JLMPS 2025 Read Me
Version 2.0
August 24, 2026

Data files

All data files are available as Stata .dta files.

2025 cross section: JLMPS 2025 xs v2.0 all.dta

- indid: unique individual identifier
- hhid: unique household identifier
- For those who cannot load the many variables in this data set, you can separately load pt 1, pt 2 and pt 3 versions.

2010-2025 panel: JLMPS 2025 panel v2.0.dta

- Findid: unique individual identifier from when individual was first observed
- The data are constructed as panel data (wide format) directly from the repeated cross section.
- This file contains only the created variables across waves.
- Variables repeat as 'variable'_# where # is wave, i.e. pn_16, pn_25 for the person numbers in 2016 and 2025. If an individual is not present in a wave, the variables for that wave will be missing.

2010/2016/2025 repeated cross section: JLMPS 2025 rep xs v2.0.dta

- round: Identifies wave of survey (2010, 2016, 2025)
- indid: unique individual identifier
- hhid: unique household identifier
- This file contains only the created, compatible variables across time.
- The data are long structure, observation is an individual & year combination

Do files

All current data processing runs through the "Master Do File JLMPS 2025.do" This do file calls sub-do files (all in do files folder)

Weights

- expan_hh: this is the cross-sectional *household* expansion weight for the full sample and can be used with the rep xs and xs files. This is the main weight you should use in cross-sectional analyses on the *household* level (using variables from household questionnaire, household level analyses). This variable is available in the panel, but should not be used for panel analyses; see panel weight below.
- expan_indiv: this is the cross-sectional *individual* expansion weight for the full sample and can be used with the rep xs and xs files. This is the main weight you should use in cross-sectional analyses on the *individual* level (using variables from individual questionnaire or individual level analyses). This variable is

available in the panel, but should not be used for panel analyses; see panel weight below.

- expan_ref_hh: this is the weight, specific to 2025, that is a cross-sectional expansion weight based on only the refresher sample. This is a weight you can use in cross-sectional analyses on the *household* level (using variables from household questionnaire, household level analyses).
- expan_ref_indiv: this is the weight, specific to 2025, that is a cross-sectional expansion weight based on only the refresher sample. This is a weight you can use in cross-sectional analyses on the *individual* level (using variables from individual questionnaire or individual level analyses).
- panel_wt_*: this is the panel weight, designed for analyses spanning different combinations of years (_*). It is not an expansion weight, but normalized to 1.

Enforcement of skip patterns

Skip patterns were only enforced uni-directionally; the programming ensured that individuals answered every question they were supposed to answer. However, if there was a misunderstanding/data entry error such that they went backwards and revised to follow a different skip path, stray/entered data remained even if no longer relevant. Analyze according to the universe who should have answered per skip patterns.

Questionnaires

The questionnaire is divided into two parts:

1. Household questionnaire
2. Individual questionnaire

English and Arabic questionnaires for each of these parts are available.

Codebooks

Created variables are documented in: Codebook JLMPS 2025 data v2.0.xlsx

For created variables only. Gives variable name, type, label, ranges and states, universe, years covered, and notes.

There are additional Arabic codebooks for:

- Country codes
- Geographic codes (gov.xls and stat_u.xls for mapping locations to urban/rural)
- Primary school codes
- Preparatory school codes
- Secondary school codes
- University codes
- Higher education specialization codes (specialization.xlsx)