

Instructions on how to use LIRNEasia AfterAccess datasets.

This document provides some general instructions on using the LIRNEasia's AfterAccess datasets.

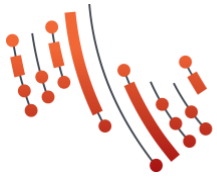
1. Representativity of data

- 1.1. Please note that this dataset is representative of respective populations at national level with a 95% confidence interval with margins of errors mentioned in table 1. The samples were not designed to be representative of any other segmentation (e.g., state-level, district-level, etc.), geographic or otherwise. For more information on the survey and sampling method, please see: <https://lirneasia.net/wp-content/uploads/2023/11/AfterAccess-ICT-access-and-use-in-Sri-Lanka-India-Pakistan-BangladeshCambodia-and-Nepal-Methodology-note-2017-18.pdf>

Table 1. Target population, sample sizes and margin of errors of the surveys

Country	Target population	Sample size (actual achieved)	Margin of Error at 95% Confidence Level (National Estimates)
Sri Lanka	<ul style="list-style-type: none"> All households Age 15-65 population 	2,017	±3.3%
India		5,069	±2.0%
Pakistan		2,002	±3.3%
Bangladesh		2,020	±3.3%
Nepal		2,008	±3.3%
Cambodia		2,123	±3.3%
Sri Lanka	<ul style="list-style-type: none"> Persons with disabilities aged 15-65 	400	±9.8%
Nepal		400	±9.8%
Sri Lanka	<ul style="list-style-type: none"> All small and medium sized enterprises 	402	±9.8%

2. Interpretation of data along with the questionnaire



2.1. The data should be interpreted in conjunction with a careful consideration of the survey questionnaire, for example to understand the full meanings/extent of the questions, but also to understand who might have been included vs excluded from answering a particular question (based on filter questions which may not be adjacent in ordering to the particular question of interest). The questionnaire contains information at the start of each section of which categories of respondents are to answer the following sections.

2.2. For household & individual surveys and persons with disability surveys:

Please also note that Part A of the questionnaire was answered by a household representative (e.g., the household head), while Part B was answered by a randomly selected individual from among those within the household who met the eligibility criteria. The responses for both of these Parts are reported in a single case (row) in the dataset.

For Enterprise survey:

There will be only one part which is answered by the enterprise owner or the enterprise manager.

3. Data weights

3.1. **For the household and individual surveys:** Separate weights have been calculated and must be applied for the household responses (Part A) versus the individual responses coming from the same household (Part B). The dataset contains two weighting variables as such (which are clearly marked in the dataset) and must be applied in order for representation of the data (either at household or individual level, respectively) to be ensured. **Raw numbers should not be reported or used at any time.**

3.1.1. Two sets of weights are utilized in this dataset to derive projections (and to use for other descriptive statistics and statistical modeling.

3.1.1.1. **Statistical projections:** The users shall use **HH_Weight_For_Projections** and **IND_Weight_For_Projections** variables to weight the datasets when deriving the projections (i.e. **estimated number of households, individuals or persons with disabilities**) from the household section and individual sections respectively.

3.1.1.2. **Other descriptive statistics, tests and modeling:** The users shall use **HH_Weight_For_Analysis** and **IND_Weight_For_Analysis** variables to weight the dataset when deriving the other descriptive statistics (i.e. Contingency tables, charts), statistical testing, and running the statistical models using the household and individual sections separately.

3.2. **For Enterprise survey dataset:** Two sets of weights have been calculated and must be applied for the entire dataset depending on the requirement. **Raw numbers should not be reported or used at any time.**

3.2.1. Statistical projections: The users shall use **Enterprise_Weight_For_Projections** variables to weight the datasets when deriving the projections (**i.e. estimated number of enterprises**) from the household section and individual sections respectively.

3.2.2. **Other descriptive statistics, tests and modeling:** The users shall use **Enterprise_Weight_For_Analysis** variables to weight the dataset when deriving the other descriptive statistics (i.e. Contingency tables, charts), statistical testing, and running the statistical models using the household and individual sections separately.

4. Bases

4.1. Some sections and sub-sections may have very low bases in certain countries (e.g., the platform use section), therefore caution should be exercised in interpretation.

5. Missing values

5.1. Missing values are coded using a range of different formats in the dataset (e.g., 99, 99999, 888888, -1, etc.). In the current SPSS format, all relevant codes have been designated as 'missing values.' However, if the dataset is to be converted into a different format, then missing values may not be captured in the same way when performing analyses by other software.

6. Income and expenditure data

6.1. The survey captured some data on income and expenditure. These are contained in the dataset, some in local currency, some converted to USD for analysis purposes. The exchange rate used is based on what was prevailing at the time of fieldwork.

6.2. However, more importantly, please consider these data for what they are: self-reported approximations of the average monthly expenditures and incomes at the time. For accurate income and expenditure data to be captured, a dedicated survey would be required, but this was not the aim of/a large priority for this survey.