant interviews. Further, although not generalizable with a known level of statistical precision, non-probability sampling can still generate indicative findings with some level of representation if participant selection is done well.

## 3.5 Primary Data Collection: Contingency plan

The main methodology outlined in section 3.4 heavily relies on the availability of a phone directory for Adamawa, Borno and Yobe states (BAY states) in North East Nigeria as a basis for random sampling; this would be made available either by the Nigerian Communications Commission or the main telecommunications companies operating in this area. In case a directory was not made available by 28 June 2020, REACH proposes the below contingency approach.

Table 2. Categorisation of LGAs, based on the applicability of a fully remote methodology and based on physical access considerations.

State	LGAs that cannot be assessed (no remote and no physical access)	LGAs that cannot be assessed remotely (partial physical access)	LGAs that can be assessed remotely
Adamawa	1	Toungo	Demsa, Fufore, Ganye, <u>*Gombi</u> , Girei, Guyuk, <u>*Hong</u> , Jada, <u>*Lamurde</u> , <u>*Madagali</u> , <u>*Maiha</u> , Mayo-Belwa, <u>*Michika</u> , Mubi North, Mubi South, <u>*Numan</u> , Shelleng, Song, Yola North, Yola South
Borno	Abadam, Guzamala, Kukawa, Marte, Nganzai	Bama, Damboa, Dikwa, Gubio, Gwoza, Kala/Balge, Mafa, Magumeri, Mobbar, Monguno, Ngala	<u>*Askira/Uba,</u> Bayo, Biu, Chibok, Hawul, <u>*Jere,</u> Kaga, <u>*Konduga,</u> Kwaya Kusar, <u>*Maiduguri,</u> Shani
Yobe	/	Geidam	Bade, Bursari, Damaturu, Fika, Fune, Gujba, Gulani, Jakusko, Karasuwa, Machina, Nangere, Nguru, Potiskum, Tarmua, Yunusari, Yusufari

\* These LGAs will be henceforth referred to as **priority LGAs**; they were identified based on their 2019 Multi-Sectoral Needs Index (MSNI) score, as well as the proportion of IDPs or Returnees households out of the total LGA population. Specifically, these LGAs were above the 25th percentile for percentage of households scoring a 3 or 4 (indicating severe or extremely severe needs) on the 2019 MSNI, and are above the 50th percentile for proportion of IDP or Returnee households out of their total population. Gombi LGA, although not fitting these criteria, was recommended as an addition to the list of priority LGAs by the AAWG, due to the dynamic context that played out in this area since the last MSNA assessment was conducted.

#### LGAs that cannot be assessed remotely and Priority LGAs

A two-stage cluster sampling strategy will be applied in **physically accessible LGAs that cannot be assessed remotely**, as defined in the original ToR, as well as in **priority LGAs** where the need to address information gaps is highest. For these LGAs, inaccessible areas due to security reasons or COVID-19 movement restrictions will be identified on a map by field officers, as well as based on Access Working Group and other security-related documents, and excluded from the sample. The feasibility of field visits will be revisited on a daily basis during data collection.

Population estimates in physically accessible locations are used to determine how many of the total surveys should be conducted in which location; this is based on VTS and IOM datasets (March and April 2020), which provide estimates of IDP and non-displaced population numbers in identified settlements, camps and informal sites, as well as estimated number of returnees at the LGA-level. A two-stage cluster sampling design will be utilized. The primary sampling unit is defined as the location (settlement, site, camp, collective centre), and the secondary sampling unit is the households within those locations. Each primary sampling unit is considered to be a location for one specific population group. This means that the same community with both IDPs, and non-displaced populations would be considered as two separate units. First locations are randomly selected within a given LGA based on probability proportional to the number of households living at the location, then households are selected within each randomly sampled location, based on the number of times that location has been selected during the previous stage. A minimum of 4 surveys per location is pre-defined, for logistical reasons.

There is an information gap on which locations host returnee households, and this target population group may not be present in some of the selected locations. A 10% buffer will be added to each target sample to account for these instances, as well as for non-responses and invalid submissions. Further, teams will continue with data collection and plans to correct incomplete targets, if necessary, during wrap-up activities at the end of data collection. The target samples are defined aiming for findings to be representative at the LGA level for the overall population, with a confidence level of 95% and 10% margin of error. The overall sample for each LGA will be complemented by a 10% buffer, to account for invalid responses, respondents being unavailable or not belonging to the target population group. The precise level of representativeness is determined post data cleaning and quality checks, and will be indicated on the final published outputs.

	LGA	Total estimated number of households (accessible areas only)	Target sample
	Gombi	13,815	432
	Hong	5,614	342
	Lamurde	900	106
Adamawa	Madagali	3,801	111
Audillawa	Maiha	15,861	104
	Michika         10,339           Numan         6,350           Toungo         6,354	10,339	173
	Numan	6,350	173
	Toungo	6,354	115
	Askira/Uba	4,462	105
	Bama	23,193	104
	Damboa	21,622	166
	Dikwa	19,283	133
	Gubio	3,238	103
Borno	Gwoza	44,666	137
DOINO	Jere	38,762	151
	Kala/Balge	12,380	104
	Konduga	28,274	155
	Mafa	4,057	116
	Magumeri	2,189	106
	Maiduguri	140,797	151

	Mobbar	5,970	108
	Monguno	50,371	144
	Ngala	12,363	173
Yobe	Geidam	7,440	209

Once the number of interviews per location has been determined, two different methods for household selection will be applied depending on the type of location:

- For big urban agglomerations (BUA, as defined in the VTS dataset), and formal camps (as defined in the IOM sites dataset), whose footprint/area is already known, random GPS points are created across the location's footprint, weighted based on population density should this vary across the target area; partner staff in the field are provided with KMZ files to use with maps.me, and printed maps, to navigate to the target locations the household located nearest to the GPS point is then interviewed\*
- For other types of location (hamlets, small settlements, informal camps and sites), whose footprint is not known, systematic random selection on site will be implemented. Partner staff will be provided with KMZ files to use with maps.me, and printed maps, to reach the target locations; once they arrive at the location, they will be using the pen spinning method, with the aid of a fixed parameter *n* pre-determined using the following calculation:

#### $n = (a \mod b) + 4$

*a* is the total number of households at the location; *b* is the number of households to be interviewed They will stand in the centre of the settlement/camp, spin a pen, then follow the pen's direction walking to the outer limit of the settled area, then turn around and target every *n*<sup>th</sup> household encountered while walking back towards their starting point\*

\*If the household is present at the target location, but refuses to give consent, the household will be counted as part of the non-responses; if the household is not present or no adult representative is available to participate in the interview, the enumerators should target the closest household to the right; if the household is present but not from the target population group, the enumerators should target the closest household from the relevant population group.

Once the household has been identified, partner staff in the field will either

- Conduct the interview in person, filling a Kobo form on their smartphones this will be the case in LGAs where
  network coverage is poor, based on information provided by field officers and KIs, or in LGAs where REACH
  enumerators are present and can conduct interviews without additional partner support. These LGAs are: Damboa,
  Dikwa, Gubio, Jere, Konduga, Mafa, Magumeri, Maiduguri, Mobbar, Monguno, Ngala Face to face data collection
- Call a dedicated REACH enumerator and hand the phone to the respondent for the interview to be conducted remotely – this will be the case in the remaining LGAs where network coverage is good, based on information provided by field officers and KIs, but mobile phone ownership rates are low or unknown, based 2019 MSNA data. These LGAs are: Gombi, Hong, Lamurde, Madagali, Maiha, Michika, Numan, Toungo, Askira/Uba, Bama, Gwoza, Kala/Balge, Geidam – Assisted data collection

Partners present and already active in target LGAs will be mobilised; they will be providing "full" data collection resources (enumerators, cars, technical and health equipment), with REACH covering phone credit expenses and providing the necessary training on tool, household selection methods and relevant health and safety SOPs.

REACH focal points will be designated, who will be responsible for coordinating with partners in the field and conduct daily briefs and debriefs. Partner enumerators will be required to submit either a copy of the preliminary consent form to the REACH focal point (Assisted data collection) or the completed survey to the secure Kobo server (Face to face data collection) on a daily basis. Progress will be tracked on dedicated tracking sheets by REACH Field Supervisors, on a daily

basis, and the raw data will be anonymised by the Database Officer before any further data cleaning can take place. Data cleaning and follow up will also take place daily.

## All remaining LGAs (non-priority LGAs that can be assessed remotely)

The last resort option outlined in section 3.4 will be applied in all remaining LGAs. A target number of households to interview (minimum 30 households per LGA, and further adjusted as relevant based on population size) will be defined as follows:

- Total LGA population < 50,000 households: 30 interviews
- Total LGA population >= 50,000 households & Total LGA population < 100,000 households: 60 interviews
- Total LGA population >= 100,000 households: 90 interviews

REACH will aim to interview an equal number of male and female headed households for each population group of interest; for example, it the target number of interviews is 30, 10 interviews per population group will be conducted, half of which with female headed households. Phone contacts of households residing in these areas will be identified through existing key informant networks and snowball sampling; to this end a snowball sampling component would be added to the Kobo tool. REACH enumerators would use their network of KIs to compile an initial list of contacts between 29 June and 3 July, and would then proceed to conduct phone interviews at the household level, filling in the Kobo survey and continuing snowball sampling, until the target number of households is reached.

State	LGA	Total estimated number of households	Target sample
	Demsa	29,229	30
	Fufore	41,367	30
	Ganye	32,212	30
	Girei	43,180	30
	Guyuk	24,261	30
Adamawa	Jada	32,935	30
Audillawa	Mayo-Belwa	33,139	30
	Mubi North	47,699	30
	Mubi South	48,779	30
	Shelleng	21,979	30
	Song	46,546	30
	Yola North	37,048	30
	Вауо	34,110	30
	Biu	64,869	60
	Chibok	31,324	30
Adamawa Borno Yobe	Hawul	58,774	60
	Kaga	36,063	30
	Kwaya Kusar	27,055	30
	Shani	45,892	30
	Bade	23,301	30
	Bursari	18,114	30
	Damaturu	37,975	30
	Fika	29,403	30
	Fune	35,991	30
Voho	Gujba	34,111	30
TODE	Gulani	26,647	30
	Jakusko	21,691	30
GireiGuyukJadaMayo-BelwaMubi NorthMubi SouthShellengSongYola NorthBayoBiuChibokHawulKagaKwaya KusarShaniShaniBadeBursariDamaturuFikaFuneGujbaGulaniJakuskoKarasuwaMachinaNangereNguru	Karasuwa	13,018	30
	13,547	30	
	Nangere	25,525	30
	Nguru	21,276	30

	Potiskum	56,119	60
	Tarmua	16,657	30
	Yunusari	27,865	30
	Yusufari	22,002	30

The overall sample for each LGA will be complemented by a 10% buffer, to account for invalid responses. Data will be collected by three teams of about 50 enumerators each, each covering one of the three BAY states, with enumerators working either from home or in call centres set up within their LGA or residence. Each team will be using a dedicated kobo form (three states, three teams, three forms) for easier tracking of data collection progress and smoother follow up. Each team, under the supervision of a Field Supervisor and a Field Assistant, will be subdivided into smaller sub-teams, each led by a team leader.

If all eligibility criteria are met by the respondent and consent is given, REACH enumerators will proceed with the interview, asking questions whilst filling in the survey on the Kobo platform. Forms will be submitted to a secure server at the end of every day of data collection, with the data being automatically deleted from the enumerator's phone.

Although findings would be indicative only, they also would be more accurate than community-level estimates based on key informant interviews. Further, although not generalizable with a known level of statistical precision, non-probability sampling can still generate indicative findings with some level of representation if participant selection is done well.

## 3.6 Data Processing & Analysis

Data entry and cleaning process: As network availability allows, data from the primary data collection will be uploaded on a daily basis by enumerators, and reviewed by the REACH Database Officer (DO) who will conduct spatial verification and then anonymise the data before further data cleaning can take place; the DO will use an R script for preliminary data cleaning, followed by additional manual cleaning conducted by Assessment Officers (AOs). Cleaning logs will then be communicated to Field Supervisors for follow up. Detailed Data Cleaning SOPs are provided in Annex to this document.

**Analytical framework**: REACH will use, to the extent possible, the updated analytical framework developed by the Joint Inter-Sectoral Analysis Group (JIAG) at the global level. Using this analytical framework will ensure that results from the MSNA will most adequately respond to the information needs of the humanitarian community.

