

## DIME Analytics Data Quality Assurance Checklist

### 0.1 Objective

In the process of acquiring data, especially primary data, many factors could lead to poor data quality. A plan to ensure quality of data is one that identifies the potential issues in the data, and tracks data quality throughout the process of acquiring data.

### Procurement

CHECKS	YES	NO	N/A
Does the TOR establish <b>minimum quality indicators</b> required for the data to be considered acceptable?			
Does the TOR establish a <b>maximum response time</b> for the survey firm to address data quality issues identified?			

## Survey Design

CHECKS	YES	NO	N/A
Does the survey have <b>relevance and constraint logic</b> set up to aid the enumerator?			
Does the survey have an <b>an ID variable</b> to identify respondents and link them to the respondents database?			
Does the survey have an <b>unique ID generator</b> for each survey submission (if survey is being done over SurveyCTO this is automatically generated as the 'key' variable.			
Does the survey have <b>duration spent</b> on completing a survey and modules within the survey?			
Has the <b>translation</b> been approved?			
Does <b>enumerator training materials</b> include explanation of informed consent, Q&A session, mock interviews and review of best practices?			
In case of projects with multiple survey instruments, does each survey have it's own detailed survey manual?			

## High-Frequency Checks (HFC)

CHECKS	YES	NO	N/A
Is there a plan in place to conduct <b>high-frequency checks</b> on the data being collected?			
Does the HFC <b>monitor survey duration</b> and start/end times and day of the week of a survey fill-out?			
Do the HFCs check for key treatment and <b>outcome variables</b> ?			
Do the HFCs check for <b>sample completeness</b> ?			
Do the HFCs report <b>duplicates, missing values and errors</b> in logic that are not automatically corrected by the survey form?			
Is the process for generating flags from HFCs <b>easy and fast to run</b> on a daily basis for FCs?			
Does the HFC process account for potential lack of <b>access to WiFi</b> or electricity for team on the field?			
Do HFCs monitor geographic location of projects based on <b>GPS coordinates</b> ?			
Has the HFC plan been communicated to all <b>stakeholders</b> in the team (incl. survey firm) with enough time before survey round?			

## Backchecks and audits

CHECKS	YES	NO	N/A
Is there a plan in place to conduct <b>back-checks</b> on the data collected?			
Does the back-checks plan include <b>variables unlikely to change</b> , as well as variables that stem from difficult or lengthy modules?			
Is there a plan to audit through <b>audio recordings</b> ? If so, is it mentioned in the informed consent?			
Does the back-checks plan include back-checking <b>all enumerators</b> ?			
Does the back-check plan include back-checking <b>20 percent of the sample</b> in the first 2 weeks and 10 percent of the sample in the following weeks?			

## Communication and Field Management

CHECKS	YES	NO	N/A
Does team have a plan for <b>communication and data quality checks feedback</b> with enumerators?			
Is there a protocol in place for the field team to <b>gather information and respond to flags</b> raised during data quality checks?			
Has the team put together <b>tracking systems</b> for enumerators to fill out as part of their daily tasks?			
Does the team have a plan for <b>reconciling data</b> from the tracking system with the responses on the server			
If feasible in the context, does the team plan on research staff <b>accompanying enumerator</b> ?			