# SCAPES - Data Management Plan

## David Simons

2024-02-05

# Qualitative data

# Participatory Rural Assessment

How will data be collected?
What form will this data take?
How/where will it be stored?
What will the data include?
What are the variables/parameters of the data?
How will data be cleaned/processed?
How will the data be anonymised/participants protected? Will data be shared? If so, how?

# Quantitative data

# Community questionnaires

Questionnaires have been designed as XLSForms and are implemented using the Kobo platform. Data is collected on local devices and sent to the KoboToolbox server (https://kf.kobotoolbox.org/) for the SCAPES project. Access can be provided for data entry, modification and download. An unprocessed data set can be downloaded directly from the KoboToolbox server.

Alternatively, a project has been created to pull the data and format it using R.

# R Repository for data processing

The robotoolbox R package is used to download the data locally through the Kobo API. This approach uses the dm R package to handle the grouped repeats in the questionnaire and the haven R package to handle variable structure (e.g., labelled factors).

#### Household questionnaires:

Household questionnaire variables and responses Variable Name: start

**Description:** Formatted datetime when the questionnaire entry was created.

Type of Data: Datetime

Variable Name: end

**Description:** Formatted datetime when the questionnaire date entry was completed. This may not equate to the length of data collection as data may be added later (e.g., images).

Type of Data: Datetime

#### Variable Name: date

**Description:** Likely not needed but was used for internal questionnaire logic. Similar to end in that it records the date of the most recent data entry in the form.

Type of Data: Date
Format: %Y\_%m\_%d

Variable Name: interviewer\_id

**Description:** Name of the interviewer who conducted the interview.

Type of Data: Factor with labels associated with values.

#### Options for Responses:

- Option 1: Diana (Code: 1)
- Option 2: Helen (Code: 2)
- Option 3: Nzube (Code: 3)
- Option 4: Sunday (Code: 4)
- Option 5: Other (Code: 5)

Relationships with other variables: interviewer\_id\_other contains the free-text entry if other is selected.

Variable Name: community

**Description:** Community questionnaire was conducted in.

Type of Data: Factor with labels associated with values.

#### Options for Responses:

- Option 1: Dyegh (Akwa Kwasi) (Code: dyegh)
- Option 2: Ikyogbakpev (Akwa Kwasi) (Code: ikyogbakpev)
- Option 3: Zugu (Code: zugu)
- Option 4: Okimbongha (Code: okimbongha)
- Option 5: Ogamanna (Code: ogamanna)
- Option 6: Ofonekom (Code: ofonekom)
- Option 7: Ezeakataka (Code: ezeakataka)
- Option 8: Enyandulogu (Code: enyandulogu)
- Option 9: Offianka (Code: offianka)

Variable Name: consent\_household

**Description:** Whether consent for data collection has been provided.

Type of Data: Binary
Options for Responses:

Yes (Code: 1)No (Code: 0)

Variable Name: household\_number

**Description:** Identification number of the household this questionnaire pertains to.

Type of Data: Integer

Validation Rules: Combined with community to produce a unique identifier for a household. Data validation

during processing to ensure each household is uniquely identified.

Variable Name: multiple\_family\_household

**Description:** Are members from this household all from the same family? See protocol for definitions used

for household and family.

Type of Data: Integer

Variable Name: n\_people

 $\textbf{Description:} \ \ \text{The number of individuals currently living within the household. Includes those who come}$ 

and go on a regular basis or are current students away for school.

Type of Data: Integer

Repeat Section: Household members demographic will be repeated for each member of the household Variable Name: sex\_person

**Description:** The sex of the individual.

Type of Data: Binary
Options for Responses:

Female (Code: 1)Male (Code: 0)

Variable Name: baby

**Description:** Are they less than 1 year old?

Type of Data: Binary (Yes or No)

Options for Responses:

Yes (Code: 1)No (Code: 0)

Variable Name: age\_person

**Description:** The age of the individual, in years. Only asked for those >1 year old.

Type of Data: Integer

Variable Name: age\_baby

**Description:** The age of the individual in months. Only asked for those <1 year old.

Type of Data: Integer

Variable Name: permanent\_transient

Question: In the last year, was this individual regularly resident within this household?

