# Research Terms of Reference
**Humanitarian Situation Monitoring (HSM)**
**AFG2109**
**Afghanistan**

**February 2022**
**Version 1**

**REACH** *Informing more effective humanitarian action*

## 1.    Executive Summary

| | | | | | |
|---|---|---|---|---|---|
| **Country of intervention** | Afghanistan | | | | |
| **Type of Emergency** | X | Natural disaster | X | Conflict | ☐ Other *(specify)* |
| **Type of Crisis** | X | Sudden onset | ☐ | Slow onset | ☐ Protracted |
| **Mandating Body/ Agency** | *United Nations Office for the Coordination of Humanitarian Affairs (OCHA)* | | | | |
| **IMPACT Project Code** | 02- ARI | | | | |
| **Overall Research Timeframe** | February 2021 - November 2022 | | | | |

| **Research Timeframe** | 1. Pilot/ training (R1): 13/02/2022 | 6. Preliminary presentation: N/A |
|---|---|---|
| | 2. Data collection started (R1): 21/02/2022 | 7. Outputs sent for validation (R1): 08/05/2022 |
| | 3. Data collected (R1): 31/03/2022 | 8. Outputs published (R1): 16/05/2022 |
| | 4. Data analysed (R1): 14/04/2022 | 9. Final presentation: To be scheduled between 21/04/2022 – 31/05/2022. |
| | 5. Data and analysis sent for validation (R1): 14/04/2022 | |

| **Number of assessments** | ☐ | Single assessment (one cycle) |
|---|---|---|
| | X | Multi assessment (more than one cycle) Round 1 (R1): 21/02/2022 – 31/03/2022 Round 2 (R2): August 2022 Round 3 (R3): November 2022 |

| **Humanitarian milestones** | **Milestone** | | **Deadline** |
|---|---|---|---|
| | ☐ | Donor plan/strategy | _ _/_ _/_ _ _ _ |
| | X | Inter-cluster plan/strategy | 16/05/2022 – Multi-Sectoral severity map + Sectoral severity maps[1] + Formatted analysis to inform emergency operations in the humanitarian response + Situation Overview of multisectoral and sectoral severity analysis at national and districts levels + Key Findings presentation to the Inter-Cluster Coordination Team (ICCT) |
| | ☐ | Cluster plan/strategy | _ _/_ _/_ _ _ _ |
| | ☐ | NGO platform plan/strategy | _ _/_ _/_ _ _ _ |
| | ☐ | Other (Specify): | _ _/_ _/_ _ _ _ |

---

[1] Education in Emergencies (EiE), Emergency Shelter and NFI (ESNFI), Food Security and Agriculture (FSAC), Health, Nutrition, Protection, and Water, Sanitation, and Hygiene (WASH).

| Audience Type & Dissemination | Audience type | Dissemination |
|---|---|---|
| | X Strategic | □ General Product Mailing |
| | X Programmatic | X Cluster and Working Group Mailing[2] |
| | X Operational | X Presentation of Key Findings to ICCT members |
| | □ [Other, Specify] | X Website Dissemination (REACH Resource Centre) |
| | | □ [Other, Specify] |

| Detailed dissemination plan required | □ | Yes | X | No |
|---|---|---|---|---|

| General Objective | To regularly collect and then triangulate information regarding service provision, sectoral needs, and vulnerabilities in Afghan communities to then support geographical and sectoral prioritizations within the 2022 humanitarian response in Afghanistan.[3] [4] HSM also seeks to complement the mid-year and annual household-level 2022 Whole of Afghanistan Assessments (WoAA 2022s) which assess the multi-sectoral and sectoral needs of target populations in the Humanitarian Needs Overview (HNO). |
|---|---|
| Specific Objective(s) | Among assessed settlements in assessed districts, HSM specifically aims to: <br> 1. With the complementary analysis from the mid-year WoAA 2022 and annual WoAA 2022, develop a severity index of sectoral needs and composite index of multi-sectoral need to define emergency need and then identify districts in emergency need. <br> 2. Determine service provision gaps that may drive multi-sectoral and sectoral need – with particular attention to emergency need. <br> 3. Determine vulnerabilities (e.g. shock, migration and reliance on unsustainable income sources) and the coping strategies (e.g. debt and child labor) that may drive multi-sectoral and sectoral need – with particular attention to emergency need. [5] <br> 4. With the complementary analysis from the mid-year WoAA 2022 and annual WoAA 2022, determine multi-sectoral and sectoral needs in assessed settlements in assessed districts – with particular attention to emergency need. <br> 5. Conduct longitudinal analyses to monitor the evolution (e.g. escalations and descalations) of service provisions, vulnerabilities, and needs (multi-sectoral and sectoral) at national and district levels – with particular attention to the evolution of emergency need. |
| Research Questions | 1. What are the service provision gaps in the assessed settlements in assessed districts? <br> 2. What are the vulnerabilities and coping strategies of the assessed settlements in assessed districts? |

---

[2] Accountability for Affected Persons working group (AAPWG), Camp Coordination Camp Management working group (CCCM), Cash and Voucher working group (CVWG), Disability and Inclusion working group (DIWG) EiE, FSAC, Gender in Humanitarian Action (GiHA), Health, Nutrition, Protection and WASH.

[3] Assessed services include services directly relevant to the Education in Emergencies (EiE), Emergency Shelter and NFI (ESNFI), Food Security and Agriculture (FSAC), Health, Nutrition, Protection, and Water, Sanitation, and Hygiene (WASH) clusters. This includes educational services, health services and drinking water infrastructure.