**Data analysis process:** REACH will primarily conduct the data analysis from the clean dataset based on **REACH global guidance**, in line with the JIAG analytical framework, and using R software. A preliminary presentation of results will be given to the AAWG, for feedback; REACH will incorporate the feedback and, with OCHA, facilitate a Joint Analysis Workshop within the humanitarian community to try and jointly contextualize results from the data analysis based on partners' experience. Results will be provided for each LGA assessed, for the overall population; wherever aggregation of LGA-level results is possible, results will also be presented at the state level, for each population group of interest.

# 4. Roles and responsibilities

Table 2: Description of roles and responsibilities

Task Description	Responsible	Accountable	Consulted	Informed
Research design	Assessment Officer	Country Focal Point	GIS Officer, HQ RDD Unit	OCHA, Sectors
Supervising data collection	Field Manager	Assessment Officer	HQ RDD Unit	OCHA, Country Focal Point
Data processing (checking, cleaning)	Database Officer	Assessment Officer, GIS Officer	HQ RDD Unit	OCHA, Country Focal Point
Data analysis	Assessment Officer, GIS Officer	Country Focal Point	HQ RDD Unit	OCHA, Sectors
Output production	Assessment Officer	Country Focal Point	HQ Reporting team	OCHA
Dissemination	Assessment Officer	Country Focal Point	HQ Comms	OCHA, Sectors, Donor
Monitoring & Evaluation	Assessment Officer	Country Focal Point	HQ RDD Unit	
Lessons learned	Assessment Officer	Country Focal Point	HQ RDD Unit	Partners

**Responsible:** the person(s) who executes the task **Accountable:** the person who validates the completion of the task and is accountable of the final output or milestone

Consulted: the person(s) who must be consulted when the task is implemented

Informed: the person(s) who need to be informed when the task is completed

# 5. Data Management Plan

Data management plan available upon request

# 6. Data Analysis Plan

Research questions	IN #	Data collection method	Indicator / Variable	Questionnaire Question	Questionnaire Responses	Data collection level
		HH Interview	Enumerator code	Enumerator code	Enter name	Household
		Remote HH interview onlv	Target Phone number	What's the phone number you're trying to contact?	Enter number	HH
		HH Interview	State	In which state is the household located?	Admin list	HH
		HH Interview	LGA	In which LGA is the community located?	Admin list	HH
NA		In person HH interview only	Ward	In which ward is the community located?	Admin list	HH
		In person HH interview only	Community / Site	What is the name of the community / site?	Admin list	HH
		In person HH interview only	Household coordinates	Log gps coordinates	Log gps coordinates	HH
	A.1.1	HH Interview	Household status	Which statement best describes your household's situation since January 2009?	<ol> <li>Household is displaced and currently not living in village or area of origin</li> <li>Household has never been displaced from village or area of origin</li> <li>Household was displaced but has since returned and is currently in village or area of origin</li> </ol>	НН
	A.1.2	HH Interview	Respondent age	What is your age in years?	Enter integer	HH
	A.1.3	HH Interview	Respondent gender	What is the gender of the respondent?	Male,Female	HH
	A.1.4	HH Interview	Head of household age (if different from respondent)	What is the age of the head of household?		
	A.1.5	HH Interview	Head of household gender (if different from respondent)	What is the gender of the head of houehold?		
What are the demographic characteristic s of the crisis- affected	A.2.1	HH Interview	Marital status	What is the marital status of the head of household?	1. Single 2. Married 3. Divorced 4. Widowed 98. NR 99. DK	HH
population, including vulnerabilities of the households?	A.2.2	HH Interview	% of IDP HHs by Area of Origin (AoO)	What is the State of origin of this household?	1. Adamawa 2. Borno 3. Yobe 4. Other (Please specify) 98. No Response 99. Don't know	HH
	A.2.3	HH Interview	% of IDP HHs by Area of Origin (AoO)	[If 1-3 selected] What is the LGA of origin of this household?	Admin list	HH
	A.2.4	HH Interview	Household size	Currently how many household members are there in your household (including the respondent)?	Enter integer	HH
	A.2.5	HH Interview	Names of household members (for reference in looping questions)	Please enter the name of the household member	Enter text	HH
	A.2.6	HH Interview	Household gender composition	Please enter the sex of the household member	Male, Female	HH

	A.2.7	HH Interview	Houshold age group composition	Please enter age of the household member in years	Enter integer	HH
	A.2.8	HH Interview	Household age group composition	(for children <1 year old) What is the age of the child in months?	Enter integer	HH
	A.2.8	HH Interview	% of households with at least one members with vulnerabilities	Does this household member fit any of the following criteria or situations?	<ol> <li>Person with chronic illness [all household members]</li> <li>Person with mental disability [all household members]</li> <li>Person with physical disability [all household members]</li> <li>Person with physical disability [all household members]</li> <li>Pregnant woman or girl [only female household members aged 12-49 years]</li> <li>Breastfeeding woman or girl [only female household members aged 12-49 years]</li> <li>Caretaker of child under 5 years old [only household members aged 12 years or older]</li> <li>Child separated from parents or usual guardian [only household members aged 0-17 years]</li> <li>[If yes] Child separated from family, both immediate and extended [only household members aged 0-17 years]</li> <li>Married household member [only household members aged 10-17 years]</li> <li>None of the above (cannot select with any other option)[No other options can be selected]</li> <li>NR [No other options can be selected]</li> <li>DK [No other options can be selected]</li> </ol>	HH
	A.2.1	HH Interview	% of children attending formal education before the COVID-19 outbreak	For the current school year, how many children in the household were attending formal school in preschool through secondary school regulady BEFORE Ramadan?	Pre-school boys (usually 3-5 years) Pre-school girls (usually 3-5 years) Primary school boys (usually 6-11 years)	HH
					Primary school girls (usually 6-11 years) Junior secondary school boys (usually (12-14 years) Junior secondary school girls (usually 12-14 years) Senior secondary school boys (usually 15-17 years) Senior secondary school girls (usually 15-17 years)	
What are the education- specific humanitarian needs of the	A.2.2	HH Interview	% of children attending non- formal education before the COVID- 19 outbreak	For the current school year, how many children were attending non-formal learning opportunities (like Islamic school, Accelerated Learning Programme, TaRL, EGRA/Kari, RANA, HASKI or vocational learning) BEFORE Ramadan?	Integer	HH
affected population, including level of access to education?	A.2.3	HH Interview	% of children continuing learning activities remotely (where schools are closed)	Of the [number] children in the household who were attending formal school regularly before Ramadan, how many have been accessing remote learning or using a home learning program since leaving school?	Integer	HH
	A.2.4	HH Interview	% of children not reached by remote learning, by reason for not learning remotely	[If reported # of children learning remotely is less than # of school-aged children in household attending formal school] For the children who attended formal school who are not learning remotely or using a home learning program, why are they not learning remotely?	<ol> <li>Can no longer afford to keep child(ren) learning</li> <li>Child(ren) started working instead</li> <li>Child(ren) helping at home / on farm</li> <li>No materials to learn remotely</li> <li>No caretaker/teacher is available to guide education</li> <li>Illness of child(ren)</li> <li>No materials for child(ren) with disabilities</li> <li>Marriage and/or pregnancy of girl child(ren)</li> <li>Remote learning schedule conflicts with household schedule for girls</li> <li>Other (specify)</li> <li>NR [No other options can be selected]</li> <li>DK [No other options can be selected]</li> </ol>	HH

	A.2.5	HH Interview	% of households by preferred education support modality	If available, what type of support would help your child(ren) with attending school or participating in regular learning activities?	<ol> <li>No support needed / wanted (cannot select with any other option)[No other options can be selected]</li> <li>Payment of school fees</li> <li>Cash for school supplies/equipment (bags, pencils, books, uniforms)</li> <li>Cash for transportation to school</li> <li>Cash for children's food</li> <li>Cash to offset opportunity cost of child working</li> <li>Direct provision of school supplies/equipment (bags, pencils, books, uniforms)</li> <li>Direct provision of transportation</li> <li>Direct provision of water and food for children</li> <li>Healthcare at school</li> <li>Provision of alternative learning curriculum</li> <li>Assistance for children with disabilities</li> <li>Assistance for children of minority groups</li> <li>Other (specify)</li> <li>NR [No other options can be selected]</li> </ol>	HH
	A.2.6	HH Interview	% of children aged 5-17 dropping out of school in the previous year	Prior to Ramadan, how many children in the household (attending primary school through secondary school) dropped out of school during the current school year?	Integer. Cannot exceed # of children aged 5- 17. Enter 0 if none, 998 if NR, 999 if DK.	HH
	A.3.1	HH Interview	% of households exclusively breastfeeding children 0-5 months	[For households with children 0-5 months] Which of the following foods or liquids did [name(s) of child(ren) aged 0-5 months] receive yesterday, during the day and at night?	<ol> <li>No liquid or food in the past 24 hours (cannot select with any other option)[No other options can be selected]</li> <li>Baby formula</li> <li>Fermented milk (kindirmo)</li> <li>Animal milk (fresh, powdered, or canned)</li> <li>Breastmilk</li> <li>Fruit juice or juice drink</li> <li>Clear broth or thin porridge</li> <li>Water</li> <li>Solid, semi-solid or soft food</li> <li>Other</li> <li>NR [No other options can be selected]</li> <li>DK [No other options can be selected]</li> </ol>	ΗΗ
What are the nutrition- specific humanitarian needs of the crisis-	A.3.2	HH Interview	% of households with continued breastfeeding of all children 1-2 years old (children 12-23 months)	[For households with children 12-23 months] Did [name(s) of child(ren) aged 12-23 months] have any breast milk yesterday, during the day or at night?	1. Yes, all were breastfed 2. No, not all were breastfed 98. NR 99. DK	HH
affected population, including challenges experienced	A.3.3	HH Interview	% of households who received nutrition counselling	[For households with PLW or caretaker(s) of children under 5 years] In the past 30 days, did [name(s) of PLW and/or caretaker(s)] receive any nutrition counselling from a health facility or community health worker?	1. Yes 2. No 98. NR 99. DK	HH
children and mothers?	A.3.4	HH Interview	% of children currently enrolled in OTP services	[For households with children 6-59 months] Of [name(s) of child(ren) 6-59 months], how many are currently enrolled in OTP (receiving RUTF like Plumpy Nut)?	Integer	HH
	A.3.5	HH Interview	% of households receiving nutrition treatment when needed	[For households with children 0-59 months] If you saw [name(s) of child(ren) 0-59 months] losing weight, where would you take them?	1. Nowhere 2. Hospital 3. PHC (primary health care) 4. Mobile/outreach clinic 5. Village outreach worker 6. Private doctor 7. Shop/vendor 8. Traditional practitioner 9. Pharmacy/dispensary 97. Other 98. No response 99. Don't know	НН