Type of Data: Binary (Yes or No)

**Options for Responses:** 

Yes (Code: 1)No (Code: 0)

End of household member repeat

Variable Name: compound

Question: Is this household within a compound?

**Description:** A compound can be an enclosed area. A compound can also be a clear but defined area.

Type of Data: Binary (Yes or No)

Options for Responses:

• Yes (Code: 1)

• No (Code: 0)

Variable Name: multiple\_households

Question: Do any other households live in the same compound?

**Description:** Only asked if the answer to the compound question is Yes.

Type of Data: Binary (Yes or No)

Options for Responses:

Yes (Code: 1)No (Code: 0)

Variable Name: n\_in\_other\_households

Question: How many people live in this compound?

Description: Not including those that are in the focal household. Only asked if the answer to the

multiple\_households question is Yes.

Type of Data: Integer

Variable Name: other\_household\_activities

Question: Which of the following activities occur regularly with these other households?

Description: Only asked if the answer to the multiple\_households question is Yes.

Type of Data: Factor with labels associated with values.

### Options for Responses:

- Option 1: Sleeping (Code: sleeping)
- Option 2: Eating (Code: eating)
- Option 3: Cooking (Code: cooking)
- Option 4: Hunting (Code: hunting)
- Option 5: Farming (Code: farming)
- Option 6: None (Code: none)
- Option 7: Other (Code: other)

Variable Name: other\_household\_activities\_specify

**Description:** Captures other activities which may be conducted with the focal household only if other is selected in other\_household\_activities

Type of Data: Free text

Variable Name: household\_ethnicity\_same

Question: Do all members of this household identify as the same ethnicity?

Type of Data: Binary (Yes or No)

**Options for Responses:** 

• Yes (Code: 1)

• No (Code: 0)

Variable Name: household\_ethnicity

**Question:** What ethnicity is the household?

**Description:** Only asked if the answer to the household\_ethnicity\_same question is Yes.

Type of Data: Factor with labels associated with values.

### Options for Responses:

- Option 1: Igbo (Izzi) (Code: igbo\_izzi)
- Option 2: Igbo (Other) (Code: igbo\_other)
- Option 3: Membe (Code: membe)
- Option 4: Tiv (Code: tiv)
- Option 5: Other (Code: other)

 ${\bf Variable\ Name:\ other\_ethnicity\_household}$ 

**Description:** Only asked if the answer to household\_ethnicity is other

Type of Data: Free-text

Variable Name: household\_religion\_same

Question: Do all members of this household share the same religion?

Type of Data: Binary (Yes or No)

Options for Responses:

Yes (Code: 1)No (Code: 0)

Variable Name: household\_religion

Question: What religion is the household?

**Description:** Factor with labels associated with values.

Options for Responses:

• Option 1: Christian (Code: christian)

• Option 2: Muslim (Code: muslim)

• Option 3: Traditionalist (Code: traditionalist)

• Option 4: Other (Code: other)

Variable Name: household\_religion\_other

Description: Only asked if the answer to household\_religion is other

Type of Data: Free-text

Variable Name: n\_compound

Question: How many buildings or other structures are there in the compound?

**Description:** This variable represents the count of buildings or other structures within the compound.

Type of Data: Integer

Variable Name: n\_individual\_buildings

Question: How many individual buildings does the household use regularly?

Description: This variable represents the count of individual buildings regularly used by the household.

Type of Data: Integer

Variable Name: building owned

**Question:** Are these buildings owned by your household?

**Description:** This variable indicates whether the buildings are owned by the household.

Type of Data: Factor with labels associated with values.

**Options for Responses:** 

• Option 1: Own themselves (Code: own themselves)

• Option 2: Owned by other members of the family (Code: own family)

• Option 3: Rented from a non-family member (Code: rented)

• Option 4: Other (Code: other)

Variable Name: building\_owned\_other

Description: This question appears if other is selected in building\_owned

Type of Data: Free-text

Variable Name: n multi room

Question: How many multi-room buildings does this household regularly use?

**Description:** This variable represents the count of multi-room buildings regularly used by the household.

Type of Data: Integer

Variable Name: n\_single\_room

Question: How many single-room buildings does this household regularly use?

Type of Data: Integer

Repeat Section: Multiroom buildings will be added, one per repeat Variable Name: building\_purpose

Question: What does the household use this building for?