[4] Sectors include the Education in Emergencies (EiE), Emergency Shelter and NFI (ESNFI), Food Security and Agriculture (FSAC), Health, Nutrition, Protection, and Water, Sanitation, and Hygiene (WASH) sectors.

[5] Shocks can include active conflict, anticipated conflict, pandemics or natural disasters. Natural disasters can include avalanches, earthquakes, floods, heavy snowfall, locusts/pests and persistent lack of rain.

3. What are the multi-sectoral and sectoral needs of assessed settlements in assessed districts?
4. Based on sectoral and multi-sectoral indices, what is emergency need? And what are the districts and sectors in emergency need?
5. How are service provisions and needs evolving between HSM rounds (February 2022 – November 2022)?
6. What are triggers that could be used to determine areas at risk of acute emergency?

| | |
|---|---|
| **Geographic Coverage** | All 420 districts across the 34 provinces of Afghanistan |
| **Secondary data sources** | 2021 OCHA Natural Disaster Database<br>Afghanistan: Famine Early Warning System Network (FEWS NET) Projection for October 2021 – January 2022<br>Afghanistan: Intergrated Food Security Phase Classification (IPC) Projection for November 2021 – March 2022<br>Displacement Tracking Matrix (DTM) Baseline Mobility Assessment (BMA) (unpublished)<br>OCHA Afghanistan Sectoral Severity Maps<br>Whole of Afghanistan Assessment (WoAA 2022) 2021 Sectoral and Multi-Sectoral Severity analysis<br>2022 Afghanistan Humanitarian Needs Overview (HNO) |

| **Population(s)** | □ | IDPs in camp | □ | IDPs in informal sites |
|---|---|---|---|---|
| | X | IDPs in host communities | □ | IDPs [Other, Specify] |
| | □ | Refugees in camp | □ | Refugees in informal sites |
| | X | Refugees in host communities | □ | Refugees [Other, Specify] |
| | X | Host communities | X | Populations in emergency need [6] |

| **Stratification** | X | Geographical #: 34 provinces [7]<br>Population size per strata is known? □ Yes X No | □ | Group #: _ _ _<br>Population size per strata is known?<br>□ Yes □ No | □ | *[Other Specify]* #: _ _<br>Population size per strata is known?<br>□ Yes □ No |
|---|---|---|---|---|---|---|

| **Data collection tool(s)** | X | Structured (Quantitative) | □ | Semi-structured (Qualitative) |
|---|---|---|---|---|

| | **Sampling method** | **Data collection method** |
|---|---|---|
| **Structured data collection tool** | **X Purposive**<br><br>□ Probability / Simple random<br><br>□ Probability / Stratified simple random<br><br>□ Probability / Cluster sampling<br><br>□ Probability / Stratified cluster sampling<br><br>**X Snowballing** | **X Key informant interview (Target #): 10,189 interviews across 34 provinces**<br><br>□ Group discussion (Target #):_ _ _ _ _<br><br>□ Household interview (Target #):_ _ _ _ _<br><br>□ Individual interview (Target #):_ _ _ _ _<br><br>□ Direct observations (Target #):_ _ _ _ _<br><br>□ [Other, Specify] (Target #):_ _ _ _ _ |
| **Data management platform(s)** | X | IMPACT | □ | UNHCR |
| | □ | [Other, Specify] | | |

---

[6] The ultimate population of interest are the communities in districts found to be in emergency need. However, for Round 1 from February 2022 – March 2022, HSM will cover all districts in all 34 provinces of Afghanistan to provide a baseline. Households in these districts may include cross-border returnee households, IDPs households, host households and Pakistani refugee households. In future rounds, HSM may adapt this methodology to accommodate changes within the response (including changes in target population groups or geographical access) in-country. If changes are substantial, a methodology note will be drafted and appended to this ToR.

[7] Comparisons between provinces will be between non-representative data.

| Expected ouput type(s) | □ | Situation overview #: _ _ | □ | Report #: _ _ | □ | Profile #: _ _ |
|---|---|---|---|---|---|---|
| | □ | Presentation (Preliminary findings) #: _ _ | X | Presentation (Final) #: 1 | X | Situational Overivew #: 1 sectoral needs and service gaps at national and district levels covering EiE, ES-NFI, FSAC, Health, Nutrition, Protection, and WASH findings |
| | □ | Interactive dashboard #:_ | □ | Webmap #: _ _ | X | Map #: 7 service gaps and sectoral needs at province level for EiE, ES-NFI, FSAC, Health, Nutrition, Protection, and WASH |
| | X | Formatted Analysis #: 1 Service gaps and sectoral needs at national and district levels for EiE, ES-NFI, FSAC, Health, Nutrition, Protection, and WASH | | | | |
| Access | X | Public (available on REACH resource center and other humanitarian platforms) | | | | |
| | □ | Restricted (bilateral dissemination only upon agreed dissemination list, no planned publication on REACH or other platforms at this stage) | | | | |
| Visibility *Specify which logos should be on outputs* | **REACH** | | | | | |
| | **Donor:** US Bureau for Humanitarian Aid (BHA) | | | | | |
| | **Coordination Framework:** OCHA | | | | | |
| | **Partners:** N/A | | | | | |

## 2. Rationale

### 2.1 Background

Endemic poverty and decades of conflict have converged with COVID-19, natural disasters, and recent changes in governance to generate disparate needs and vulnerabilities within Afghanistan. The year-on-year increase, from 17.7 million in 2021 to 22.1 million in 2022, in the projected number of people in acute and multi-sectoral need highlights the overall necessity for the identification of priority areas that urgently require humanitarian assistance via joint multi-sectoral analysis.

