

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 minimum quality indicators required for the data to be considered acceptable?			
Does the TOR establish a maximum response time for the survey firm to address data quality issues identified?			

Survey Design

CHECKS	YES	NO	N/A
Does the survey have relevance and constraint logic set up to aid the enumerator?			
Does the survey have an an ID variable to identify respondents and link them to the respondents database?			
Does the survey have an unique ID generator for each survey submission (if survey is being done over SurveyCTO this is automatically generated as the 'key' variable.			
Does the survey have duration spent on completing a survey and modules within the survey?			
Has the translation been approved?			
Does enumerator training materials 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 high-frequency checks on the data being collected?			
Does the HFC monitor survey duration and start/end times and day of the week of a survey fill-out?			
Do the HFCs check for key treatment and outcome variables ?			
Do the HFCs check for sample completeness ?			
Do the HFCs report duplicates, missing values and errors in logic that are not automatically corrected by the survey form?			
Is the process for generating flags from HFCs easy and fast to run on a daily basis for FCs?			
Does the HFC process account for potential lack of access to WiFi or electricity for team on the field?			
Do HFCs monitor geographic location of projects based on GPS coordinates ?			
Has the HFC plan been communicated to all stakeholders 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 back-checks on the data collected?			
Does the back-checks plan include variables unlikely to change , as well as variables that stem from difficult or lengthy modules?			
Is there a plan to audit through audio recordings ? If so, is it mentioned in the informed consent?			
Does the back-checks plan include back-checking all enumerators ?			
Does the back-check plan include back-checking 20 percent of the sample 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 communication and data quality checks feedback with enumerators?			
Is there a protocol in place for the field team to gather information and respond to flags raised during data quality checks?			
Has the team put together tracking systems for enumerators to fill out as part of their daily tasks?			
Does the team have a plan for reconciling data from the tracking system with the responses on the server			
If feasible in the context, does the team plan on research staff accompanying enumerator ?			