**Description:** For each building used by a household this variable captures information on what purpose it is used for.

Type of Data: Character with each option in a string separated by a space

#### Options for Responses:

- Option 1: Sleeping (Code: sleeping)
- Option 2: Food preparation (Code: food\_preparation)
- Option 3: Cooking (Code: cooking)
- Option 4: Eating (Code: eating)
- Option 5: Socialising/Parlour (Code: socialising\_parlour)
- Option 6: Cooked food storage (Code: food storage)
- Option 7: Packaged food storage (Noodle, Indomie etc.) (Code: packaged food storage)
- Option 8: Crop storage (Garri, Rice, Yam etc.) (Code: crop\_storage)
- Option 9: Seed storage (Code: seed\_storage)
- Option 10: Animal storage (Chicken, Goat etc.) (Code: animal\_storage)
- Option 11: Other storage (Code: other storage)
- Option 12: Other (Code: other)

Variable Name: specify\_building\_purpose

Description: Only asked if other is selected as one of the options in building\_purpose

Type of Data: Free-text

Variable Name: building\_location

Question: Where is the building located?

Type of Data: Factor with labels associated with values

## Options for Responses:

- Option 1: Within the compound (Code: compound)
- Option 2: No compound, within the village (Code: no compound)
- Option 3: Elsewhere in the village, outside of the compound (Code: village)
- Option 4: In the field, outside of the compound (Code: fields)
- Option 5: In a different village (Code: other\_village)
- Option 6: In a different town/city (Code: in\_town)
- Option 7: Other (Code: other)

Variable Name: specify\_building\_location

Description: Only asked if other is selected as the option in building\_location

Type of Data: Free-text

\_\_\_\_\_

Variable Name: roof\_material

**Question:** What is the roof made of?

Type of Data: Character with each option in a string separated by a space

#### Options for Responses:

- Option 1: Mat/thatch (Code: mat\_thatch)
- Option 2: Zinc/metal (Code: zinc\_metal)
- Option 3: Deck (Code: deck)
- Option 4: None (Code: none)
- Option 5: Other (Code: other)

#### **Definitions:**

- Roof refers to external material of the roof
- Deck . . .

Variable Name: roof\_other

**Description:** Only asked if other is selected as the option in roof material

Type of Data: Free-text

\_\_\_\_\_

Variable Name: wall\_material

**Question:** What are the walls made of?

**Type of Data:** Character with each option in a string separated by a space

#### **Options for Responses:**

- Option 1: Mud block (Code: mud\_block)
- Option 2: Brick (Code: brick)
- Option 3: Cement blocks (Code: cement\_block)
- Option 4: Plaster (Code: plaster)

- Option 5: Wood (Slats, Lumber etc.) (Code: wood)
- Option 6: Trees/Sticks (Code: wood\_other)
- Option 7: Other (Code: other)

#### **Definitions:**

- Mud block an unfired mud block
- Brick a fired block
- Plaster ...

Variable Name: walls\_other

Description: Only asked if other is selected as the option in wall\_material

Type of Data: Free-text

Variable Name: door\_material

**Question:** What is the door of the main entrance made of?

Type of Data: Select one Options for Responses:

• Option 1: Wood (Code: wood)

• Option 2: Metal (Code: metal)

• Option 3: None (Code: none)

• Option 4: Other (Code: other)

Variable Name: door\_other

Description: Only asked if other is selected as the option in door\_material

Type of Data: Free-text

Variable Name: window\_material

Question: What are the windows made of?

Type of Data: Character with each option in a string separated by a space

Options for Responses:

- Option 1: Permanently open (Code: permanently open)
- Option 2: Permanently closed (Code: permanently closed)
- Option 3: Wooden shutters (Code: wooden\_shutters)
- Option 4: Glass panes (Code: glass)
- Option 5: Mosquito net (Code: screen)
- Option 6: Metal grill/sheet (Code: metal)
- Option 7: No windows (Code: none)
- Option 8: Other (Code: other)

Variable Name: specify\_window

Description: Only asked if other is selected as the option in window\_material

Type of Data: Free-text

Variable Name: ceiling\_material

Question: What material is used for the ceiling?