	A.3.6	HH Interview	% of households receiving nutrition treatment when needed	[If 2-97 are selected] Have you gone to the [option selected] for nutrition treatment in the past 30 days?	<ol> <li>Yes, and the facility/provider was open and functioning</li> <li>Yes, but the facility/provider was closed, not functioning or did not have supplies at the time</li> <li>No, the facility/provider was closed, not functioning or did not have supplies at the time</li> <li>No, facility/provider was open and functioning but chose to go to a different facility/provider at the time</li> <li>No, could not contact facility/provider at the time</li> <li>No, couldn't move due to the lockdown</li> <li>No, did not need nutrition treatment in the past 30 days</li> <li>No, afraid of COVID-19</li> <li>Other (specify)</li> <li>NR [No other options can be selected]</li> <li>DK [No other options can be selected]</li> </ol>	HH
	A.4.1	HH Interview	% of households where all children 6 months-15 years have a vaccination card	Do you have a National Child Immunization Record, immunization records from a private health provider or any other document where [name of child 0 months to 15 years]'s vaccinations are written down?	1. Yes 2. No 98. NR 99. DK	HH
	A.4.2	HH Interview	% of households where all children 6 months-15 years have a vaccination card	[Face-to-face interview only] May we see the card(s)/record(s)?	1. Yes, card(s) shown 2. No, not shown	HH
	A.4.3	HH Interview	% of households where all children <1 received a PENTA3 vaccination	Has [name of child 0-23 months old] ever received a PENTA vaccine?	1. Yes 2. No 98. NR 99. DK	HH
What are the health- specific humanitarian needs of the	A.4.4	HH Interview	% of households where all children <1 received full coverage of PENTA3, by administrative unit	[If yes] How many times was the PENTA vaccine received?	1. Once 2. Twice 3. Three times 98. NR 99. DK	Η
affected population, including challenges of accessing health	A.4.5	HH Interview	% of households where all children <1 received full vaccination coverage for polio	Has [name of child 0-23 months old] ever received a polio vaccine? If yes, was it administered through a drop by mouth or through an injection?	1. Yes, by mouth 2. Yes, by injection 3. No 98. NR 99. DK	HH
households and what is the vaccination status of children?	A.4.6	HH Interview	% of households where all children <1 received full vaccination coverage for polio	[If yes] How many times was the polio vaccine received?	<ol> <li>Once by mouth or by injection</li> <li>Twice by mouth</li> <li>Three time by mouth</li> <li>Four times by mouth</li> <li>98. NR</li> <li>99. DK</li> </ol>	HH
	A.4.7	HH Interview	% of households where all children 6 months-15 years have received measles vaccination	Has [name of child 6 months to 15 years old] ever received a measles vaccine?	1. Yes 2. No 98. NR 99. DK	HH
	A.4.8	HH Interview	% of households where all children 6 months-15 years have received measles vaccination	[If yes] How many times was the measles vaccine received?	1. Once 2. Twice 98. NR 99. DK	HH
	A.4.9	HH Interview	% of households where all children aged 12-23 months received BCG containing vaccine at any time before the survey	Has [name of child 0-23 months old] ever received a BCG vaccine?	1. Yes 2. No 98. NR 99. DK	ΗΗ

A.4.1 0	HH Interview	% of households that can access primary healthcare within one hour's walk from dwellings	How long does it take you to reach the nearest healthcare facility by walking?	1. 0-14 mins 2. 15-29 mins 3. 30-59 mins 4. 1-3 hours 5. More than 3 hours 98. NR 99. DK	HH
A.4.1 1	HH Interview	% of households with most recent delivery attended by a skilled birth attendant	[If household has female member(s) aged 12- 49 years] Has anyone in your household given birth in the past year?	1. Yes 2. No 98. NR 99. DK	HH
A.4.1 2	HH Interview	% of households with most recent delivery attended by a skilled birth attendant	[If yes] In the most recent delivery, who helped with this birth?	<ol> <li>Skilled birth attendant (doctor, nurse, midwife)</li> <li>Other health care worker (health volunteer, CHEW)</li> <li>Traditional birth attendant</li> <li>Other women in the community</li> <li>No support</li> <li>Other</li> <li>No response</li> <li>Don't know</li> </ol>	HH
A.4.1 3	HH Interview	% of households with members ill in two weeks prior	Has anyone in your household been sick in the past two weeks?	1. Yes 2. No 98. NR 99. DK	HH
A.4.1 4	HH Interview	% of households with members ill in two weeks prior	[If yes] What symptoms or illnesses did the person have?	<ol> <li>Fever</li> <li>Chills</li> <li>Cough, dry and persistent</li> <li>Cough, wet and productive</li> <li>Shortness of breath or trouble breathing</li> <li>Sore throat</li> <li>Diarrhoea</li> <li>Skin infection</li> <li>Eye infection</li> <li>Dizziness</li> <li>Vomiting</li> <li>Injury</li> <li>Persisting headache</li> <li>Stomach pains</li> <li>Heart problem</li> <li>Tooth ache</li> <li>Measles</li> <li>Cholera</li> <li>Sexually transmitted infection</li> <li>Mental illness symptoms</li> <li>Other (specify)</li> <li>NR [No other options can be selected]</li> </ol>	HH
A.4.1 5	HH Interview	% of households who have or would seek treatment at healthcare facilities	Where would be the first place you would go for treatment, if a household member was sick?	<ol> <li>Would not seek treatment in a health facility</li> <li>Hospital</li> <li>PHC (primary healthcare)</li> <li>Mobile / outreach clinic</li> <li>Village outreach worker</li> <li>Private doctor</li> <li>Patent medicine store / Chemist</li> <li>Traditional practitioner</li> <li>Pharmacy / Dispensary</li> <li>Other</li> <li>NR</li> <li>DK</li> </ol>	ΗΗ

What are the WASH- specific humanitarian needs of the crisis- affected population, including issues related to water quantity and use, hygiene	A.5.1	HH Interview	% of households having access to an improved water source	What is the main source of water used by your household for drinking?	<ol> <li>Public tap/standpipe</li> <li>Handpumps/boreholes</li> <li>Protected well</li> <li>Unprotected well</li> <li>Water seller/kiosks</li> <li>Piped connection to house (or neighbor's house)</li> <li>Rain water collection</li> <li>Bottled water, water sachets</li> <li>Tanker trucks</li> <li>Surface water (lake, pond, dam, river)</li> <li>Other (please specify)</li> <li>NR</li> <li>Don't know</li> </ol>	НН
access to sanitation for households?	A.5.2	HH Interview	% of households by time (minutes) taken to fetch water (round trip by walking, queuing and time needed to fetch water)	How long does it take to go to your main water source, fetch water, and return (including queuing at the water source)?	<ol> <li>Water on premises</li> <li>Less than 5 min to fetch and return</li> <li>Between 5 and 15 min to fetch and return</li> <li>Between 16 and 30 min to fetch and return</li> <li>More than 30 min to fetch and return</li> <li>98. Don't want to say/no answer</li> <li>99. Don't know</li> </ol>	HH
	A.5.3	HH Interview	% of households having access to a sufficient quantity of water for drinking, cooking, bathing, washing or other domestic use	Does your household currently have enough water to meet the following needs?	<ol> <li>Have enough water for drinking</li> <li>Have enough water for cooking</li> <li>Have enough water for personal hygiene (washing or bathing)</li> <li>Have enough water for other domestic purposes (cleaning house, floor, etc.)</li> <li>Not enough water to meet any of the above needs (cannot select with any other option)[No other options can be selected]</li> <li>NR [No other options can be selected]</li> <li>DK [No other options can be selected]</li> </ol>	НН
	A.5.4	HH Interview	% of households having access to a functional and improved sanitation facility	What kind of sanitation facility (latrine/toilet) does your household usually use?	<ol> <li>Flush or pour-flush toilet</li> <li>Pit latrine without a slab and/or platform</li> <li>Pit latrine with a slab and platform</li> <li>Open hole</li> <li>Pit VIP toilet</li> <li>Bucket toilet</li> <li>Plastic bag</li> <li>Hanging toilet/latrine</li> <li>None of the above, open defecation</li> <li>Other (specify)</li> <li>Don't want to say/no answer</li> <li>Don't know</li> </ol>	НН
	A.5.5	HH Interview	% of households sharing sanitation facility - by number of household per sanitation facility	Do you share this sanitation facility with other households?	1. Yes 2. No 98. Don't want to say/no answer 99. Don't know	HH
	A.5.6	HH Interview	% of households sharing sanitation facility - by number of household per sanitation facility	If yes, how many households use this sanitation facility (latrine/toilet)?	# of households	HH
	A.5.7	HH Interview	% of households with access to functioning handwashing facilities, with water and soap available	What kind of handwashing facility do your household members usually use to wash their hands?	<ol> <li>No specific handwashing device (no device at all or only pouring device or simple basin/bucket, with no taps)</li> <li>Sink with tap water</li> <li>Buckets with taps</li> <li>Tippy tap</li> <li>Sani-Taps</li> <li>Other (specify)</li> <li>Don't want to say/no answer</li> <li>Don't know</li> </ol>	НН
	A.5.8	HH Interview	% of households with access to soap	Do you have any soap in your household?	1. Yes 2. No 98. Don't want to say/no answer 99. Don't know	НН

What are the food security humanitarian needs of the crisis- affected population, including issues related to	A.6.1	HH Interview	% of households by most commonly used fuel type for cooking	What is the most common source of fuel used for cooking in your household?	<ol> <li>Firewood</li> <li>Charcoal</li> <li>Kerosene</li> <li>Animal dung</li> <li>Agricultural waste (crop residue)</li> <li>Gas</li> <li>Other</li> <li>NR</li> <li>DK</li> </ol>	HH
food quantity and access, availability of firewood/fuel and practices of agriculture for households?	A.6.2	HH Interview	% of households by means of obtaining firewood/fuel for cooking	What is your primary means of obtaining your main source of fuel?	<ol> <li>Purchase from local seller</li> <li>Collect directly from within the community</li> <li>Collect directly from outside the community</li> <li>Trade goods / items for fuel</li> <li>From organisation assistance / aid</li> <li>Other</li> <li>NR</li> <li>DK</li> </ol>	HH
	A.6.3	HH Interview	% of households with sufficient access to firewood/fuel in the surrounding environment	Do you have sufficient access to firewood/fuel to meet your daily energy needs?	1. Yes 2. No 98. NR 99. DK	HH
	A.6.4	HH Interview	% of households using coping strategies for insufficient fuel	[If no] What does your household do when there is insufficient fuel for your needs?	<ol> <li>Nothing (cannot select with any other option)[No other options can be selected]</li> <li>Reduce number of meals</li> <li>Selling food/rations to buy fuel</li> <li>Undercooking food</li> <li>Use less preferred fuel source (animal dung, etc.)</li> <li>Eating raw food</li> <li>Cook with other families together (communal cooking)</li> <li>Other (specify)</li> <li>NR [No other options can be selected]</li> <li>DK [No other options can be selected]</li> </ol>	HH
	A.6.5	HH Interview	Household Hunger Scale	In the past 30 days, was there ever no food to eat of any kind in your house because of lack of resources to get food?	1. Yes 2. No 98. NR 99. DK	HH
	A.6.6	HH Interview	Household Hunger Scale	How often did this happen in the past 30 days?	1. Rarely (1-2 times) 2. Sometimes (3-10 times) 3. Often (More than 10 times) 98. NR 99. DK	HH
	A.6.7	HH Interview	Household Hunger Scale	In the past 30 days, did you or any household member go to sleep at night hungry because there was not enough food?	1. Yes 2. No 98. NR 99. DK	HH
	A.6.8	HH Interview	Household Hunger Scale	How often did this happen in the past 30 days?	1. Rarely (1-2 times) 2. Sometimes (3-10 times) 3. Often (More than 10 times) 98. NR 99. DK	ΗH
	A.6.9	HH Interview	Household Hunger Scale	In the past 30 days, did you or any household member go a whole day and night without eating anything at all because there was not enough food?	1. Yes 2. No 98. NR 99. DK	HH
	A.6.1 0	HH Interview	Household Hunger Scale	How often did this happen in the past 30 days?	1. Rarely (1-2 times) 2. Sometimes (3-10 times) 3. Often (More than 10 times) 98. NR 99. DK	HH

What are the early recovery humanitarian needs of the crisis- affected population, including issues related to household livelihoods and access to essential public infrastructure ?	A.7.1	HH Interview	% of household by main source of income	In the past 30 days, what were the three main sources of income for the household?	<ol> <li>Salaried work</li> <li>Casual or daily labour</li> <li>Income from own business or commerce.</li> <li>Income from household's agricultural products</li> <li>Government social benefits or assistance</li> <li>Support from family and friends (including remittances)</li> <li>Assistance from organisation(s) (including cash for work)</li> <li>Charitable donations such as zakat (not including assistance from organisations)</li> <li>No income (cannot select with any other option)[No other options can be selected]</li> <li>Other (please specify) 98. NR [No other options can be selected]</li> </ol>	ΗΗ
	A.7.2	HH Interview	% of household by main source of income	How much income was generated by the household in total in the past 30 days?	Integer	HH
	A.7.3	HH Interview	% of households by main coping strategies for lack of income or resources	In the past 30 days, was this income sufficient to address your household's basic needs (such as food, shelter, water and hygiene)?	1. Yes 2. No 98. Don't want to say/no answer 99. Don't know	HH
	A.7.4	HH Interview	% of households by main coping strategies for lack of income or resources	[If no] In the past 30 days, did your household do one of the following things to cope with lack of income / or because you were unable to meet your needs?	<ol> <li>Sell non-productive household or agricultural products assets/good (furniture, television, jewellery, non-breeding animals.)</li> <li>Spend savings</li> <li>Send household members to eat elsewhere</li> <li>Purchase food on credit or borrow food</li> <li>Borrow money</li> <li>Rely on support from family/host family/external assistance</li> <li>Sell productive household or agricultural productive assets, or means of transport (sewing machine, agricultural machinery/vehicles, cars, bicycle, milk/egg producing livestock, breeding livestock, etc.)</li> <li>Withdraw children from school</li> <li>Reduce expenditure on other services like health and education</li> <li>Harvest immature crops (green maize)</li> <li>Consume seed stocks that were to be saved for the next season</li> <li>Decrease expenditure on fertilizer, pesticide, animal feed, veterinary care, etc.G101</li> <li>Sell and or property</li> <li>Beg for money and/or food</li> <li>Engage in dangerous or illegal work/activity (incl. theft, illegal substances dealing, prostitution)</li> <li>Marriage of female household member under the age of 18</li> <li>None of the above (cannot select with any other option)[No other options can be selected]</li> <li>Other (specify)</li> <li>NR [No other options can be selected]</li> <li>Don't know[No other options can be selected]</li> </ol>	HH

	A.7.5	HH Interview	% of households without access to functional basic public facilities and infrastructure, % of household without access to debris/ waste management services, % of households without access to functional community safety & security initiatives	Which of the following public facilities or services are functional and within a 30 minute walk of your household?	<ol> <li>Nigeria registered bank</li> <li>Mobile money platforms</li> <li>Nursery school</li> <li>Primary school</li> <li>Secondary school</li> <li>Internet access available to community members</li> <li>Public facility where community members can meet, such as woman's hall or community centre</li> <li>Market</li> <li>Debris or waste management collection service</li> <li>Police station</li> <li>Community security initiative (apart from Nigeria Police, Nigerian Security and Civil Defense Corps, Nigerian Immigration, National Drug Law Enforcement Agency (NDLEA), Department of State Security (DSS) and Nigerian Customs Service)?</li> <li>Local government building</li> <li>None of the above (cannot select with any other option) [No other options can be selected]</li> <li>NR [No other options can be selected]</li> </ol>	ΗΗ
What are the shelter and NFI-specific humanitarian needs of the crisis- affected population?	A.8.1	HH Interview	% of households with access to a safe and healthy housing enclosure unit (1)	What type of shelter does the household live in?	<ol> <li>Household without shelter - living outside/living in open</li> <li>Household living in makeshift shelter made from blanket or local materials</li> <li>Emergency shelter provided by organisation</li> <li>Communal shelter or Transit shade constructed by an organisation</li> <li>Transitional shelter solution using organisation materials</li> <li>Transitional shelter solution - mud brick</li> <li>Partitioned pre-existing Structure (mosque, school, other public building)</li> <li>Communal pre-existing Structure (mosque, school, other public building)</li> <li>Masonry house</li> <li>Mud brick / traditional House</li> <li>Don't know</li> <li>No response</li> </ol>	HH
	A.8.2	HH Interview	% of households with access to a safe and healthy housing enclosure unit (2)	Does the shelter have any of the following enclosure issues?	<ol> <li>Lack of insulation from cold</li> <li>Leaks during light rain</li> <li>Leaks during heavy rain</li> <li>Limited ventilation (not sufficient to provide appropriate air circulation)</li> <li>Presence of dirt or debris (removable)</li> <li>Presence of dirt or debris (non-removable)</li> <li>None of the above (cannot select with any other option)[No other options can be selected]</li> <li>Other (specify)</li> <li>Don't want to say/no answer [No other options can be selected]</li> <li>Don't know [No other options can be selected]</li> </ol>	ΗΗ
	A.8.3	HH Interview	% of household living in overcrowded shelters	Apart from the members of your household, how many people share this shelter?	Integer	HH
	A.8.4	HH Interview	% of households whose shelter solutions meet agreed technical and performance standards	[If 7, 8, 9, 10, 11 or 12 selected for shelter type] Is this shelter damaged?	<ol> <li>No to little damage only, does not present any structural risk</li> <li>Yes. Partially damaged - presenting some structural risk but still liveable</li> <li>Yes. Completely damaged</li> <li>NR</li> <li>DK</li> </ol>	ΗΗ

	A.8.5	HH Interview	% of households whose shelter solutions meet agreed technical and performance standards	[If 4 or 5 selected for shelter type] Are any elements of this structure damaged?	<ol> <li>No to little damage</li> <li>Tarpaulin damaged</li> <li>Timber damaged</li> <li>Completely damaged (i.e. both Tarpaulin and Timber damaged)</li> <li>NR</li> <li>DK</li> </ol>	ΗΗ
	A.8.6	HH Interview	% of households whose shelter solutions meet agreed technical and performance standards	[If 6 selected for shelter type] Are any elements of this structure damaged?	<ol> <li>No to little damage (cannot select with any other option)[No other options can be selected]</li> <li>Tarpaulin damaged</li> <li>Timber frame damaged</li> <li>Roof Damaged</li> <li>Completely damaged (cannot select with any other option)[No other options can be selected]</li> <li>NR</li> <li>DK</li> </ol>	ΗH
	A.8.7	HH Interview	% of households with housing, land and property issues	Do you currently have any of the following problems related to housing, land and property?	<ol> <li>Disputed ownership</li> <li>Property unlawfully occupied by others (secondary occupation)</li> <li>Disputes about rent (including payment) between landlord and tenant</li> <li>Rules and processes on housing and land not clear</li> <li>Inheritance issues</li> <li>Lack or loss of housing land tenancy or ownership documents</li> <li>Looting of private property</li> <li>Threat or fear of eviction/harassment by landlord or others</li> <li>None of the above (cannot select with any other option)[No other options can be selected]</li> <li>Other (specify)</li> <li>Don't want to say/no answer [No other options can be selected]</li> <li>Don't know [No other options can be selected]</li> </ol>	H
	A.8.8	HH Interview	% of households in need of non-food items	What are the essential non-food items that your household needs that it currently does not have?	<ol> <li>None, have all NFI items (cannot select with any other option) [No other options can be selected]</li> <li>Blankets</li> <li>Sleeping mat or mattress</li> <li>Mosquito net</li> <li>Jerry can</li> <li>Laundry detergent / bars</li> <li>Soap</li> <li>Menstrual hygiene products</li> <li>Solar lamp</li> <li>Kettle</li> <li>10L basin</li> <li>Kitchen items - pots</li> <li>Kitchen items - cutlery</li> <li>Aquatabs</li> <li>Education materials for children (notebook, textbook, school bag)</li> <li>Other (specify)</li> <li>NR [No other options can be selected]</li> <li>DK [No other options can be selected]</li> </ol>	ΗΗ
What are the protection- specific humanitarian needs of the crisis- affected	A.9.1	HH Interview	% of households that have experienced movement restrictions in the past 30 days	In the past 30 days, have any members of your household faced restrictions when moving outside your neighbourhood/block/camp to go to another nearby city/village/camp?	Yes, due to COVID-19 restrictions     Yes, due to other reasons     No     Have not tried to move around     Don't want to say/no answer     99. Don't know	ΗΗ
population, including issues related to safety and security, legal documentatio n and	A.9.2	HH Interview	% of households that have experienced movement restrictions in the past 30 days (3)	In the past 30 days, have any members of your household faced movement restrictions when travelling or wanting to travel to/from another state/LGA?	<ol> <li>Yes, due to COVID-19 restrictions</li> <li>Yes, due to other reasons</li> <li>No</li> <li>Have not tried to move around</li> <li>Don't want to say/no answer</li> <li>Don't know</li> </ol>	ΗΗ

movement restrictions for households?	A.9.3	HH Interview	% of households that have suffered incidents affecting household members in the past 3 months	Have any household members been affected by a safety or security incident in the past 3 months?	1. Yes 2. No 98. Don't want to say/no answer 99. Don't know	HH
	A.9.4	HH Interview	% of households that have suffered incidents affecting household members in the past 3 months	[If yes] What was the nature of this safety or security incident?	<ol> <li>Killings</li> <li>Physical violence</li> <li>Adult household members missing, detained, abducted or forcefully recruited</li> <li>Child household members missing, detained, abducted or forcefully recruited</li> <li>Armed attacks/Attack by armed group</li> <li>Fire outbreak</li> <li>Destruction of properties / Looting</li> <li>Incident related to explosive hazards affecting adult household member</li> <li>Incident related to explosive hazards affecting child household member</li> <li>Other (specify)</li> <li>Don't want to say/no answer [No other options can be selected]</li> <li>Don't know [No other options can be selected]</li> </ol>	ΗΗ
	A.9.5	HH Interview	% of households reporting at least one member disappeared or abducted in the last 3 months	[If 3 selected for type of security incident] Do you know what happened to them?	<ol> <li>Adult man taken by AOGs</li> <li>Adult woman taken by AOGs</li> <li>Adult man detained by police/military</li> <li>Adult woman detained by police/military</li> <li>Adult lost during household displacement / movement</li> <li>Adult man missing</li> <li>Adult woman missing</li> <li>No response [No other options can be selected]</li> <li>Don't know [No other options can be selected]</li> </ol>	HH
	A.9.6	HH Interview	% of households reporting at least one member disappeared or abducted in the last 3 months	[If 4 selected for type of security incident] Do you know what happened to them?	<ol> <li>Boy child taken by AOGs</li> <li>Girl child taken by AOGs</li> <li>Boy child taken by CJTF</li> <li>Girl child taken by CJTF</li> <li>Child trafficked out of state/country</li> <li>Child lost during household displacement / movement</li> <li>Child missing (not during household displacement)</li> <li>Child detained by police / military</li> <li>Orbit detained by police / military</li> <li>Don't want to say/no answer [No other options can be selected]</li> <li>Don't know [No other options can be selected]</li> </ol>	HH

A.9.7	HH Interview	% of households that have suffered incidents involving explosive ordnance in the past 3 months	[If 8 selected for type of security incident] Do you know what happened to them?	<ol> <li>Male household member travelling on frequently travelled roads (i.e. to work or market)</li> <li>Female household member travelling on frequently travelled roads (i.e. to work or market)</li> <li>Male household member when working on the farms</li> <li>Female household member when working on the farms</li> <li>Male household member when collecting firewood</li> <li>Female household member when collecting firewood</li> <li>Female household member when collecting water</li> <li>Female household member when collecting water</li> <li>Female household member when collecting scrap metal</li> <li>Female household member when collecting scrap metal</li> <li>Female household member when collecting scrap metal</li> <li>Female household member when scllecting scrap metal</li> <li>Female household member when shepherding</li> <li>Female household member when shepherding</li> <li>Female household member when shepherding</li> <li>Female household member whilst trying to remove an explosive device</li> <li>Male household member whilst relocating between LGAs or camps</li> <li>Female household member whilst relocating between LGAs or camps</li> <li>No response [No other options can be selected]</li> <li>Don't know [No other options can be</li> </ol>	HH
A.9.8	HH Interview	% of households that have suffered incidents involving explosive ordnance in the 3 months	[If 7, 8, 9, 10, 13, 14, 15 or 16 selected for type of explosive ordinance incident] What type of land was this on?	<ol> <li>Agricultural land</li> <li>Grazing land</li> <li>Forested areas</li> <li>In and around trenches</li> <li>Residential areas / homes</li> <li>Abandoned villages / homes</li> <li>Roadside</li> <li>In and around military areas (checkpoints, roadblocks)</li> <li>Other</li> <li>No response</li> <li>Don't know</li> </ol>	HH