With the recent political-economical shifts in-country, average incomes are estimated to be decreasing whilst prices of key goods, especially food items, are increasing, driving the risk of greater food insecurity and/or debt accumulation. The Whole of Afghanistan Assessment (WoAA 2022) 2021 reported that 68% of assessed households said that they had earned a lower income than usual in the 30 days prior (compared to 40% in WoAA 2022 2020). Separately, the Joint Market Monitoring Initiative (JMMI) reported that the average food basket price escalated by 7% between June 2021 and November 2021. WoAA 2022 2021, collected from August 2021 – October 2021, also reported that more than half of all households reportedly struggled to obtain food or money to buy food and at least a third (37%) were found to have 'Poor' Food Consumption Scores (FCS) – compared to 23% of all households in WoAA 2022 2020.

Following the official onset of drought in June 2021, 80% of the country remains in severe or serious drought. Drought, particularly among rural communities reliant on agriculture and livestock as key income sources, is driving economic precarity and food insecurity. WoAA 2022 2021 found that drought-affected households more frequently reported loss of income (81% compared to 67% non-drought affected), lack of access to food, and wider use of emergency coping strategies compared to non-drought affected households (57% compared to 37%). It is likely that natural disasters will continue to drive need - the Notre Dame Global Adaptation Index ranks Afghanistan as the 11th least prepared country against climatic shocks

and the 10th most vulnerable country in the world to climate change. The Food and Agriculture Organization of the United Nations (FAO) is planning for recurrent droughts into 2030.

The rapidity with which humanitarian needs are subsequently expected to evolve calls for analysis at regular intervals to support geographical and sectoral prioritizations within the humanitarian response in Afghanistan.

## 2.2 Intended impact

HSM seeks to inform both the geographical and sectoral prioritization of emergency needs and any subsequent emergency interventions by actors in the response – including OCHA and clusters – by monitoring the evolution of service gap provisions, vulnerabilities, coping strategies and needs of assessed settlements in assessed districts. HSM would complement both the mid-year and annual household-level 2022 Whole of Afghanistan Assessments (WoAAs 2022) by filling in information gaps on a regular basis.

# 3.    Methodology

## 3.1 Methodology overview

HSM Round 1 will assess all 420 districts in all 34 provinces in Afghanistan. In future rounds, HSM may adapt to accommodate changes within the response (including changes in target population groups or geographical access) in-country. If changes are substantial, a methodology note will be drafted and appended to this ToR.

Enumerators will conduct approximately 10,189 face-to-face or area of knowledge (AoK) Key Informant Interviews (KIIs), according to accessibility levels across all districts in all 34 provinces of Afghanistan. KI networks will be developed through a snowballing approach to cover the entirety of the selected districts, stratifying each district's network by basic service unit (BSU)[8] to ensure a minimum of 3 KIs per BSU, each with information covering a different settlement than the others[9] and a minimum of 10% of settlements covered per district.

The structured questionnaire covers service provisions, leadership structures and sectoral needs at the settlement level. The questionnaire aims to ensure complementarity with the annual, household-level Whole of Afghanistan Assessment (WoAA 2022).

### Population of interest

The ultimate population of interest are the settlements in districts found to be in emergency need. However, for Round 1 from February 2022 – March 2022, HSM will assess settlements in  all 420 districts in all 34 provinces of Afghanistan to provide a baseline. In future rounds, HSM will adapt this methodology to accommodate both changes within the response (including changes in target population groups or geographical access) in-country and to re-assess districts found to be in emergency need during Round 1. A methodology note will then be drafted and appended to this ToR.

## 3.2 Secondary data review

To triangulate against HSM findings and inform REACH-designed sectoral and multi-sectoral need indices, HSM will rely on the following secondary data sources:

- 2021 OCHA Natural Disaster Database
- Afghanistan: Famine Early Warning System Network (FEWS NET) Projection for October 2021 – January 2022
- Afghanistan: Intergrated Food Security Phase Classification (IPC) Projection for November 2021 – March 2022
- International Office of Migration (IOM) Emergency Events Tracking (EET)
- OCHA Afghanistan Sectoral Severity Maps
- Whole of Afghanistan Assessment 2021 Sectoral and Multi-Sectoral Severity analysis
- 2022 Afghanistan Humanitarian Needs Overview (HNO)
- WorldPop Database

---

[8] BSUs are defined as economic and geographic service units, each of which representing settlements and populations with similar levels of service access and humanitarian needs.
[9] Unless less than 3 settlements are present within the BSU.

### 3.3 Primary Data Collection

HSM will rely on an adaption of the Hard to Reach (HtR) methodology. At least three key informant interviews (KIIs), reporting each on a different settlement, are conducted in each Basic Service Unit (BSU). To determine a BSU, face-to-face participatory mapping of settlements and services available are conducted prior to data collection by REACH enumerators with community leaders. Key informants (KIs), e.g. community leaders, teachers, doctors – individuals with extensive knowledge of the basic services and humanitarian needs within assessed settlements, are identified by enumerators through local councils during the participatory mapping process. REACH has mapped every BSU in every district in every province of Afghanistan as of January 2022.