Type of Data: Character with each option in a string separated by a space

#### **Options for Responses:**

- Option 1: Mat/thatch (Code: mat thatch)
- Option 2: Wood (Code: wood)
- Option 3: Cement (Code: cement)
- Option 4: Synthetic (Code: synthetic)
- Option 5: Sack bag/Carpet/Lino (Code: lining)
- Option 6: No ceiling (Code: no\_ceiling)
- Option 7: Other (Code: other)

#### **Definitions:**

• Sack bag/Carpet/Lino - . . .

Variable Name: specify\_ceiling

Description: Only asked if other is selected as the option in ceiling\_material

Type of Data: Free-text

Variable Name: ceiling\_storage

Question: Is there anything stored between the ceiling and the roof?

Type of Data: Binary (yes/no)

# Options for Responses:

Yes (Code: 1)No (Code: 0)

Variable Name: storage\_items\_ceiling

Question: What is stored in this space?

Type of Data: Character with each option in a string separated by a space

#### **Options for Responses:**

- Option 1: Cooked food (Code: cooked food)
- Option 2: Uncooked food (Code: uncooked food)
- Option 3: Packaged food (Code: packaged\_food)
- Option 4: Crops for sale (Code: crops\_for\_sale)
- Option 5: Crops for household use (Code: crops for household use)
- Option 6: Seeds for sale (Code: seeds for sale)
- Option 7: Seeds for household use (Code: seeds for household use)
- Option 8: Clothing (Code: clothing)
- Option 9: Other (Code: other)

Variable Name: ceiling\_storage\_other

Description: Only asked if other is selected as the option in storage\_items\_ceiling

Type of Data: Free-text

Variable Name: floor

**Question:** What is the floor made of?

Type of Data: Character with each option in a string separated by a space

**Options for Responses:** 

• Option 1: Mud (Code: mud)

• Option 2: Cement (Code: cement)

• Option 3: Tile (Code: tile)

• Option 4: Other (Code: other)

Variable Name: internal\_door

Question: Are there doors between the rooms in the building?

**Description:** This variable indicates whether there are doors between the rooms in the building.

Type of Data: Single-select with binary responses.

Options for Responses:

Yes (Code: 1)No (Code: 0)

Variable Name: internal\_door\_fit

Question: Do the doors fit well within the building?

**Description:** This variable indicates whether the doors fit well within the building. i.e., If your finger can fit through any space in the door, or between the door and the frame, it does not fit well.

Type of Data: Single-select with binary responses.

Options for Responses:

Yes (Code: 1)No (Code: 0)

\_\_\_\_

Variable Name: rooms\_in\_building

Question: How many rooms are there in this building?

**Description:** This variable captures the number of rooms within the same building, including kitchens, bedrooms, storerooms, parlors, etc.

Type of Data: Integer

\_\_\_\_

Variable Name: n\_household\_sleep

Question: How many people from this household sleep in this building?

Description: This variable indicates the number of people from this household who sleep in this building.

Type of Data: Integer

Variable Name: sleep\_same\_room

Question: Does everyone from this household that sleeps in this building sleep in the same room?

**Description:** This variable indicates whether everyone from this household who sleeps in this building sleeps in the same room.

Type of Data: Single-select with binary responses.

#### **Options for Responses:**

Yes (Code: 1)No (Code: 0)

Variable Name: n\_rooms\_household\_sleep

Question: How many different rooms do people from this household sleep in?

**Description:** This variable captures the number of different rooms in which people from this household

sleep.

Type of Data: Integer

Variable Name: other\_household\_sleep

Question: Do members of a different household sleep in this building?

**Description:** This variable indicates whether members of a different household sleep in this building.

Type of Data: Single-select with binary responses.

#### **Options for Responses:**

Yes (Code: 1)No (Code: 0)

Variable Name: other\_household\_sleep\_shared

Question: Do members of a different household sleep in the same space/place as members of this household?

**Description:** This variable indicates whether members of a different household sleep in the same space/place as members of this household.

Type of Data: Single-select with binary responses.

#### **Options for Responses:**

Yes (Code: 1)No (Code: 0)

Variable Name: n\_other\_household\_sleep\_multi

Question: How many members of a different household sleep in this building?