A.9.9	HH Interview	% of households that have suffered incidents involving explosive ordnance in the 3 months	[If 9 for type of security incident] Do you know what happened to them?	<ol> <li>Boy child household member travelling on frequently travelled roads (i.e. to school)</li> <li>Girl child household member travelling on frequently travelled roads (i.e. to school)</li> <li>Boy child household member at school or near school</li> <li>Girl child household member at school or near school</li> <li>Boy child household member when working on the farms</li> <li>Girl child household member when working on the farms</li> <li>Girl child household member when working on the farms</li> <li>Girl child household member when collecting firewood</li> <li>Girl child household member when collecting water</li> <li>Girl child household member when collecting water</li> <li>Girl child household member when collecting water</li> <li>Girl child household member when collecting scrap metal</li> <li>Boy child household member when collecting scrap metal</li> <li>Boy child household member when shepherding</li> <li>Boy child household member whilst trying to remove an explosive device</li> <li>Girl child household member whilst trying to remove an explosive device.</li> <li>Boy child household member whilst relocating between LGAs or camps</li> <li>No response [No other options can be selected]</li> <li>Don't know [No other options can be selected]</li> </ol>	H
A.9.1 0	HH Interview	% of households that have suffered incidents involving explosive ordnance in the 3 months	[If 9, 10, 11, 12, 15, 16, 17 or 18 selected for type of explosive ordinance incident] What type of land was this on?	<ol> <li>Agricultural land</li> <li>Grazing land</li> <li>Forested areas</li> <li>In and around trenches</li> <li>Residential areas / homes</li> <li>Abandoned villages / homes</li> <li>Roadside</li> <li>In and around military areas (checkpoints, roadblocks)</li> <li>Other</li> <li>No response</li> <li>Don't know</li> </ol>	H
A.9.1 1	HH Interview	% of households reporting one or several children that report experiencing signs of psychological distress following a safety or security incident	[If ANY security incident selected] Following this event, did any of your children's behaviour change? How did it change?	<ol> <li>No (cannot select with any other option) [No other options can be selected]</li> <li>Unwillingness to go to school</li> <li>Difficulties with sleep</li> <li>Isolation</li> <li>Less willingness to help parents/caregivers</li> <li>Sadness</li> <li>Violence against other children</li> <li>Violent behaviour in general in the community</li> <li>Eagerness to join armed/vigilante group</li> <li>Substance abuse</li> <li>Committing crime</li> <li>Engaging in high risk sexual behaviour</li> <li>Other (specify)</li> <li>No response [No other options can be selected]</li> <li>Don't know [No other options can be selected]</li> </ol>	ΗΗ

	A.9.1 2	HH Interview	% of households that have suffered incidents affecting household members in the past 30 days (2)	[If ANY security incident selected] Was this/Were any of these incidents in the past 30 days?	1. Yes 2. No 98. Don't want to say/no answer 99. Don't know	HH
	A.9.1 3	HH Interview	% of men, women, boys and girls without a valid Passport and/or valid national ID, at the time of data collection (1)	Do all adult household members currently have an International Passport and/or a valid National Identification document at this time?	<ol> <li>Yes, in our possession</li> <li>Yes, we all have IDs but they are not in our possession</li> <li>No, some household members are missing IDs</li> <li>Don't want to say/no answer</li> <li>Don't Know</li> </ol>	НН
	A.9.1 4	HH Interview	% of men, women, boys and girls without a valid Passport and/or valid national ID, at the time of data collection (1)	[If 2 or 3 selected] In the past 3 months, did this/these adult(s) experience any problems accessing basic services due to not having an ID?	1. Yes 2. No 98. Don't want to say/no response 99. Don't know	HH
	A.9.1 5	HH Interview	% of men, women, boys and girls without a valid Passport and/or valid national ID, at the time of data collection (1)	Do all child household members currently have an International Passport and/or valid national Identification document at this time?	<ol> <li>Yes, in our possession</li> <li>Yes, we all have IDs but they are not in our possession</li> <li>No, some household members are missing IDs</li> <li>Don't want to say/no response</li> <li>Don't know</li> </ol>	HH
	A.9.1 6	HH Interview	% of men, women, boys and girls without a valid Passport and/or valid national ID, at the time of data collection (1)	[If 2 or 3 selected] In the past 3 months, did this/these child(ren) experience any problems accessing basic services due to not having an ID?	1. Yes 2. No 98. Don't want to say/no answer 99. Don't know	HH
	A.9.1 7	HH Interview	% of HHs with children under 18 currently not residing in the HH	Do you have any other child, son or daughter under 18 years not living in the HH? If yes, how many?	1. Yes 2. No 98. Don't want to say 99. Don't know	HH
	A.9.1 8	HH Interview	% of HHs with children under 18 currently not residing in the HH	[If yes] How many other children under 18 years are not living in the HH?	Integer	HH
	A.9.1 9	HH Interview	% of HHs with children under 18 currently not residing in the HH, by protection incident	We would like to understand why those children are not living under your roof. I will read you a list of possibilities, let me know how many children currently under 18 years old fall in each category:	<ol> <li>Married and left the house</li> <li>Left the house to seek employment</li> <li>Left the house to study</li> <li>Left the house to engage with the army or armed groups</li> <li>Kidnapped/abducted</li> <li>Missing (left and no news)</li> <li>Arbitrarily detained</li> <li>Don't want to say</li> <li>Don't know</li> </ol>	НН
What are the humanitarian response specific needs of the crisis-	A.10. 1	HH Interview	% of households who received aid in the past 3 months	Has your household received aid from an organisation in the past 3 months?	1. Yes 2. No 98. NR 99. DK	HH
affected population, including issues related to	A.10. 2	HH Interview	[Of those who received aid] % of households who were satisfied with the aid they received	[If yes] Was your household satisfied with the aid you received?	1. Yes 2. No 98. NR 99. DK	HH

access to assistance?	A.10. 3	HH Interview	[Of those who received aid and were dissatisfied with aid received] Most commonly reported reasons for dissatisfaction with the aid received	[If 2 selected] Why were you not satisfied with the aid received?	<ol> <li>Quality was not good enough</li> <li>Quantity was not good enough</li> <li>Did not receive the aid on time / Delays in delivery of aid</li> <li>Type of assistance provided was not suitable for my need</li> <li>Other (specify)</li> <li>NR [No other options can be selected]</li> <li>DK [No other options can be selected]</li> </ol>	ΗΗ
What are the COVID-19 specific needs of the crisis- affected population, including strategies adopted by households in response to the current COVID-19 health emergency	A.11. 1	HH Interview	% of households who know about COIVD-19 virus	Have you heard about COVID-19, also known as coronavirus?	Yes No	Ŧ
	A.11. 1	HH Interview	% of households adapting behaviours to try to prevent COVID-19 spreading	Since you heard about COVID-19, have you and your household members taken any action to prevent yourselves from getting COVID-19?	<ol> <li>No, no action taken (cannot select with any other option) [No other options can be selected]</li> <li>Not leaving the house at all</li> <li>Reducing movement outside the house</li> <li>Stopping handshakes or physical contact</li> <li>Keeping distance from people</li> <li>Avoiding public places and gatherings</li> <li>Avoiding public transport</li> <li>Wearing a face mask</li> <li>Wearing gloves</li> <li>Washing hands more regularly</li> <li>Keeping surfaces clean</li> <li>Praying to god</li> <li>Staying away from animals</li> <li>Other (specify)</li> <li>NR [No other options can be selected]</li> <li>DK [No other options can be selected]</li> </ol>	ΗΗ
	A.11. 2	HH Interview	% of households by type of action taken if suffering from COVID-19	What would you do if you think you or someone in your household has COVID-19?	<ol> <li>Nothing, continue daily life as normal [No other options can be selected]</li> <li>Stay home and do nothing / take no medicine</li> <li>Stay home and self-medicate</li> <li>Call hospital or emergency services</li> <li>Call dedicated COVID-19 number</li> <li>Speak to a religious leader</li> <li>Speak to a settlement leader</li> <li>Go to a pharmacy</li> <li>Go to a doctor's office or health centre</li> <li>Go to a traditional / local healer</li> <li>Other (specify)</li> <li>NR [No other options can be selected]</li> <li>DK [No other options can be selected]</li> </ol>	ΗΗ

Most commonly reported priority nee among assessed household	d	What do you think is the main priority need for your household overall?	<ol> <li>No needs</li> <li>Food</li> <li>Water</li> <li>Latrines</li> <li>Health</li> <li>Reunification with family</li> <li>Psychosocial support services</li> <li>Legal documentation</li> <li>Education / schools</li> <li>Livelihoods / income / cash</li> <li>Shelter support</li> <li>Non-food items (blankets, jerry cans, buckets, etc.)</li> <li>Other</li> <li>NR</li> <li>DK</li> </ol>	ΗΗ
Final comments		Do you have any final comment on your situation and needs?	Text (not required)	HH

# 7. Monitoring & Evaluation Plan

IMPACT Objective	External M&E Indicator	Internal M&E Indicator	Focal point	Tool	Will indicator be tracked?
		# of downloads of x product from Resource Center	Country request to HQ		x Yes
	Number of humanitarian	# of downloads of x product from Relief Web request to	x Yes		
Humanitarian stakeholders are	organisations accessing IMPACT services/products	# of downloads of x product from Country level platforms	Country team		□ Yes
accessing IMPACT products	Number of individuals accessing IMPACT services/products	# of page clicks on x product from REACH global newsletter	Country request to HQ	User_log	□ Yes
		# of page clicks on x product from country newsletter, sendingBlue, bit.ly	Country team		x Yes
		# of visits to x webmap/x dashboard	Country request to HQ	-	□ Yes
IMPACT activities contribute to better		# references in HPC documents (HNO, SRP, Flash appeals, Cluster/sector strategies)			HNO/HRP for Nigeria 2021 Donor strategy documents
program implementation and coordination of the humanitarian response	Number of humanitarian organisations utilizing IMPACT services/products	# references in single agency documents	Country team	Reference_I og	Sector strategy documents Any other relevant strategic document Any other relevant assessment or research study
Humanitarian stakeholders are	Humanitarian actors use IMPACT evidence/products as a	Perceived relevance of IMPACT country-programs Perceived usefulness and influence of IMPACT outputs	Country	Usage_Feed back <i>and</i>	
using IMPACT products	basis for decision making, aid planning and delivery	Recommendations to strengthen IMPACT programs Perceived capacity of IMPACT staff Perceived quality of outputs/programs	team	Usage_Surv ey template	Usage survey to be conducted at the end of the year for the whole

	Number of humanitarian documents (HNO, HRP, cluster/agency strategic plans, etc.) directly informed by IMPACT products	Recommendations to strengthen IMPACT programs			Nigeria mission; no need for a usage survey specific to the MSNA 2020
Humanitarian stakeholders are	Number and/or percentage of humanitarian organizations directly	# of organisations providing resources (i.e.staff, vehicles, meeting space, budget, etc.) for activity implementation			x Yes
engaged in IMPACT programs throughout the	contributing to IMPACT programs (providing	# of organisations/clusters inputting in research design and joint analysis	Country team	Engagement _log	x Yes
research cycle	resources, participating to presentations, etc.)	# of organisations/clusters attending briefings on findings;			x Yes

## ANNEX 1: DATA CLEANING SOPS

# During data collection: overview of responsibilities for data cleaning and follow up

Staff	Responsibility
Enumerators	Conduct snowball sampling, data collection and follow up, on a daily basis.
	Flag any issues to their team lead
Team leads	Supervise their team
	Ensure daily briefs and debriefs on a daily basis, following Field Assistant instructions
	Ensure all forms are submitted at the end of the day
	Support the Field Assistant with follow up
	Flag any issues to the Field Assistant
Field Assistants (FA)	Supervise team leads
	Ensure the smooth running of data collection and follow up activities with team leads
	Flag any issues to their Field Supervisors and Assessment Officer
Field Supervisors	Supervise Field Assistants
(Field Manager, Field Officers)	Track data collection progress, daily
	Compile results of snowball sampling and distribute contact numbers among teams
	Coordinate with the Database Officer regarding data collection progress
	Liaise with the Assessment Officer to coordinate follow up
Assessment Officers	Clean data on a daily basis
(AO)	Provide follow up questions to Field Assistants and coordinate with them to ensure
	follow up questions are addressed in a timely manner
	Support with the tracking of data collection progress
Database Officer	Download raw data at the end of each day
(DO)	

Overview of responsibilities (data cleaning and follow up)

	Anonymise the data and conduct preliminary R cleaning
	Share anonymised data and preliminary follow up with the relevant Assessment Officers
	Tracking of data collection progress overall, in coordination with Field supervisors and AOs
GIS Officer (GISO)	Conduct spatial verification checks daily (while data collection in the field is on-going)

Overview of weekly data cleaning activities

		Mon	Tue	Wed	Thu	Fri			
Enumerators	Morning Afternoon	<ul> <li>Snowball sampling, data collection and follow up</li> <li>Submit all forms to the server before the end of the day</li> </ul>							
Team Leads	Morning	- Brief team flagged by	n on daily targets y Field Assistant	and discuss iss	sues with data qu	uality as			
(TL)	Afternoon	<ul> <li>Supervise</li> <li>Ensure fo</li> <li>Ensure all</li> </ul>	Supervise teams Ensure follow up is done and communicated back to the field assistant Ensure all forms are submitted to the server by the end of the day						
Field	Morning	- Provide team leads with daily sample / targets							
Assistants	Afternoon	- Communi	Communicate any follow up needed to team leads, based on Assessment Officer feedback Compile follow up responses and send back to relevant Assessment Officer						
(FA)		- Compile for Officer							
Field	Morning	- Provide ta	arget samples to	teams, with the	support of field a	assistants			
Manager and Field	Afternoon	- Download	Download data and check data collection progress, after all forms have been uploaded						
Officers		<ul> <li>Inform Da</li> <li>Prepare ta</li> </ul>	Inform Database Officer and Assessment Officers of progress Prepare target samples for the following day						
Assessment Officers	Morning	- Full data o	cleaning of previ leaning log with	ous day's data new follow up gu	estions to Field	supervisors			
	Afternoon	- Work on f	ollow up respons	ses received from	m previous days				
Database	Morning	- Full data o	cleaning of previ	ous day's data					

Officer		Provide cleaning log with new follow up questions to Field supervisors				
(DO)	Afternoon	<ul> <li>Download raw data and anonymise</li> <li>Preliminary data cleaning (R script), share with AOs as relevant</li> <li>Keep track of overall data collection progress, in coordination with Field managers and AOs</li> </ul>				
GIS Officer (GISO)	Morning	NA				
()	Afternoon	- Spatial verification of the day's submissions (location, gps coordinates), when needed, as part of preliminary data cleaning				

# During data collection: process outline

Three forms will be on Kobo, one per state. Similarly, three enumerator teams, three FAs, three FSs and three AOs will be designated, one per state.

At the end of every day of data collection, enumerators will upload survey data to the KoBo server (by 5:30 p.m.). Team Leads, with the oversight of Field Assistants, will ensure that this is done on a daily basis.

**After data is uploaded**, the Field supervisors (FS) and Database Officer will download the raw data from the KoBo server. Only the Database Officer is authorised to download raw data from all three Kobo forms, while each FS will be able to access their respective form for snowball sampling and data collection tracking purposes.

#### Preliminary checks of the day's data:

The Database Officer will

- a. Download the 3 raw datasets
- b. Filter the data to include only submissions from that day;
- c. Filter out "snowball sampling only" submissions;
  - "sampling/snowball"
  - "sampling/snowball\_2"
  - "sampling/snowball\_nr"
  - "sampling/snowball\_number"
  - "sampling/phone\_numbers"
  - "sampling/loopcount\_2"
  - "sampling/lga\_snowball"
  - "sampling/enumerator"
  - "sampling/type\_interview"
  - "sampling/group\_target"
  - "sampling/hoh\_gender\_target"
  - "sampling/new\_number"
  - "sampling/snowball\_call\_back"
  - "sampling/snowball\_time"
  - "sampling/phone\_numbers"
  - "sampling/enum\_comment"
  - "sampling/note end sampling"
- d. Check for and remove duplicates in the day's raw data to ensure that all observations are unique;
- e. *Raw Dataset for Spatial Verification*: the GISO will conduct spatial verification checks, where relevant, flagging entries deemed suspicious. Data columns to be included to the *Spatial Verification checks* files are the following:
  - `\_uuid
  - `pre\_survey\_inperson/state`
  - `pre\_survey\_inperson/lga`
  - `pre\_survey\_inperson/ward`
  - `pre\_survey\_inperson/village\_site`
  - `pre\_survey\_inperson/village\_site\_other`
  - `pre\_survey\_inperson/gps`
  - `pre\_survey\_inperson/\_gps\_latitude`

- `pre\_survey\_inperson/\_gps\_longitude`
- `pre\_survey\_inperson/\_gps\_altitude`
- `pre\_survey\_inperson/\_gps\_precision`

f. Anonymised Raw Dataset: The DO will generate a copy of the raw data with identifying information completely removed, and save it to the relevant folder on the Dropbox. Data columns to be removed are the following:

- target\_phone\_number
- hh\_phone\_number
- village\_site
- village\_site\_other
- gps
- name\_member
- age\_group\_3\_17\_name
- age\_group\_0\_59\_months\_name
- age\_group\_12\_23\_months\_name
- age\_group\_6\_59\_months\_name
- plw\_caretaker\_name
- school\_aged\_children\_names
- nutrition\_group\_names
- nutrition\_group\_12\_23m\_names
- nutrition\_group\_6\_59m\_names
- plw\_caretaker\_group\_names
- g. R script: the DO will run the R script on the Anonymised raw data. (R-script Checks explained in detail in ANNEX I). Outputs of the R script are as follows:
  - i. **Issues Log:** A log in xlsx format listing each specific case where checks have found a discrepancy in the data. Each specific case will include a comment to the Field Officer on how to address the issue. Please find the script at the end of this document.
- h. The DO will merge the *Spatial verification log* and *Issues Log* into a standardised, xlsx format, *Cleaning Log*;
- i. The DO will send the *anonymised raw dataset* and the *Cleaning Log*, <u>disaggregated by state</u>, to the relevant AO.

#### Completing data checks of previous day's data

The following day, Assessment Officers will conduct the following processes on their assigned dataset:

- a. Data Checks (Macro Tool): Using the macro-enabled excel cleaning tool, check for and investigate any discrepancies in the dataset. Main Checks to be conducted are as follows:
  - Enumerator checks: Instances where similar response paths for one enumerator are identified should be flagged to the relevant Flat Assistant for immediate follow-up (Does a specific enumerator(s) always tend to select the fastest possible response combinations?);
  - Enumerator checks: Multiple records filled by the same enumerator showing the same responses to a
    particular question/ series of questions; should be flagged to the relevant Flat Assistant for immediate followup;
  - Contradiction Checks: Instances in the dataset where reported answers appear to contradict other reported answers. To address these issues, discuss with enumerators to make sure it is understood how to ask these questions and how to collect responses.
  - More than Three Responses Checks: Instances in the dataset where enumerators are found to be reporting
    more than three responses for select multiple choices significantly more than other enumerators (a pattern is
    noticed). To address these issues, discuss with enumerators to make sure it is understood how to ask these
    questions and how to collect responses.

- Do not know / No answer Checks: Instances in the dataset where enumerators are found to be commonly reporting *Do not know / No answer* for select multiple choices significantly more than other enumerators (a pattern is noticed). To address these issues, discuss with enumerators to make sure it is understood how to ask these questions and how to collect responses.
- **Other Response Checks:** Instances in the data where enumerators selected *Other option* and entered a text response to explain an option. To address these issues, review and translate the text response and add to the data frame. If the text translation is unclear the enumerator will be contacted for clarification.
- **Unusually fast interviewing times:** Instances where an enumerator is significantly faster than other enumerators (less than 10 minutes per interview).
- **Any other Checks**: All other potential issues in the dataset based on the situational knowledge and experience of the Field Officer to be addressed as necessary and recorded in the cleaning log.
- b. Compile all needed checks into the Cleaning Log (xlsx document);
- c. Submit the Cleaning Log to the relevant Field Officer for follow up;
- d. Ensure, in coordination with the Field Officer, that follow up for data collected on day *n* is completed by day *n*+4 at the very latest; data collection should be organised in a way that data collection and follow up are both occurring on a daily basis and carried out equally rigorously.

#### Final cleaning (day *n* dataset)

- AOs will, based on follow up received from the field, update the **Cleaning Log** and clean the dataset as necessary using the macro-enabled excel cleaning tool.
- They will send **Anonymised Datasets** (3), **Cleaned Datasets** (3) and **Cleaning Logs** (3) to the Senior Assessment Officer for review.
- They will update a live **Data Deletion Log**, where they will log all records that are deleted and the reason for deletion.

#### Tracking of data collection progress:

FSs will conduct a daily tracking of snowball sampling and of data collection by checking the day's raw data.

Based on the raw data they will identify:

a. number of surveys conducted that day per LGA and per population group;

b. phone numbers that were contacted successfully and phone numbers that need to be contacted at a later time (due to no responses or to the respondent asking to reschedule the call).

c. among phone numbers that were contacted successfully, select 10% randomly for FAs to conduct back checks. d. for 'Assisted remote' surveys, they'll provide the **number of surveys conducted in each location** to Field Assistants, who will **cross-check this information with copies of the preliminary consent forms** filled in by partners on the ground.

They will use this information to plan data collection and back checks for the following day, and flag any issues to the DO.

The DO will be monitoring data collection progress across the three states on a **data collection tracking document**, accessible by FSs and AOs. This will include number of successful interviews per LGA, per population group, as well as rate of non-response at the LGA level. This will be done to cross-check the sample collected against the originally intended sample, and plan in time for re-sampling if needed.

The **data collection tracking** document will be accompanied by a live **Data Deletion Log**, where AOs and DO will log all records that are deleted during the data cleaning process and the reasons for deletion. Field Supervisors will be immediately notified of any record deletions and its impact on data collection progress.

# After data collection

The above processes will be repeated for every day of data collection. Once data collection is completed, the relevant final checks will be conducted in accordance with **IMPACT Data Cleaning Guidelines**. Risk indicators will be removed, as per the data management plan. Data and the accompanying documentation will be submitted to the Country Coordinator for review, prior to sending to HQ for final check and validation.

# **R-script Checks**

# Title : Master Code
# Developer : Cal Pearson
# Email : callum.pearson@reach-initiative.org // callum-pearson@outlook.com
# Version : v1
# Date Mofified : 24/06/2020
# Modified by who : Cal Pearson
# PACKAGES
library(tidyverse)
library(lubridate)
library(readxl)
library(httr)
library(devtools)
library(ggplot2)
library(summarytools)
library(datasets)
library(Hmisc)
library(scales)
library(roxygen2)
library(naniar)
library(hablar)
library(writexI)
library(xlsx)
library(XLConnect)
library(openxlsx)

# Part : 1

# Title : Download data from the KOBO humanitarian response.

#### # 1.1) Locate raw file in the github.

source\_url("https://raw.githubusercontent.com/ppsapkota/kobohr\_apitoolbox/master/R/r\_func\_ps\_kobo\_utils.R")

#### # 1.2) Download form/project list

kobo\_user <- "calpearson" # kobo user account name as string (example "nnkbuser") kobo\_pw <- "# # password for kobo user account as string (example "nnkbpassword")</pre>

# 1.3) This is getting your 6 figure codes for each of your forms. url <- "https://kc.humanitarianresponse.info/api/v1/data.csv" d\_formlist\_csv <- kobohr\_getforms\_csv(url,kobo\_user, kobo\_pw) d\_formlist\_csv <- as.data.frame(d\_formlist\_csv)</pre>

# Plug in your 6 figure code of the form you want to take data from.

# 1.4) Download data in CSV format

form\_id <- 562058 # id of the deployed project (for example 112233)

# remember you need to put the number at the end .csv of the code below.

# In this case its 553863.csv

url<-"https://kc.humanitarianresponse.info/api/v1/data/562058.csv" d\_raw <- kobohr\_getdata\_csv(url,kobo\_user,kobo\_pw) d\_raw <- as.data.frame(d\_raw)

# 1.5) Remove all the functions and values as we don't need them anymore.

rm(form\_id, kobo\_pw,kobo\_user,url, kobohr\_count\_submission, kobohr\_create\_export, kobohr\_download\_csv, kobohr\_getdata, kobohr\_getdata\_csv, kobohr\_kpi\_deploy\_asset, kobohr\_kpi\_get\_asset\_list, kobohr\_kpi\_get\_asset\_uid, kobohr\_kpi\_share\_asset, kobohr\_kpi\_upload\_xlsform, kobohr\_latest\_exports, kobohr\_list\_exports, kobohr\_getforms, kobohr\_getforms\_csv, d\_formlist\_csv)

# Part : 2

# Title : Raw Data Group

# 2.1) Tidy up the data

d\_raw <- as.tibble(d\_raw)

# 2.2) Save the raw data to file.

write.csv(d\_raw,"C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/1) daily raw data/d\_raw.csv", row.names = FALSE)

# 2.3) De-duplicate convert all n/a to a true NA.

d\_raw\_dedupe <- d\_raw %>%

distinct() %>%

replace\_with\_na\_all(condition = ~.x %in% common\_na\_strings)

# 2.4) Save the de-duplicated raw data

write.csv(d\_raw\_dedupe,"C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/2) daily raw dedupe/d\_raw\_dedupe.csv", row.names = FALSE)

# 2.5) How many duplicates are there total? duplicates\_total <- nrow(d\_raw)-nrow(d\_raw\_dedupe) duplicates\_total

# 2.6) Filter data from today

dedupe\_today <- d\_raw\_dedupe %>%

filter(today == Sys.Date())