Once an initial KI is identified during the participatory mapping process, a district network of KIs is subsequently developed via the snowballing approach. KIIs will be conducted in person by enumerators (with the oversight of regionally-based Senior Field Officers). When access constraints will prevent in-person interviews, interviews will be conducted with former residents with current knowledge of the settlement that they are reporting on by enumerators (with the oversight of regionally-based Senior Field Officers). This knowledge may draw from either regular contact with households in the assessed settlement and/or their departure was within 3 months of data collection (Area of Knowledge methodology).

Prior to data collection, REACH field teams will be piloting the tool and a training of enumerators will take place to ensure that all parties are familiar with the tool and the application (Kobo tool) used for data collection. Following SFO and enumerator trainings, data collection will be from 21 Feb – 31 Mar. Each enumerator will collect between 5 – 7 KIIs per day with a total of 218 KIIs per day.

### 3.4 Data Processing & Analysis

Throughout data collection, data cleaning and verification will take place daily. Feedback will be provided directly to REACH Senior Field Officers (SFOs) by the REACH Operations Program Manager. REACH SFOs will then inform the province REACH supervisor overseeing the data collection team or inform the data collection team directly. Data quality checking using the IMPACT Data Cleaning Minimum Standards Checklist will be used to outline the process and criteria for data deletion and determine staff responsibilities. Data cleaning will focus on identifying outliers in the data, contradictory or unlikely response options (logical inconsistencies), and suspicious patterns from enumerators. A cleaning log of all changes will be kept and will be available upon request after the REACH publication of datasets. Please refer to Annex 2 for further details.

Relying on R, REACH will draw on the Data Analysis Plan (DAP), sectoral need composites, and multi-sectoral need composites to offer key (indicative) findings of needs at district level, with data weighted based on the numbers of settlement per district. To address the research questions regarding emergency need, REACH will also develop a severity index of sectoral needs and composite index of multi-sectoral need and analyse the evolution of services and need over time. Tabulated analysis at district levels, including sectoral and multi-sectoral analysis relevant to emergency planning as well as and a multi-sectoral severity map, and an overall national overview will then be drafted and presented to OCHA and clusters.

## 4.    Key ethical considerations and related risks

The proposed research design meets / does not meet the following criteria:

| *The proposed research design…* | *Yes/ No* | *Details if no (including mitigation)* |
|---|---|---|
| … Has been coordinated with relevant stakeholders to **avoid unnecessary duplication** of data collection efforts? | Yes | |
| … **Respects respondents, their rights and dignity** (*specifically by: seeking informed consent, designing length of survey/ discussion while being considerate of participants' time, ensuring accurate reporting of information provided*)? | Yes | |

| | | |
|---|---|---|
| … Does not **expose data collectors to any risks as a direct result** of participation in data collection? | Yes | |
| … Does not **expose respondents / their communities to any risks as a direct result** of participation in data collection? | Yes | |
| … Does not involve **collecting information on specific topics which may be stressful and/ or re-traumatising** for research participants (both respondents and data collectors)? | Yes | |
| … Does not involve **data collection with minors** i.e. anyone less than 18 years old? | Yes | |
| … Does not involve **data collection with other vulnerable groups** e.g. persons with disabilities, victims/ survivors of protection incidents, etc.? | No | While vulnerable groups are not explicitly interviewed for this assessment, it is possible that vulnerable individuals may be interviewed. Enumerators are prepared and trained for this. |
| … Follows IMPACT SOPs for management of **personally identifiable information**? | Yes | |

# 5. Roles and responsibilities

**Table 3: Description of roles and responsibilities**

| Task Description | Responsible | Accountable | Consulted | Informed |
|---|---|---|---|---|
| *Research design* | Assessment Officer (AO) | Research Manager (RM) | IMPACT HQ – Research Design and Data Unit (RDDU) | Country Coordinator (CC) |
| *Supervising data collection* | Operations Program Manager (PM) | Country Operations Manager (COM) | RM/ IMPACT HQ – RDDU | CC |
| *Data processing (checking, cleaning)* | Database Manager (DM) | AO | RM/ IMPACT HQ – RDDU | CC |
| *Data analysis* | DM | AO | PM/RM/ IMPACT HQ – RDDU | CC |
| *Output production* | AO | RM | PM/DM/IMPACT HQ – Research Reporting Unit (RRU) | CC |
| *Dissemination* | AO | RM | IMPACT HQ – RRU | CC |
| *Monitoring & Evaluation* | AO | RM | IMPACT HQ – Research | CC |

| | | Department (RD) | |
|---|---|---|---|
| *Lessons learned* | AO | RM | IMPACT HQ - RD | CC |

**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*

# 6. Data Analysis Plan

Please see refer to the following link.

# 7. Data Management Plan

Available on request.

