

**STEP Survey Weighting Procedures Summary
(Based on The World Bank Weight Requirement)**

Kenya

March 28, 2014

Table of Contents

1	Survey Design Overview.....	1
1.1	Target Population.....	1
1.2	Sample Frame	1
1.3	Sample Design.....	2
2	Data Collection Outcomes	3
3	Weighting Overview.....	5
4	Files - Weighting Process	6
4.1	Input Files.....	6
4.1.1	Revisions to Input Data File.....	6
4.1.1.1	Input File – Revised for Weighting	8
4.2	Final Output Files from Weighting Process.....	10
5	Weight Calculations based on World Bank (WB) Weight Requirement	13
5.1	First-stage Weight (PSU Weight).....	13
5.1.1	First-stage Response Adjustment	14
5.1.1.1	Response Adjustment Factor for PSU Non-participation	15
5.1.2	Final PSU Weight.....	15
5.2	Second-stage Weight (Household Weight).....	16
5.2.1	Basic Household Weight	16
5.2.2	Second-stage Response-adjusted Household Weight	17
5.2.2.1	Second-stage Household Response Categories	17
5.2.2.2	Eligible Proportion of Households	20
5.2.2.3	Response Adjustment Factor for Household Non-participation	21
5.2.2.4	Response-adjusted Household Weight.....	22
5.2.3	Second-stage Final Household Weight	22
5.2.4	Weighted Household Counts	23
5.3	Third-stage Weight (Person Weight)	24
5.3.1	Basic Third-stage Person Weight	24
5.3.2	Third-stage Response-adjusted Person Weight.....	25
5.3.2.1	Third-stage Person-level Response Categories.....	25
5.3.2.2	Response Adjustment Factor for Selected Person Non-participation	28
5.3.2.3	Response-adjusted Person Weight	29
5.3.3	Third-stage Final Person Weight.....	30

5.3.4	Weighted Person Counts	31
6	Response Rate.....	32

March 28, 2014

1 Survey Design Overview

The Kenya survey firm implemented a Full Literacy Assessment design. The Full assessment required each selected person to attempt to complete a General Booklet comprising Reading Components and a set of Core Literacy Items. Subsequently, those respondents that correctly answered at least 3 Core Literacy items were assigned one of four assessment Exercise Booklets to complete. The full assessment sampling objective was to have a minimum of about 2400 completed Exercise Booklets, with a distribution of about 600 complete cases for each of the four Exercise Booklets.

1.1 Target Population

The target population for the Kenya STEP survey comprises all non-institutionalized persons 15 to 64 years of age (inclusive) living in private dwellings in urban areas of the country at the time of data collection. This includes all residents except foreign diplomats and non-nationals working for international organizations.

Exclusions

- Itinerants (as classified in the Population Census 2009 in Kenya);
- War marred and unstable regions of KENYA
- Residents of senior homes and hospices;
- Residents of other group dwellings such as college dormitories, halfway homes, workers' quarters, etc.;
- Persons living outside the country at the time of data collection.

1.2 Sample Frame

The sample frame for the selection of first stage sample units is the list of core urban Enumeration Areas (EAs, i.e., PSUs) from the Population Census of 2009.

Table 1 provides the sample frame counts by stratum of the number of households in the population.

Table 1: Kenya Sample Frame, # of Urban Households by Stratum

Stratum	# of Urban HHs
1-Nairobi	984,825
2-Other Large cities (over 100,000 HHs)	723,075
3- Medium cities (60,000 to 100,000 HHs)	372,248
4-Other Urban Areas	612,230
Total	2,692,378

1.3 Sample Design

The Kenya sample design is a stratified 3 stage sample design. The sample was stratified by 4 geographic areas: 1-Nairobi, 2-Other Large Cities (over 100,000 households), 3- Medium cities (60,000 to 100,000 HHs), and 4-Other Urban Areas.

First Stage Sample

The primary sample unit (PSU) is a census Enumeration Area; each PSU is uniquely defined a Sequential PSU Number 'SRN'.

The first stage units were selected by the Kenya National Bureau of Statistics. The sampling objective was to select 268 PSUs.

The PSUs were selected using a probability proportional to size (PPS) sampling method, where the measure of size was the number of households in a PSU.

Second Stage Sample

The second stage sample unit (SSU) is a household. The sampling objective was to obtain interviews at 15 households within each selected PSU. The households were selected in each PSU using a systematic random method.

Table 2 provides the overall sample allocation of PSUs and households.

Table 2: Kenya Overall Sample Allocation by Stratum				
Stratum	# of PSUs	Initial HH Sample	Reserve HH Sample	Total
1-Nairobi	67	1005	1005	2010
2-Other Large cities (over 100,000 HHs)	67	1005	1005	2010
3- Medium cities (60,000 to 100,000 HHs)	67	1005	1005	2010
4-Other Urban Areas	67	1005	1005	2010
Total	268	4020	4020	8040

Third Stage Sample

The third stage sample unit was an individual aged 15-64 (inclusive); one individual was selected with equal probability from each participating household.

