

Dataset Organization

The questionnaire for the 2018 is divided into 5 books:

Book 1: Household Member Characteristics

- Section A: Household Roster
- Section B: General Education
- Section C: Alternative Education and ICT
- Section D: General Health and Disability
- Section E: Child Preventative Health
- Section F: Women's Reproductive Health
- Section G: Health Knowledge, Behaviors and Attitudes
- Section H: Employment and Time Use - 7 Days
- Section I: Employment and Time Use - 12 Months
- Section J: Migration

Book 2: Household Characteristics

- Section K: Housing
- Section L: Durable Goods
- Section M: Social Assistance and Subjective Wellbeing
- Section N: Non-Food Consumption (Infrequently Purchased Items)
- Section O: Financial Services
- Section P: Non-Farm Enterprises

Book 3: Agriculture

- Section R: Agricultural Assets
- Section S: Annual Crops
- Section T: Permanent Crops
- Section U: Forestry Activities
- Section V: Fishing
- Section W: Livestock

Book 4: Consumption

- Section X: Food Consumption
- Section Y: Non-Food Consumption (Frequently Purchased Items)

[Book 4A: secs X and Y for the first 10 days and sec Z; Book 4B: X and Y for days 11 to 20]

- Section Z: Bulk Purchases

The different sections / questions are mapped into the different datasets as shown on the next page.

All datasets contain:

- ID variables (`_cluster`, `_hhno` and a third (and fourth) id if needed such as `_ind` or `_line`).
- variables for survey data analysis: `_cluster` (psu), `_stratum` and `_pweight`
- serial numbers of the questionnaire books from which they take data.
- In general, variables whose names start with an underscore are administrative variables.
- Other variable name correspond directly to the section and question number on the questionnaire: b2 is section B, question 2. Multipart questions (first, second and third reasons etc) and usually postfixed a, b, c etc (for example, d4a, d4b, d4c). When the "other" option was answered for a question, the text given to specify is the variable with the postfix `_other` (for example, d4a_other).

Datasets can be merged using the variable `_cluster` and `_hhno` (or just `_cluster` if merging with `slhs2018_cluster.dta`)

Book	Section	Questions	Dataset
cover page and administrative variables			slihs2018_cluster.dta
consumption aggregates and poverty lines			slihs2018_consexp.dta
1	all		slihs2018_ind.dta
2	K	1-41, 48-53	slihs2018_hh.dta
2	K	42-47	slihs2018_k42.dta
2	L	all	slihs2018_l.dta
2	M	2-8	slihs2018_m2.dta
2	M	9-20	slihs2018_hh.dta
2	M	21-23	slihs2018_m21.dta
2	M	24-26	slihs2018_m24.dta
2	N	all	slihs2018_n.dta
2	O	1, 2, 11, 15, 23, 30, 40	slihs2018_hh.dta
2	O	3-10, 12-14, 16-22, 24-29, 31-39	slihs2018_o.dta *
2	O	41-48	slihs2018_o41.dta
2	O	49-54	slihs2018_o49.dta
2	P	all	slihs2018_p.dta
2	Q	all but Q1	slihs2018_q.dta
3	R	1-7	slihs2018_r1.dta
3	R	9-24	slihs2018_r9.dta
3	S	1-42	slihs2018_s.dta
3	S	43-44	slihs2018_hh.dta
3	T	1-39	slihs2018_t.dta
3	T	40-41	slihs2018_hh.dta
3	U	all	slihs2018_u.dta
3	V	1-8, 29	slihs2018_hh.dta
3	V	9-17, 18-22, 23-26	slihs2018_v9.dta *
3	V	30-40	slihs2018_v30.dta
3	W	all but W0	slihs2018_w.dta
2-4	misc	P0, Q1, R8, V1, V0, W0, Z0	slihs2018_hh.dta
4A/B	X	all	slihs2018_x.dta
4A/B	Y	all	slihs2018_y.dta
4B	Z	all but Z0	slihs2018_z.dta

* Note that datasets **slihs2018_o.dta** and **slihs2018_v9.dta** have data from multiple parallel tables.