# 7. Monitoring & Evaluation Plan

| IMPACT Objective | External M&E Indicator | Internal M&E Indicator | Focal point | Tool | Will indicator be tracked? |
|---|---|---|---|---|---|
| **Humanitarian stakeholders are accessing IMPACT products** | Number of humanitarian organisations accessing IMPACT services/products<br><br>Number of individuals accessing IMPACT services/products | # of downloads of x product from Resource Center | Country request to HQ | User_log | x Yes |
| | | # of downloads of x product from Relief Web | Country request to HQ | | x Yes |
| | | # of downloads of x product from Country level platforms | Country team | | □ Yes |
| | | # of page clicks on x product from REACH global newsletter | Country request to HQ | | □ Yes |
| | | # of page clicks on x product from country newsletter, sendingBlue, bit.ly | Country team | | □ Yes |
| | | # of visits to x webmap/x dashboard | Country request to HQ | | □ Yes |
| **IMPACT activities contribute to better program implementation and coordination of the humanitarian response** | Number of humanitarian organisations utilizing IMPACT services/products | # references in HPC documents (HNO, SRP, Flash appeals, Cluster/sector strategies) | Country team | Reference_log | Afghanistan HNO and HRP 2022 |
| | | # references in single agency documents | | | OCHA country documents, Cluster-specific documents |

| | | | | | |
|---|---|---|---|---|---|
| **Humanitarian stakeholders are using IMPACT products** | Humanitarian actors use IMPACT evidence/products as a basis for decision making, aid planning and delivery<br><br>Number of humanitarian documents (HNO, HRP, cluster/agency strategic plans, etc.) directly informed by IMPACT products | Perceived relevance of IMPACT country-programs | Country team | Usage_ Feedback *and* Usage_ Survey template | Usage survey to be conducted at the end of the research cycle related to all outputs, targeting at least 20 partners<br><br>Usage survey to be conducted at the end of the research cycle related to all outputs, targeting at least 20 partners |
| | | Perceived usefulness and influence of IMPACT outputs | | | |
| | | Recommendations to strengthen IMPACT programs | | | |
| | | Perceived capacity of IMPACT staff | | | |
| | | Perceived quality of outputs/programs | | | |
| | | Recommendations to strengthen IMPACT programs | | | |
| **Humanitarian stakeholders are engaged in IMPACT programs throughout the research cycle** | Number and/or percentage of humanitarian organizations directly contributing to IMPACT programs *(providing resources, participating to presentations, etc.)* | # of organisations providing resources (i.e.staff, vehicles, meeting space, budget, etc.) for activity implementation | Country team | Engagement_log | x Yes |
| | | # of organisations/clusters inputting in research design and joint analysis | | | x Yes |
| | | # of organisations/clusters attending briefings on findings; | | | x Yes |

# Annex 1: SAMPLING FRAMEWORK FOR KEY INFORMANT SURVEYS

| Province | N° of districts | N° of BSUs | N° of Settlements | N° of KI Interviews |
|---|---|---|---|---|
| Badakhshan | 28 | 277 | 2341 | 831 |
| Badghis | 7 | 68 | 1443 | 232 |
| Baghlan | 15 | 126 | 1851 | 378 |
| Balkh | 16 | 139 | 1537 | 417 |
| Bamyan | 8 | 53 | 2291 | 260 |
| Daykundi | 9 | 72 | 2479 | 316 |
| Farah | 11 | 58 | 1597 | 225 |
| Faryab | 14 | 89 | 1660 | 290 |
| Ghazni | 19 | 122 | 3763 | 496 |
| Ghor | 10 | 86 | 2582 | 310 |
| Helmand | 15 | 143 | 2456 | 451 |
| Herat | 20 | 167 | 2734 | 527 |
| Jawzjan | 11 | 44 | 708 | 146 |
| Kabul | 15 | 80 | 1307 | 250 |
| Kandahar | 18 | 145 | 1960 | 455 |
| Kapisa | 7 | 37 | 800 | 122 |
| Khost | 13 | 63 | 1544 | 193 |
| Kunar | 16 | 101 | 1420 | 303 |
| Kunduz | 10 | 121 | 1170 | 363 |
| Laghman | 6 | 54 | 946 | 170 |
| Logar | 7 | 42 | 892 | 138 |
| Maidan Wardak | 9 | 65 | 2277 | 277 |
| Nangarhar | 22 | 160 | 2226 | 483 |
| Nimroz | 6 | 36 | 697 | 112 |
| Nuristan | 8 | 46 | 662 | 138 |
| Paktika | 20 | 101 | 1862 | 320 |
| Paktya | 13 | 65 | 1578 | 220 |
| Panjsher | 8 | 39 | 629 | 122 |
| Parwan | 10 | 62 | 1543 | 207 |
| Samangan | 7 | 60 | 811 | 180 |
| Sar-e-Pul | 7 | 42 | 1003 | 154 |
| Takhar | 17 | 219 | 1848 | 657 |
| Uruzgan | 7 | 52 | 806 | 163 |
| Zabul | 11 | 92 | 1365 | 283 |
| **Total** | **420** | **3,126** | **54,788** | **10,189** |

# Annex 2: DATA CLEANING STANDARD OPERATING PROCEDURES (SoPs)

*Below are the data cleaning standard operating procedures (SoPs) for HSM key informant surveys.*

## OVERVIEW OF DAILY RESPONSIBILITIES

### Program Manager (PM)
- Responsible for communicating security concerns from Senior Field Officers (SFOs) and implementing partners to larger Senior Management Team (SMT), who can then decide on whether or not to change the sampling framework and communicate that to Assessment Officers (AOs) and Database Manager (DM)
- Knowledgeable of AO to SFO communications regarding data cleaning issues via Skype group and e-mails

### Senior Field Officers (SFOs)
- Responsibility for Afghanistan regions divided between SFOs, each assigned to contact enumerator team leaders, individual enumerators and/or implementing partners if any issues with the data
- In constant communication between the database manager (DM) and field SFOs / individual enumerators regarding issues with data collection and data quality issues
- Ultimately responsible for progress tracking in google sheets
- Responsible for following up on feedback given to enumerators via google sheets