2 Data Collection Outcomes

The final survey weights depend on the outcome from the attempt to obtain an interview at each selected household. The treatment of each sampled case in the weighting process depends on the result code assigned to the sampled household and the result code assigned to the selected person. The result codes used in STEP are listed in Table 3. A variable called 'final_status' with these code values was created during the weighting process and a 'final_status' was assigned to each sampled case.

Table 3: <i>Full Literacy Assessment</i> Final Status Codes for Sampled Households and Selected Persons	
I. Household-level Codes	
Code	Code Description
01	Refusal-Household
02	Household unusual circumstance (e.g., death in family, illness, fire in dwelling, etc.)
03	No knowledgeable household member found after 3 visits (only a child, non-competent adult, etc.)
04	Temporarily absent/unavailable for field period (information from other persons)
05	No competent household member to interview (because of severe illness, mental disability, etc.)
06	Language problem – no household member spoke a language understandable by survey team, no translator available
07	Dwelling could not be found / selected address has no household
08	Vacant dwelling / selected address has no household
09	Dwelling not habitable/dwelling destroyed/dwelling converted to commercial use
21	Household Refusal to continue Household Module (i.e., begun but not complete)
22	Household unusual circumstance (e.g., death in family, illness, fire in dwelling, etc.) prevented continuance
31	No household member in the eligible range of 15-64
94	Case dropped due to data quality issue
95	Sample case justifiably not activated, e.g., reserve sample household not needed , sample goal achieved
96	General non-response; Reason unspecified by survey firm
97	Household not interviewed due to sample selection error – wrong household selected in its place
II. Person-level Codes	
Code	Code Description
	Individual modules 2-7 <u>not</u> begun or unacceptable interview
32	Refusal by Selected Person to begin modules 2-7
33	Unusual circumstance - Selected Person (e.g., death in family, illness, fire in dwelling, etc.)
34	Refusal - Another household member refused to let selected individual do individual modules
35	Unable to contact Selected Person after three visits during field period
36	Temporarily absence/unavailability of Selected Person for field period (info from household member)
37	Hearing Problem - Selected Person is deaf or so hard-of-hearing, no translator available
38	Learning or mental disability (includes emotional conditions like severe depression) of Selected Person
39	Speech impairment of Selected Person prevented participation, no translator available
40	Language problem – Selected Person's language not understood by survey team, no translator available
98	Sample selection error – an incorrect eligible person was selected
99	Sample selection error – an ineligible person was selected
	Individual modules 2-7 begun but not completed
51	Refusal by Selected Person to continue modules 2-7
52	Unusual circumstance - Selected Person
	General Booklet (module 9) not begun
61	Unable to read or write in language of General Booklet – Selected Person refused to begin
62	Refusal by Selected Person to begin General Booklet
63	Unusual circumstance - Selected Person
64	Blindness or visual impairment of Selected Person prevented completion of General Booklet
65	Physical disability of Selected Person prevented completion of General Booklet

Table 3: <i>Full Literacy Assessment</i> Final Status Codes for Sampled Households and Selected Persons	
	General booklet (module 9) looked through or attempted but not completed
71	Selected Person looked at Booklet but unable to read & write the language of Booklet - did not begin
72	Refusal by Selected Person to continue the General Booklet
73	Interruption – unacceptably long, General Booklet administration terminated
74	Unusual circumstance - Selected Person
75	Attempted General Booklet but no answers marked
76	Marked some items but did not attempt all items in General Booklet
	Exercise Booklet (module 9)
81	No attempt at any of the exercise booklet items / refusal to start
82	Interruption too long during exercise booklet administration
83	Unusual circumstance, Exercise Booklet administration terminated
84	Attempt at some Exercise Booklet items but no answers marked
85	Attempt at all Exercise Booklet items but no answers marked
86	Marked some Exercise Booklet items, not all items attempted
87	Marked some Exercise Booklet items, all items attempted
88	Marked all Exercise Booklet items

3 Weighting Overview

While the Kenya three-stage stratified cluster design greatly enhanced the operational feasibility of data collection, it resulted in differential probabilities of selection for the selected persons. Consequently, each selected person in the survey does not necessarily represent the same number of persons in the target population. To account for differential probabilities of selection due to the nature of the design and to ensure accurate survey estimates, STEP requires a sampling weight for each person that participated in the survey.

In general, the objectives of the STEP weighting are to construct a set of survey weights to,

- 1) compensate for unequal probabilities of selection;
- 2) compensate for household-level non-response and person-level non-response.

The general weighting procedure for the Kenya STEP survey required the following tasks.

- 1) Creation of a data file to input into the weighting process;

2) **World Bank (WB) Weight Requirement:**

Create survey weights for sampled cases of households and persons that provided sufficient data to be considered a participant in the survey. This requirement does not necessarily include the completion of an assessment General Booklet, nor does it necessarily include the completion of all household and individual questionnaire modules.

- a) Calculation of a PSU weight for 268 sampled PSUs;
 - b) Calculation of a household weight for each sampled household;
 - i) Calculation of a household-level non-response adjustment independently for each PSU.
 - c) Calculation of a person weight for each selected person (SP);
 - i) