Beginner – Data Quality

Data Quality

This application will allow you to execute data quality rules upon your project data to check for discrepancies in your data. Pre-defined data rules that are already developed for you in the application are for such things as missing values, field validation errors and outliers for numerical fields.

A "data quality rule" is a calculation expression that must evaluate to True or False. When a rule is executed on your project data any records for which the expression evaluates False can be viewed and potentially corrected.

The Data Quality page includes detailed instructions.

Execute Rules

Execute rules one at a time using Execute, or all together using All.

The Clear button refreshes the view to the state before any rules have been executed.

Discrepancies

Rules that identify discrepancies (i.e. where your project contains data for which the calculation expression returns False) are highlighted in red.

- Click view to see the records that fail the test.

- Navigate to the record / form that contains the discrepant value by clicking on the value

- Clicking exclude will mark a discrepant value as not to be included as a discrepancy in future executions of this rule *Add a New Rule*

  Add a new rule by entering a description of the rule and the calculation expression. Then click Add.

For more information on calculation expressions see Calculated Fields. Remember that data quality rule expressions must evaluate to True or False.

Important Tip! If your expression includes a "less than" sign (<), ensure that you include a space after it (as shown above). This stops the page thinking that you're entering a potential harmful HTML tag and stripping out the remainder of your expression! Using "Less than or equal to" (<=) is fine – this problem does not occur.

Real Time Execution

Ticking the Execute in real time option is a useful mechanism for performing cross-field validation checks during data entry. Rules where this option is ticked will be execute each time you save a data entry form that contains a field that is part of the expression.

Note that the rules are executed on data entry forms only: NOT when importing data via Data Import or API, and NOT during survey data entry.