### Database Manager (DM)
- Responsible for downloading, deleting and anonymizing raw data and passing to S/DBOs for cleaning
- Runs daily R data checking script with clean data and raw data from most recent day to identify errors for SFOs to follow up on
- Checks/reviews the logged issues, and then the feedback that will be provided SFOs and AOs by Lead DBO Updates Daily Progress Tracker (HTML)
- Makes final call on survey deletions. Updates deletion columns in online tracking sheets
- Responsible for overseeing changes to sampling framework and adjusting sampling targets accordingly

### Lead Database Officer (Lead DBO)
- Responsible for supervising the visual check, split data for visual checks for J/DBOs and consolidate all the visual check results in single file.
- Sends cleaning result to AOs for providing feedback for and follow up on it
- Consolidate all cleaning logs which feedback provided for them

### Assessment Officers (AOs)
- Responsible for reviewing cleaning result on daily basis and providing feedback to either DM, SFOs, GISO
- Responsible for compiling all subsequent feedback from DM, SFOs, and GISO and sending to DM
- In communication regarding security or logistical concerns that change sampling framework – communication lead by PM, DM, and SMT

### Database Officers (J/DBOs)
- Visually check for patterns in the data in questions to see if there is any evidence of enumerators developing habits of always entering the same values
- DBO's responsible for ensuring **daily backups of their cleaned data** occur and that proper file naming protocol is followed for cleaned data and cleaning log AND cleaning of data at end of data collection
- **DBOs** communicate all data issues to **Lead DBO,** who communicates feedback and issues to responsible **AO** to contact responsible SFO, who contacts field teams and individual enumerators to clarify any issues with the data

### Individual Enumerators

- Ensure phones are fully charged prior to next day of data collection
- Ensure phones are set to the correct time and date prior to data collection. Achieved with steps below: Settings→General management→Date and time→Automatic date and time **AND** Use 24-hour format **ON**
- Collect coordinates and finalize survey after asking final questions
- Upload forms to the kobo server daily
- <u>Enumerator ID</u> is in the format 'organization-province-number' ie. reach-badakhshan-5

### <u>GIS Officer (GISO)</u>
- Cross-checks household locations with village selected by enumerator. If location issues or spatial duplicates found, they are to be brought up with DM, who will contact responsible SFO.

### **Additional information for DBOs:**

- We've been automating most of the data processing stages and encourage our DBOs to implement their new ideas for developing this process. In addition, DBOs should familiarize their self with all data cleaning scripts, in particular, translation, replacing cleaning log and pattern check logger, time check and etc.
- DBOs responsible for familiarizing themselves with the constraints of the **kobo data collection tool**
- DBOs to have the **latest version of the kobo tool and this SOP** open to understand each question code
- The column structure must be followed in order
- Any consistent pattern issues related to specific enumerators **should be noted** by DBOs
- Translations from Dari or Pashto first harmonized to match existing categories if possible. If the received option in other has occurred in a ratio of 15% of other options, then it can be a new choice.

## DATA PROTECTION
To maintain the safety and security of both respondents and enumerators, the following procedures will be followed:

o Key informant names and phone numbers are deleted from final dataset
o Settlement location, enumerator information and interview times are deleted from final dataset
o Settlement location will be removed from processed raw dataset that is shared among AOs and PM
o Province codes, district codes, and enumerators will be used for raw dataset that is shared among AOs and PM
o Database Manager is the only individual with full access to raw dataset

## DELETION OF DATA

o Considering the deletion criteria, the deletion command will be coded in R script to apply on daily data processing
o Database Manager is responsible for deletion of surveys
o Database Manager receives cleaning logs from lead DBO and feedback from GISO, and compiles uuid of surveys to delete from dataset.

### <u>Criteria for deletion:</u>
- All surveys completed in under 25 minutes or more than 100 minutes, without reasonable explanation
- All surveys with a pattern match of 90% or higher to another survey
- All surveys with a respondent 17 years of age or younger
- All surveys without consent
- All surveys with 6 or more flags, without reasonable explanation
- Surveys located 5km or more from selected village, without reasonable explanation

## DATA CHECKING SUMMARY

### A – DUPLICATE CHECKS: PARENT WORKSHEET
**Duplicate uuid's should not be present**

| # | STEP |
|---|------|
| 1. | **Duplicate surveys are flagged for deletion**. <br><br> We have already coded this in our data processing script. If there is a duplicate, then interview is DELETED. |

## B – TIME CHECKS: PARENT WORKSHEET
**Survey should take between 30 and 90 minutes**

| # | STEP |
|---|------|
| 1. | This survey should take maximum between 30 and 90 minutes to complete. 25 minutes and under: Surveys will be DELETED as invalid. Surveys between 30 minutes and 26 minutes will be FLAGGED. Surveys between 90 minutes and 100 minutes will be FLAGGED. Any interviews that are more than 100 minutes are DELETED. |
| 2. | We're using audit files to check the duration of each survey and using R script for doing this operation. |

## C - TRANSLATIONS FROM DARI / PASHTO
**Organization and Village names may be entered in Pashto or Dari**

| # | QUESTION | ACTION |
|---|----------|--------|
| 1. | **village_other, organisation_other** <br> *Parent Worksheet* | **'Other' answers, entered initially in Dari or Pashto.** <br> Translations from Dari/Pashto done within same cell in the *Data Checking* Tab. If possible, translations are first harmonized into groups that match existing categories. If answer falls into category that already exists, the 'other' category answer is corrected in the same cell. |

