**The SASA! Together Radio Booster Study**

**Data dictionary for TWG\_merged\_baseline\_followup.dta**

**NB: This data dictionary is to be used in conjunction with TWG\_merged\_baseline\_followup.dta** and **TWG\_Questionnaire**

The baseline and follow-up surveys were cross-sectional (each conducted among different samples of community members). The same data were collected at each time-point thus allowing the two datasets to be appended. Each record corresponds to data from one community member and contains data from one data-collection time-point only (either baseline or follow-up, but not both). The survey round (baseline or follow-up) is indicated by the variable *survey\_round*. The total number of records in the dataset is 1662, comprising 644 baseline records and 1018 follow-up records.

* Variables named *q#* or *q##* correspond to survey question number Q# or Q##.
* Variables named *q##\_#* or *q##\_##* correspond to survey question number Q## for which multiple responses may be selected (‘select all that apply’). The dataset contains a binary (yes/no) variable for each response option, with the first part of the variable name (*q##*) denoting the overall question number, and the second part of the variable name (*\_#*) denoting the numerical code for the relevant response option as indicated in the questionnaire. For example, if question Q13 is a ‘select all that apply’ question with eleven yes/no options to choose from, the dataset will contain eleven corresponding variables for this question: q13\_1 to q13\_11.
* Variables named q18a# and q18b# correspond to the respective sub-components of question Q18 as indicated in the questionnaire.

Please refer to the survey questionnaire (**TWG\_Questionnaire**) to identify these variables and interpret response codes.

Variables that deviate from this naming convention or do not correspond directly to a question from the baseline or follow-up survey are listed and described in Table 1 (below).

Whether or not specific variables contain data will depend on whether or not filter questions required the respondent to skip certain questions. The column in Table 1 entitled ‘Respondents with data’ indicates which respondents should have data for the respective variable.

**Table 1: Additional variables to those in the baseline and follow-up survey, and variables deviating from the above detailed naming convention**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable name** | **N** | **Type** | **Respondents with data** | **Value labels/range** | **Variable description** |
| **Administrative variables** |  |  |  |  |  |
| survey\_round | 1,662 | Numeric | All | 1=Baseline  2=Follow-up | Data-collection time-point (Baseline or Follow-up) |
| respondent\_id | 1,662 | string | All |  | Unique respondent ID |
| interviewer\_id | 1,662 | Numeric | All | 1-25 (excluding 18 and 24) | Interviewer ID number |
| intyear | 1,662 | Numeric | All | 2022-2023 | Year of interview |
| intmonth | 1,662 | Numeric | All | 3=March  4=April  7=July  8=August | Month of interview |
| **Village-level variables** |  |  |  |  |  |
| village\_id | 1,662 | Numeric | All | 1-23 (excluding 3, 5, 6, 12 and 17) | Village ID number |
| ca\_active | 1,662 | Numeric | All | 1=No community activist (CA) in village  2=Village CA is currently inactive  3=Active CA in village | Presence and status (currently active/inactive) of SASA! Together community activist (CA) in village |
| ca\_active\_bin | 1,662 | Numeric | All | 0=No/inactive CA in village  1=Active CA in village | Presence of active SASA! Together CA in village |
| urban | 1,662 | Numeric | All | 0=Rural  1=Urban | Urban/rural designation of village (NB. Urban denotes small towns with access to amenities such as electricity, piped water supplies, improved housing and trading centres, rather than large urban centres) |
| distantvillage | 1,662 | Numeric | All | 0=Not distant  1=Distant (3-4 hour walk from other villages) | Geographical remoteness of village |
| **Respondent characteristics** |  |  |  |  |  |
| agecat | 1,661 | Numeric | All | 1=18-19 years  2=20-24 years  3=25-29 years  4=30-34 years  5=35-39 years  6=40-44 years  7=45-49 years  8=50-54 years  9=55-59 years  10=60-64 years  11=65-69 years  12=70-74 years  13=75-79 years  14=80-90 years | Categorical age variable  Exact age was calculated using year of birth (Q3) and interview date (q3 has been removed from the dataset as part of the anonymisation process, and q4 was removed due to absence of any data for this variable). Exact age was then removed from the dataset, and replaced with this categorical variable, in order to reduce risk of reidentification of individuals through unique values and/or outliers. |
|  |  |  |  |  |  |
| **Variables from questionnaire that deviate from above naming convention** |  |  |  |  |  |
| q22\_years | 1,223 | Numeric | Respondents currently in a relationship (married/living as married/regular partner) | 0-70 | Number of years respondent has been in current relationship - (part of response to Q22) |
| q22\_months | 1,223 | Numeric | Respondents currently in a relationship (married/living as married/regular partner) | 0-11 | Number of additional months (on top of reported years) that respondent has been in current relationship - (part of response to Q22) |
| q27\_total | 1,661 | Numeric | All | 0-15 | Number of children up to the age of 18 respondent is responsible for - (part of response to Q27) |
| q27\_livewith | 986 | Numeric | Respondents reporting being responsible for at least one child up to the age of 18 | 0-10 | Number of children (respondent is responsible for) that live in respondent’s household - (part of response to Q27) |
| q27\_liveelsewhere | 986 | Numeric | Respondents reporting being responsible for at least one child up to the age of 18 | 0-12 | Number of children (respondent is responsible for) that live elsewhere - (part of response to Q27) |
| q52\_baseline | 263 | Numeric | Ever-listeners of TWG among baseline survey respondents only | 0=No  1=Yes | Baseline version of Q52: Have you listened to ‘Together with Gloria’ in the last four weeks? |
| q52\_followup | 526 | Numeric | Ever listeners of TWG among follow-up survey respondents only | 0=No  1=Yes | Follow-up version of Q52: Together with Gloria has now finished airing, and we are interested in how many people were listening to the final few episodes.  Have you listened to any Together with Gloria episodes since Christmas? |