**Description:** This variable captures the number of members of a different household who sleep in this building.

Type of Data: Integer

Variable Name: rodents\_enter\_building

Question: Do any rats enter this building?

Description: This variable indicates whether any rats are known to enter this building.

Type of Data: Single-select with three responses

#### **Options for Responses:**

• Option 1: Yes (Code: yes)

- Option 2: No (Code: no)
- Option 3: Unknown (Code: unknown)

Variable Name: rodents\_in\_building

Question: Do any rats live in this building?

**Description:** This variable indicates whether any rats are known to live in this building.

Type of Data: Single-select with three responses

#### Options for Responses:

• Option 1: Yes (Code: yes)

• Option 2: No (Code: no)

• Option 3: Unknown (Code: unknown)

Variable Name: rodents\_in\_building\_evidence

Question: How do you know rats live here?

**Description:** This variable captures the evidence of rats living in the building.

Type of Data: Multi-select

# Options for Responses:

- Option 2: See dead rats (Code: see\_dead\_rats)
- Option 3: See rat urine (Code: see rat urine)
- Option 4: See rat faeces (Code: see rat faeces)
- Option 5: See rat burrows (Code: see\_rat\_burrows)
- Option 6: Hear them (Code: hear\_them)
- Option 7: Smell them (Code: smell them)
- Option 8: Direct contact with rats (Code: direct contact)
- Option 9: Seen the damage they have done to items (Code: item damage)
- Option 10: Other (Code: other)

Variable Name: rodents\_in\_building\_specify

Question: Please specify.

**Description:** This variable allows for specifying other evidence of rats living in the building. Only asked if other is selected in rodents\_in\_building\_evidence.

Type of Data: Free text

Variable Name: rodent faeces multi sleep

Question: Do you ever see the excreta/shit of rats in the places where people sleep?

**Description:** This variable indicates whether rat excreta are observed in sleeping areas.

Type of Data: Single-select with three responses (Yes, No, Unknown)

**Options for Responses:** 

• Option 1: Yes (Code: yes)

• Option 2: No (Code: no)

• Option 3: Unknown (Code: unknown)

Variable Name: rodent\_sleeping\_contact\_multi

Question: Has anyone been bitten or scratched by a rat while sleeping in this building?

**Description:** This variable indicates whether there has been any contact that could potentially result in direct pathogen transmission with rats while sleeping. Sleeping areas may not be the same thing as bedrooms as individuals may sleep in additional rooms of the building.

Type of Data: Single-select with three responses (Yes, No, Unknown)

**Options for Responses:** 

• Option 1: Yes (Code: yes)

• Option 2: No (Code: no)

• Option 3: Unknown (Code: unknown)

Variable Name: mastomys\_multi

**Question:** Do you notice a rat called *Mastomys natalensis* (the multimammate rat, or the rat with many offspring) in this building?

**Description:** This variable captures awareness of the presence of *Mastomys natalensis* in the building.

Type of Data: Single-select with three responses (Yes, No, Unknown)

**Options for Responses:** 

• Option 1: Yes (Code: yes)

• Option 2: No (Code: no)

• Option 3: Unknown (Code: unknown)

Variable Name: mastomys\_timing\_multi

**Question:** What time of the day do you see this rat?

**Description:** This variable captures the timing of *Mastomys natalensis* sightings.

Type of Data: Single-select Options for Responses:

- Option 1: Daytime (Code: daytime)
- Option 2: Nighttime (Code: nighttime)
- Option 3: Anytime (Code: anytime)

Variable Name: mastomys\_season\_multi

Question: Which season do you see them in?

**Description:** This variable captures the seasonality of *Mastomys natalensis* sightings.

Type of Data: Single-select

Options for Responses:

• Option 1: Dry (Code: dry)

• Option 2: Rainy (Code: rainy)

• Option 3: All seasons (Code: all\_seasons)

Variable Name: designated\_kitchen

Question: Is there a room that is designated as the kitchen in this building?

**Description:** This variable indicates whether there is a room designated as the kitchen in the building. Most kitchen activities occur outside but some preparation may occur within buildings or the kitchen may be attached to the building.