## D – SPECIFIC QUESTION CHECKS
**Parent Worksheet**

| # | QUESTION | ACTION |
|---|----------|--------|
| 1. | **reason_idp_displacement** | 1. Flag if any of the following are selected, but event_conflict is NOT selected for **events_one_year**: <br> active_conflict <br> anticipated_conflict <br> 2. Flag if unemployment_poverty is selected but event_economy is NOT selected for **events_one_year** <br> 3. Flag if "no_food" is selected but **hunger_level** is "no_hunger" or "small_hunger" |

| 2. | **reason_returnee_displacement** | 1. Flag if any of the following are selected, but event_conflict is NOT selected for **events_one_year**: active_conflict anticipated_conflict 2. Flag if unemployment_poverty is selected but event_economy is NOT selected for **events_one_year** 3. Flag if "no_food" is selected but **hunger_level** is "no_hunger" or "small_hunger" |
|---|---|---|
| 3. | **forcibly_displaced_why** | 1. Flag if any of the following are selected, but event_conflict is NOT selected for **events_one_year**: active_conflict anticipated_conflict 2. Flag if unemployment_poverty is selected but event_economy is NOT selected for **events_one_year** 3. Flag if "no_food" is selected but **hunger_level** is "no_hunger" or "small_hunger" |
| 4. | **forcibly_displaced_why_return** | Flag if "access_services" is selected but **education_available** is "no". Flag if "access_services" is selected but **healthcare_available** is "no". |
| 5. | **events_one_year** | Flag if any of the following are selected: event_earthquake event_flood event_locust event_drought event_snow But not listed at province level in the [OCHA 2021 Natural Disaster database](OCHA 2021 Natural Disaster database) |
| 6. | **events_one_year** | Flag if event_economy but **change_employment** is "no_hh" |
| 7. | **food_access** | Flag if "no" but **food_price_change** is decreased_little OR decreased_lot |
| 8. | **hunger_level** | Flag if **hunger_level** is "bad_hunger" or "worst_hunger" but **food_price_change** is decreased_little OR decreased_lot |
| 9. | **coping_mechanisms** | Flag if **"coping_no"** is selected but "no" is selected for **food_access** |
| 10. | **reason_change_farming** | Flag if **reason_change_farming** is event_flood, but **events_one_year** is not event_flood. Flag if **reason_change_farming** is event_conflict, but **events_one_year** is not event_conflict. Flag if **reason_change_farming** is event_drought, but **events_one_year** is not event_drought. Flag if **reason_change_farming** is event_locust, but **events_one_year** is not event_locust. Flag if **reason_change_farming** is event_snow, but **events_one_year** is not event_snow. Flag if **reason_change_farming** is event_earthquake, but **events_one_year** is not event_earthquake. Flag if **reason_change_farming** is event_covid, but **events_one_year** is not event_covid |
| 11. | **reason_change_livestock** | Flag if **reason_change_livestock** is event_flood, but **events_one_year** is not event_flood. Flag if **reason_change_livestock** is event_conflict, but **events_one_year** is not event_conflict. Flag if **reason_change_livestock** is event_drought, but **events_one_year** is not event_drought. Flag if **reason_change_livestock** is event_locust, but **events_one_year** is not event_locust. Flag if **reason_change_livestock** is event_snow, but **events_one_year** is not event_snow. Flag if **reason_change_livestock** is event_earthquake, but **events_one_year** is not event_earthquake. Flag if **reason_change_livestock** is event_covid, but **events_one_year** is not event_covid |
| 12. | **reason_change_employment** | Flag if **reason_change_employment** is event_flood, but **events_one_year** is not event_flood. Flag if **reason_change_employment** is event_conflict, but **events_one_year** is not event_conflict. Flag if **reason_change_employment** is event_drought, but **events_one_year** is not event_drought. Flag if **reason_change_employment** is event_locust, but **events_one_year** is not event_locust. Flag if **reason_change_employment** is event_snow, but **events_one_year** is not event_snow. Flag if **reason_change_employment** is event_earthquake, but **events_one_year** is |