# ###### This bit is just to make sure that i can keep testing with old data please remove after testing is complete.

dedupe\_today <- d\_raw\_dedupe

# 2.7) Find the number of duplicates in todays haull

d\_raw\_today <- d\_raw %>%

filter(today == Sys.Date())

duplicates\_today <- nrow(dedupe\_today)-nrow(d\_raw\_today)

duplicates\_today

# 2.8) Save the de-duplicated data for today.

write.csv(dedupe\_today,"C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/3) daily raw dedupe today/dedupe\_today.csv", row.names = FALSE)

# 2.9) Make a master dataframe backup

write.csv(d\_raw\_dedupe,"C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/4) master dataframe/master\_dataframe.csv", row.names = FALSE)

rm(d\_raw, d\_raw\_dedupe, d\_raw\_today)

# Part : 3

# Title : Daily Spatial Verification (DSV\_1)

# 3.1) Seperate information into | UUID, Localisation and GPS\_coordinates |

DSV\_1 <- dedupe\_today %>%

select("\_uuid",

today,

"pre\_survey\_inperson/state",

'pre\_survey\_inperson/lga',

"pre\_survey\_inperson/ward",

"pre\_survey\_inperson/village\_site",

"pre\_survey\_inperson/village\_site\_other",

"pre\_survey\_inperson/gps",

"pre\_survey\_inperson/\_gps\_latitude",

"pre\_survey\_inperson/\_gps\_longitude",

"pre\_survey\_inperson/\_gps\_altitude",

"pre\_survey\_inperson/\_gps\_precision")

# 3.2) Download old set of data.

DSV\_1\_old <- read\_excel("C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/5) daily dedupe raw spatial verification.xlsx", sheet = "Sheet1")

# 3.3) Attach new data to old data.

DSV\_AllToDate <- rbind(DSV\_1,DSV\_1\_old)

write\_xlsx(DSV\_AllToDate, "C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/5) daily dedupe raw spatial verification.xlsx")

# 3.4) Remove old data

rm(DSV\_1,DSV\_1\_old)

```
# Part : 4
```

# Title : Splitting the data into two dataframes one for hh\_level and another for ind\_level

# 4.1) This will split all the iterations for the loop with the hh\_members

```
compile_indloop<-function(n,dedupe_today){
index<-paste0("[",n,"]")
hh <- dedupe_today %>% select("_uuid",enumerator_code,contains(index)))
if(ncol(hh)!=2){
    names(hh) <- gsub(index, "", names(hh), fixed = TRUE)
    hh <- hh %>% filter(`A0_demographics/A_0_8_hh_composition/name_member` != "n/a")
    hh$hh_number <- n
    hh <- retype(hh)
    return(hh)
}
</pre>
```

```
# 4.2) Bind the dataframes together
```

ind\_level<-lapply(1:40,compile\_indloop,dedupe\_today=dedupe\_today) %>% bind\_rows

# 4.3) Remove any true NA's from the name\_member so your getting all the information !

```
ind_level <- ind_level %>%
```

filter(!is.na(`A0\_demographics/A\_0\_8\_hh\_composition/name\_member`)) %>%

arrange(`\_uuid`)

# 4.4) Remove all the [] from dedupe\_today making dedupe\_today\_hh - This gives only data that is on household level hh\_level <- dedupe\_today %>% select(!contains("["))

# Part : 5

#### # Title : Daily Anonnymised Raw Data for individual and household (DARD\_hh & DARD\_ind)

# 5.1) Filter out any identifiable information at hh\_level

DARD\_hh <- hh\_level %>%

select(!contains(c("target\_phone\_number","hh\_phone\_number","village\_site","village\_site\_other","gps","name\_member","age\_group\_3\_17\_name","a ge\_group\_0\_59\_months\_name","age\_group\_12\_23\_months\_name","age\_group\_6\_59\_months\_name","plw\_caretaker\_name","school\_aged\_childre n\_names","nutrition\_group\_names","nutrition\_group\_12\_23m\_names","nutrition\_group\_6\_59m\_names","plw\_caretaker\_group\_names","sampling/sn owball","sampling/snowball\_2","sampling/snowball\_nr","sampling/snowball\_number","sampling/phone\_numbers","sampling/loopcount\_2","sampling/loopcount\_2","sampling/group\_target","sampling/hoh\_gender\_target","sampling/now\_number","sampling/snowball\_call\_back","sampling/snowball\_time","sampling/phone\_numbers","sampling/enum\_comment","sampling/note\_end\_sampling")))

# 5.2) Save the daily D	ARD_hh
-------------------------	--------

write_xlsx(DARD_hh,"C:/Users/krull/Desktop/IMPACT	start/1)	Work/Automation	deposit/6)	daily	anon	raw	dedupe
today/daily_anon_raw_data_hh.xlsx")							

# 5.3) Remove the DARD\_hh once it is saved

rm(DARD\_hh)

# 5.4) Filter out any identifiable information at hh\_level

DARD\_ind <- ind\_level %>%

select(!contains(c("target\_phone\_number","hh\_phone\_number","village\_site","village\_site\_other","gps","name\_member","age\_group\_3\_17\_name","age\_group\_0\_59\_months\_name","age\_group\_0\_59\_months\_name","plw\_caretaker\_name","school\_aged\_childre n\_names","nutrition\_group\_names","nutrition\_group\_names","nutrition\_group\_12\_23m\_names","nutrition\_group\_6\_59m\_names","plw\_caretaker\_group\_names","sampling/sn owball","sampling/snowball\_2","sampling/snowball\_nr","sampling/snowball\_number","sampling/phone\_numbers","sampling/hone\_group\_target","sampling/hone\_group\_target","sampling/hone\_group\_target","sampling/hone\_number","sampling/nowball\_call\_back","sampling/snowball\_time","sampling/phone\_numbers","sampling/enumeration","sampling/note\_end\_sampling/phone\_numbers","sampling/enumert","sampling/note\_end\_sampling","sampling/note\_end\_sampling","sampling/note\_end\_sampling","sampling/note\_end\_sampling","sampling/note\_end\_sampling","sampling/note\_end\_sampling","sampling/note\_end\_sampling","sampling/note\_end\_sampling","sampling/note\_end\_sampling","sampling/note\_end\_sampling","sampling/note\_end\_sampling","sampling/note\_end\_sampling","sampling/note\_end\_sampling","sampling/note\_end\_sampling","sampling/note\_end\_sampling","sampling,","sampling/note\_end\_sampling","sampling/note\_end\_sampling","sampling,","sampling/note\_end\_sampling","sampling,","samp

write_xlsx(DARD_ind,"C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/6) daily anon raw today/daily_anon_raw_data_ind.xlsx")	dedupe

# 5.6) Remove the DARD\_hh once it is saved

rm(DARD\_ind)

# Part : 6

# Title : Issues Check

#### # 1) Make a list of unusually fast and slow interviews

# 1.0) Check that all the time stamps have a particular time zone and make a histogram of the days work hh\_level\$difference <- hh\_level\$end - hh\_level\$start hh\_level\$difference <- as.numeric(hh\_level\$difference) # 1.1) Make a histogram to see the time distribution. hist(hh\_level\$difference) # 1.2) Flag any results where the difference is greater than x and lower than y fast <- hh\_level %>% select(`\_uuid`, enumerator\_code, start, end, today, difference) %>% filter( difference < 20) slow <- hh\_level %>% select(`\_uuid`, enumerator\_code, start, end, today, difference) %>% filter(difference > 50)

#### # 2) Flag up any thing that displays other in both hh\_level and ind\_level

# 2.0) Flag up anything that shows "other" # 2.1) Make a quick function to remove a column that contains all NA rm\_col\_all\_na <- function(x) any(!is.na(x)) # 2.2) Pipe in and select for any questions that have other in the name of the question for hh\_level other\_hh <- hh\_level %>% select('\_uuid',enumerator\_code, contains("other")) %>% select\_if(rm\_col\_all\_na) other\_hh\$enumerator\_code <- as.character(other\_hh\$enumerator\_code)</pre>

other\_hh <- select\_if(other\_hh, is.character) other\_hh <- other\_hh %>% select\_if(~!any(. == c("yes","no"), na.rm = TRUE))

# 2.3) Pipe in and select for any questions that have other in the name of the question for ind\_level other\_ind <- ind\_level %>% select('\_uuid',enumerator\_code, contains("other")) %>% select\_if(rm\_col\_all\_na) other\_ind\$enumerator\_code <- as.character(other\_ind\$enumerator\_code)</pre>

other\_ind <- select\_if(other\_ind, is.character) other\_ind <- other\_ind %>% select\_if(~!any(. == c("yes","no"), na.rm = TRUE))

#### # 3) All the enumerators age and the hoh age do not match the hh\_ind ages.

# 3.0) Find all ages household level age\_check <- hh\_level %>% select(`\_uuid`, enumerator\_code,`pre\_survey\_inperson/age\_respondent`, `pre\_survey\_phone/pre\_survey\_respondent\_eligibility/respondent\_age`,`A0\_demographics/A0\_2\_age\_hoh`)

# 3.1) Bind ages into age\_respondent age\_check\$age\_respondent <- ifelse(!is.na(age\_check\$`pre\_survey\_inperson/age\_respondent`), age\_check\$`pre\_survey\_inperson/age\_respondent`, age\_check\$`pre\_survey\_phone/pre\_survey\_respondent\_eligibility/respondent\_age`)

# 3.2) Select only the useful information age\_check <- age\_check %>% select(`\_uuid`, enumerator\_code, age\_respondent, `A0\_demographics/A0\_2\_age\_hoh`)

# 3.3) Find all the ages at individual level age\_check\_ind <- dedupe\_today %>% select(`\_uuid`, contains("age\_member")) # 3.4) Bind into one table to look at the ages of household members and reported heads of household ages as well as respondent ages. age\_check\_tot <- left\_join(age\_check,age\_check\_ind, by = '\_uuid') age\_check\_tot <- retype(age\_check\_tot) # Using the age\_check\_tot select out the uuid of each which dont match.

# 3.5) Age check for the respondent # Find the maximum household size hh\_n <- max(hh\_level\$`A0\_demographics/A0\_7\_hh\_size`) hh\_n <- hh\_n + 4</pre>

# 3.5) Plug in to find the rows that show FALSE (i.e. where the hh\_members don't include the respondents age) show\_false <- as.data.frame(age\_check\_tot[,5:hh\_n] == age\_check\_tot\$age\_respondent)

# 3.6) Find the rows that contain a True inside it show\_true <- as.data.frame(apply(show\_false, 1, any))

# Now we bind the data together into the age\_check\_tot to make age\_check\_tot age\_check\_tot <- cbind(show\_true, age\_check\_tot)

# 3.7) Finally we then filter out all the TRUE to give only the rows or surveys where the respondents age did not match any of the hh\_ind age\_check\_resp <- age\_check\_tot %>% filter(`apply(show\_false, 1, any)` == FALSE)

#

# 3.8) Remove the `apply(false, 1, any)` from the dataframe so we can do it again age\_check\_tot <- age\_check\_resp %>% select(-`apply(show\_false, 1, any)`)

#

# 3.9) Age check for the head of hoursehold # Find the maximum household size hh\_n <- max(hh\_level\$`A0\_demographics/A0\_7\_hh\_size`) hh\_n <- hh\_n + 4</pre>

# 3.10) Plug in to find the rows that show FALSE (i.e. where the hh\_members dont include the respondents age) show\_false <- as.data.frame(age\_check\_tot[,5:hh\_n] == age\_check\_tot\$`A0\_demographics/A0\_2\_age\_hoh`)

# 3.11) Find the rows that contain a True inside it show\_true <- as.data.frame(apply(show\_false, 1, any))

# 3.12) Now we bind the data together into the age\_check\_tot to make age\_check\_tot age\_check\_tot <- cbind(show\_true, age\_check\_tot)

# 3.13) Finally we then filter out all the TRUE to give only the rows or surveys where the respondents age did not match any of the hh\_hoh age\_check\_hoh <- age\_check\_tot %>% filter(`apply(show\_false, 1, any)` == FALSE)