Type of Data: Single-select with three responses (Yes, No, Unknown)

**Options for Responses:** - Option 1: Yes (Code: yes) - Option 2: No (Code: no) - Option 3: Unknown (Code: unknown)

Variable Name: sleep\_in\_kitchen

Question: Do people from this household regularly sleep in the room that is designated as the kitchen?

**Description:** This variable captures whether individuals from the household regularly sleep in the room designated as the kitchen. Only asked if the answer to designated\_kitchen is Yes.

Type of Data: Single-select with three responses (Yes, No, Sometimes)

**Options for Responses:** - Option 1: Yes (Code: yes) - Option 2: No (Code: no) - Option 3: Sometimes (Code: sometimes)

Variable Name: n\_sleep\_in\_kitchen

Question: How many individuals from this household sleep in the kitchen?

**Description:** This variable indicates the number of individuals from the household who regularly sleep in the kitchen. Only asked if the answer to sleep\_in\_kitchen is Yes.

Type of Data: Integer

Variable Name: rodent\_damage\_kitchen

**Question:** Do rats eat or destroy food or ingredients in the kitchen?

**Description:** This variable captures whether rats cause damage to food or ingredients in the kitchen. It considers the damage they do to household items stored within kitchens.

Type of Data: Single-select with three responses (Yes, No, Unknown)

#### **Options for Responses:**

• Option 1: Yes (Code: yes)

• Option 2: No (Code: no)

• Option 3: Unknown (Code: unknown)

Variable Name: rodent\_damage\_kitchen\_items

Question: What do rats eat or destroy in the kitchen?

**Description:** This variable captures the items that rats eat or destroy in the kitchen.

Type of Data: Multi-select

#### Options for Responses:

• Option 1: Prepared but uncooked food (Garri, Pounded Yam etc.) (Code: uncooked\_food)

- Option 2: Other ingredients for cooking (Spices, Dried Fish etc.) (Code: ingredients)
- Option 3: Cooked food (Code: cooked\_food)
- Option 4: Rice for cooking (Code: rice)
- Option 5: Yams for cooking (Code: yam)
- Option 6: Groundnut for cooking (Code: groundnut)
- Option 7: Other crops for cooking (Code: other\_crops)

Variable Name: specify\_other\_kitchen\_items

Question: Please specify other items that rats eat or destroy in the kitchen.

**Description:** This variable allows for specifying other items that rats may eat or destroy in the kitchen. This question is only asked if other crops is selected in the rodent\_damage\_kitchen\_items.

Type of Data: Free-text

Options for Responses:

\_\_\_\_

Variable Name: prepared food damage

Question: How much of the prepared food kept in this kitchen is damaged?

**Description:** This variable indicates the extent of damage to prepared but uncooked food kept in the kitchen by rats.

Type of Data: Single-select with three responses.

Variable Name: cooked\_food\_damage

Question: How much of the cooked food kept in this kitchen is damaged?

**Description:** This variable indicates the extent of damage to cooked food kept in the kitchen by rats.

Type of Data: Single-select with three responses (A bit (<25%), A lot (25-75%), Most (>75%))

Variable Name: rice\_damage

Question: How much of the rice kept in this kitchen is damaged?

**Description:** This variable indicates the extent of damage to rice kept in the kitchen by rats.

Type of Data: Single-select with three responses (A bit (<25%), A lot (25-75%), Most (>75%))

Variable Name: yam damage

Question: How much of the yam kept in this kitchen is damaged?

**Description:** This variable indicates the extent of damage to yams kept in the kitchen by rats.

Type of Data: Single-select with three responses (A bit (<25%), A lot (25-75%), Most (>75%))

Variable Name: groundnut\_damage

Question: How much of the groundnut kept in this kitchen is damaged?

**Description:** This variable indicates the extent of damage to groundnuts kept in the kitchen by rats.

Type of Data: Single-select with three responses (A bit (<25%), A lot (25-75%), Most (>75%))

Variable Name: ingredients\_damage

Question: How much of the other ingredients kept in this kitchen is damaged?

**Description:** This variable indicates the extent of damage to other ingredients kept in the kitchen by rats.

Type of Data: Single-select with three responses (A bit (<25%), A lot (25-75%), Most (>75%))

## Individual questionnaires

Linking IDs to samples?

**Human serology** 

Rodent trapping

Rodent serology/PCR

Other