| | | |
|---|---|---|
| | | not event_earthquake. Flag if **reason_change_employment** is event_covid, but **events_one_year** is not event_covid |
| 13. | **male_unemployment** | Flag if **male_unemployment** is "all" or many" but event_economy is NOT selected for **events_one_year** |
| 14. | **female_breadwinners** | Flag if **female_breadwinners** is "yes" but **female_employment** is "no" |
| 15. | **women_business** | Flag if **women_business** is "yes" but **female_employment** is "no" |
| 16. | **financial_support** | Flag if **financial_services** is "yes" but **financial_support** is "no" |
| 17. | **reason_debt** | Flag if **reason_debt** is "healthcare_food" but **coping_mechanisms** is "coping_no" |
| 18. | **reason_debt** | Flag if **reason_debt** is "cost_displacement" but **proportion_host** is "all_hh" |
| 19. | **safe_school_commute_girls** | Flag if "not_at_all_safe" but **gbv_unsafe_place** is NOT "unsafe_way_school" |
| 20. | **school_equipment** | Flag if **school_equipment** is "yes" but education_issues is "lack_of_equipment". |
| 21. | **teachers** | Flag if **teachers** is "yes" but education_issues is "lack_of_teachers". |
| 22. | **staff** | Flag if **staff** is "yes" but **healthcare_issues** is "lack_admin_staff" OR "lack_medical_staff" OR "lack_doctors" OR "lack_midwives" OR "lack_female_staff" OR "lack_specialised_staff" |
| 23. | **eviction_why** | Flag if **eviction_why** is "unable_pay" but **shelter_tenancy** is "own_shelter_doc" or "own_shelter_wdoc" |
| 24. | **heating_source** | Flag if **heating_source** is no source, but **access_fuel** is "no_hh" or "few_hh". |
| 25. | **heating_source** | Flag if **heating_source** is electricity, but **electricity_avail** is "no_hh" or "few_hh. |
| 26. | **electricity_hours** | Flag if **electricity_hours** is greater than 8 hours, but **electricity_avail** is "no_hh" or "few_hh" |
| 27. | **child_marriage** | Flag if "yes" but "coping_marriage" is not selected for **coping_mechanisms** |
| 28. | **incidents_men** OR **incidents_women** OR **incidents_boys** OR **incidents_girls** | Flag if "forced_to_leave" but **eviction_risk** is "no" |
| 29. | **incidents_men** OR **incidents_women** OR **incidents_boys** OR **incidents_girls** | Flag if "injured" but **mines** is "no" |
| 30. | **incidents_boys** | Flag if "hindered_move" but **safe_school_commute_boys** is "not_at_all_safe" |
| 31. | **incidents_women** OR **incidents_girls** | Flag if "hindered_move" but **gbv_unsafe** is "no" |
| 32. | **gbv_unsafe** | Flag if "no" but **women_water_access** is not_allowed. Flag if "no" but **women_health_access** is not_allowed. Flag if "no" but **women_market_access** is not_allowed |
| 33. | **gbv_unsafe_place** | Flag if "unsafe_water_point" but **women_water_access** is alone |
| 34. | **gbv_unsafe_place** | Flag if "unsafe_hospital" but **women_health_access** is alone |

| 35. | **gbv_unsafe_place** | Flag if "unsafe_market" but **women_market_access** is alone |
|-----|----------------------|--------------------------------------------------------------|

## E – ENUMERATOR PATTERN CHECKS

| # | QUESTION | ACTION |
|---|----------|--------|
| 1. | **Select_one** and **select_multiple** questions | Visual check that individual enumerators are not developing patterns for answers to these questions or answering the same answers for every survey. |
| 2. | **proportion_hh and proportion_hh_2** | Visual check that individual enumerators are not developing patterns regarding proportions of households. |

## G – DAILY GIS SPATIAL CHECKS
**Detailed summary found in Annex II**

| # | STEP |
|---|------|
| 1. | Database Manager provides location data of the days' surveys to GIS Officer (GISO) at the end of each day of data collection, in *csv format*. The following information is included in the spreadsheet exported to the GISO: **data collection date, uuid, village code, village name (in kobo format).** It is not necessary to include other fields in the export to the GISO. |
| 2. | GISO loads new points into ArcGIS software, along with a reference shapefile with all villages covered in the assessment. The field **'kobo_code'** in the village reference shapefile will be equivalent to the **'village'** field from the kobo entries. GISO then *performs a table join in the kobo points layer* using the two matching codes. 2 sets of coordinates will now be present in the attribute table of the kobo points. |
| 3. | GISO check for duplication of coordinates **(within 10m)** collected in kobo. A summary of duplicate uuid is exported as a table.<br><br>**Surveys with duplicate coordinates are flagged to the Database Manager to flag to SFOs who ensure information reaches individual enumerators** |
| 4. | **'XY to line'** function used from ArcGIS Toolbox to draw line between the pair of coordinates present in the point attribute table. Surveys greater than 5km **and greater than 10km** from villages where they are indicated to have taken place are flagged. |
| 5. | GISO exports a report table with uuid, village name, and distance (km) columns for surveys located greater than 5km from indicated village. **Uuid's with wrong village assigned are flagged to the Database Manager, and then to regional SFO to discuss with individual enumerators** |

## H – DAILY R SCRIPT FOR RAW DATA CHECK

Producing reports broken by region, the purpose of the daily R script check is to provide guidance to SFOs and DBOs for enumerator follow up and provide an additional check against enumerator error.

| R Script incorporates cleaned data as well as raw data from the most recent day of data collection |
| --- |

| # | QUESTION | ACTION |
| --- | --- | --- |
| 1. | **progress check** | At province level. Shows progress at the province, district, and basic service unit (BSU) level |
| 2. | **time checks** | Flags surveys with illogical time stamps including an end time that is before the start time, interviews submitted in the future, interviews that are either **under 30 minutes** or **more than 90 minutes**. |
| 3. | **survey frequency** | Flags enumerator productivity in most recent day of data collection. Enumerators with more than 7 surveys per day maybe notified to slow down and interviews may be deleted if number exceeds 7. |

## ANNEX I

*Detailed summary of Sampling Framework Changes*

- **Step 1:**
  Individual enumerator, team leader, or implementing partner raises security or logistical concern (including active conflict, natural disaster, change in mobile connectivity in area, hostility from local actors etc.) to SFO. PM then alerts DM and SMT.
- **Step 2:**
  SFO escalates security or logistical concern to PM.
- **Step 3:**
  SMT determines sampling framework change and communicates sampling change to PM, AOs, and DM. Alternatively, PM communicates sampling change to AOs and DM.
- **Step 4:**
  DM adjusts sampling framework to accommodate changes.