#

# 3.14) Remove all the data that is not needed rm(age\_check, age\_check\_tot, show\_false, show\_true ,age\_check\_fin)

#### # 4) Household size is the same as the reported number of household individuals.

# 4.0) Find the frequency of household number reported\_hh\_members <- as.data.frame(table(ind\_level\$`\_uuid`))

# 4.1) Rename id reported\_hh\_members <- reported\_hh\_members %>% rename(uuid = Var1)

# 4.2) Select all the reported sizes hh\_size\_errors <- dedupe\_today %>% select(`\_uuid`, start, end, today, enumerator\_code, `A0\_demographics/A0\_7\_hh\_size`) %>% rename(uuid = `\_uuid`)

# 4.3) Join the two dataframes hh\_size\_errors <- left\_join(reported\_hh\_members,hh\_size\_errors, by = "uuid")

```
# 5.4) Filter only those that matching = FALSE
hh_size_errors <- hh_size_errors %>%
mutate(matching = Freq == `A0_demographics/A0_7_hh_size`) %>%
filter(matching == FALSE)
```

#### # 5) Make sure that at least one hh member has a Yes if there is a child under 5 years.

# 5.0) Select all columns that are needed and rename them. child\_care <- ind\_level %>% select(`\_uuid`, ,enumerator\_code,`A0\_demographics/A\_0\_8\_hh\_composition/vulnerability/vulnerability\_4\_yn` ,`A0\_demographics/A\_0\_8\_hh\_composition/age\_member`)%>% rename(age = `A0\_demographics/A\_0\_8\_hh\_composition/age\_member`, care = `A0\_demographics/A\_0\_8\_hh\_composition/vulnerability/vulnerability\_4\_yn`)

# 5.1) Create a young column to find all the ages that are below 5 and create a new column that equals True. young <- child\_care %>% filter(age <= 5) %>% mutate(needs\_carer = T) %>% select(`\_uuid`, needs\_carer)

# 5.2) Now I need to bind this into the child\_care filter, this just gives a list of values where you can see who needs to go where.
# This will also change the yes to 1 and no to 0 so i can sum each by \_uuid and flag anything that is < 1.</p>
child\_care <- child\_care %>%
left\_join(young, by = "\_uuid") %>%
filter(!is.na(needs\_carer), !is.na(care)) %>%
mutate(care = ifelse(care == "No".0.1)) %>%

# 5.2) Remove superfluous data. rm(young)

filter(care < 1)

# 6) Has anyone in the household given birth in the past year. Make sure that there is a hh\_member very young.

# 6.0) Gain a dataframe of all the data to then analyse. # Find all ages household level age\_check <- hh\_level %>%

select(`\_uuid`, enumerator\_code,start,end,today,interview\_type,`pre\_survey\_inperson/age\_respondent`, `pre\_survey\_phone/pre\_survey\_respondent\_eligibility/respondent\_age`,`A0\_demographics/A0\_2\_age\_hoh`,`A3\_health/A3\_2\_births\_in\_the\_past\_ye ar`)

# 6.1) Bind ages into age\_respondent (there should be a gap ) age\_check\$age\_respondent <- ifelse(lis.na(age\_check\$`pre\_survey\_inperson/age\_respondent`), age\_check\$`pre\_survey\_inperson/age\_respondent`, age\_check\$`pre\_survey\_phone/pre\_survey\_respondent\_eligibility/respondent\_age`)

# 6.2) Select only the useful information

age\_check <- age\_check %>% select(`\_uuid`, enumerator\_code,start,end,today,interview\_type, age\_respondent, `A0\_demographics/A0\_2\_age\_hoh`,`A3\_health/A3\_2\_births\_in\_the\_past\_year`)

# 6.3) Find all the ages at individual level age\_check\_ind <- dedupe\_today %>% select(`\_uuid`, contains("age\_member"))

# 6.4) Bind together age\_check\_ind and age\_check births <- left\_join(age\_check, age\_check\_ind, by = "\_uuid")

# 6.5) delete all superfulous data rm(age,age\_check,age\_check\_ind,age\_check\_tot, ages)

#\_

# 6.6) Transform the data into a readable format # if births in the last year == "yes" and there is an NA in is.na(age\_member\_in\_months) births\_errors\_yes <- births %>%

filter(`A3\_health/A3\_2\_births\_in\_the\_past\_year` == "yes", is.na(`A0\_demographics/A\_0\_8\_hh\_composition[1]/age\_member\_in\_months`))

# 6.7) If births in the last year == "no" and there is a number present in the age\_member\_in\_months. Then there is something wrong as this household has aquired a child from nowhere. births errors no <- births %>%

filter(`A3\_health/A3\_2\_births\_in\_the\_past\_year` == "no", `A0\_demographics/A\_0\_8\_hh\_composition[1]/age\_member\_in\_months` < 13)

# 6.8) Remove superfluous data rm(births)

# Part : 7 # Title : Create a finalised sheet.

# 7.0) Create question\_name column and reason\_for\_flagging Then i can make my master sheet with (uuid, question\_name, flagged value, reason\_for\_flagging, response, changed\_to)

# 7.1) Makes list of unusually fast interviews
Afast <- fast %>%
mutate(question\_name = "time difference too slow", explanation = "the survey was too fast") %>%
select(`\_uuid`,enumerator\_code, question\_name, explanation)

#7.2) Make a list of unusually slow interviews

Aslow <- slow %>%

mutate(question\_name = "time difference too high", explanation = "the survey was too long") %>%
select(`\_uuid`,enumerator\_code, question\_name, explanation)

# 7.3) Flag up any thing that displays other in hh\_level Aother\_hh <- other\_hh %>% mutate(question\_name = "") %>% select(`\_uuid`,enumerator\_code, question\_name, explanation)

# 7.4) Flag up any thing that displays other in ind\_level Aother\_ind <- other\_ind %>% mutate(question\_name = "") %>% select(`\_uuid`,enumerator\_code, question\_name, explanation)

# 7.5) All the enumerators age do not match the hh\_ind ages.

Aage\_check\_resp <- age\_check\_resp %>%

mutate(question\_name = "respondent\_age", explanation = "age of respondent does not match any of the ages observed in the individual information (possible mistake of respondent not giving information on him/herself, thus missing a household members information") %>% select(`\_uuid`,enumerator\_code, question\_name, explanation)

# 7.6) All the hoh age do not match the hh\_ind ages. Aage\_check\_hoh <- age\_check\_hoh %>% mutate(question\_name = "A0\_2\_age\_hoh", explanation = "age of head of household does not match any of the ages observed in the individual information (possible mistake head of household not giving information on him/herself, thus missing a household members information") %>% select(`\_uuid`,enumerator\_code, question\_name, explanation)

# 7.7) GPS should be in the area (unique)
Agps\_duplications <- gps\_duplications %>%
mutate(question\_name = "gps", explanation = "interviews that both have the same gps coordinates") %>%
select(`\_uuid`,enumerator\_code, question\_name, explanation)
# 7.8) Household size is the same as the reported number of household individuals.
Ahh\_size\_errors <- hh\_size\_errors %>%
mutate(question\_name = "A0\_7\_hh\_size", explanation = "the reported household size does not match the number of household members we have
information on") %>%
rename(`\_uuid`,enumerator\_code, question\_name, explanation)

# 7.9) Make sure that at least one hh member has a Yes if there is a child under 5 years Achild\_care <- child\_care %>% mutate(question\_name = "vulnerability\_4\_yn", explanation = "there is a child under 5 who has no stated guardian over the age 12") %>% select(` uuid`.enumerator code, guestion name, explanation)

# 7.10) Has anyone in the household given birth in the last year = yes, then there should be another child that is under a year old (Missing infant from data)

Abirths\_errors\_yes <- births\_errors\_yes %>%

mutate(question\_name = "A3\_2\_births\_in\_the\_past\_year", explanation = "a household member has reported to have given birth, however there is no infant reported under the age of 1") %>%

select(`\_uuid`,enumerator\_code, question\_name, explanation)

# 7.11) Has anyone in the household given birth in the last year = no, then there should NOT be another child that is under a year old (Additional infant in household)

Abirths\_errors\_no <- births\_errors\_no %>%

mutate(question\_name = "A3\_2\_births\_in\_the\_past\_year", explanation = "there is no one in the household who has reported to have given birth, however there is an infant reported under the age of 1") %>%

select(`\_uuid`,enumerator\_code, question\_name, explanation)

7.12) Bind together

final\_reading <-

rbind(Afast,Åslow,Aage\_check\_resp,Aage\_check\_hoh,Agps\_duplications,Ahh\_size\_errors,Achild\_care,Abirths\_errors\_yes,Abirths\_errors\_no) rm(Afast,Aslow,Aage\_check\_resp,Aage\_check\_hoh,Agps\_duplications,Ahh\_size\_errors,Achild\_care,Abirths\_errors\_yes,Abirths\_errors\_no)

# Part : 9

# Title : Downloading and Appending to the old sheet

# 9.1) Final download the old datasets and label them as old.

fast\_old <- read\_excel("C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/7) issues log/issues\_log.xlsx", sheet = "fast")

slow\_old <- read\_excel("C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/7) issues log/issues\_log.xlsx", sheet = "slow")

other\_hh\_old <- read\_excel("C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/7) issues log/issues\_log.xlsx", sheet = "other\_hh")

other\_ind\_old <- read\_excel("C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/7) issues log/issues\_log.xlsx", sheet = "other\_ind")

age\_check\_resp\_old <- read\_excel("C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/7) issues log/issues\_log.xlsx", sheet = "age\_check\_resp")

age\_check\_hoh\_old <- read\_excel("C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/7) issues log/issues\_log.xlsx", sheet = "age\_check\_hoh")

gps\_duplications\_old <- read\_excel("C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/7) issues log/issues\_log.xlsx", sheet = "gps\_duplications")

hh\_size\_errors\_old <- read\_excel("C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/7) issues log/issues\_log.xlsx", sheet = "hh\_size\_errors")

child\_care\_old <- read\_excel("C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/7) issues log/issues\_log.xlsx", sheet = "child\_care")

births\_errors\_yes\_old <- read\_excel("C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/7) issues log/issues\_log.xlsx", sheet = "births\_errors\_yes")

births\_errors\_no\_old <- read\_excel("C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/7) issues log/issues\_log.xlsx", sheet = "births\_errors\_no")

final\_reading\_old <- read\_excel("C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/7) issues log/issues\_log.xlsx", sheet = "final\_reading")

# 9.2) Bind the old datasets issues with the issues found today.

fast <- rbind(fast, fast\_old)

slow <- rbind(slow, slow\_old)</pre>

other\_hh <- rbind(other\_hh, other\_hh\_old)

other\_ind <- rbind(other\_ind, other\_ind\_old)

age\_check\_resp <- rbind(age\_check\_resp,age\_check\_resp\_old)</pre>

age\_check\_hoh <- rbind(age\_check\_hoh, age\_check\_hoh\_old)</pre>

gps\_duplications <- rbind(gps\_duplications, gps\_duplications\_old)

hh\_size\_errors <- rbind(hh\_size\_errors, hh\_size\_errors\_old)

child\_care <- rbind(child\_care, child\_care\_old)</pre>

births\_errors\_yes <- rbind(births\_errors\_yes, births\_errors\_yes\_old)</pre>

births\_errors\_no <- rbind(births\_errors\_no, births\_errors\_no\_old)</pre>

final\_reading <- rbind(final\_reading,final\_reading\_old)</pre>

# 9.3) Create sheets for each and every error tab

sheets <- list("fast" = fast, "slow" = slow, "other\_hh" = other\_hh, "other\_ind" = other\_ind, "age\_check\_resp" =
age\_check\_resp, "age\_check\_hoh" = age\_check\_hoh,"gps\_duplications" = gps\_duplications, "hh\_size\_errors" =
hh\_size\_errors, "child\_care" = child\_care,"births\_errors\_yes" = births\_errors\_yes, "births\_errors\_no" = births\_errors\_no,
"final\_reading" = final\_reading)</pre>

# 9.4) Read and replace the excel sheet into the issues log

write\_xlsx(sheets, "C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/7) issues log/issues\_log.xlsx")

# Part : 10

# Title : Downloading and Appending to create the Cleaning Log ( a mix of Daily Spatial Verification and Issues Log)

sheets <- list("DSV\_AllToDate" = DSV\_AllToDate, "final\_reading" = final\_reading, "other\_hh" = other\_hh, "other\_ind" = other\_ind)

write\_xlsx(sheets, "C:/Users/krull/Desktop/IMPACT start/1) Work/Automation deposit/8) cleaning log/cleaning\_log.xlsx")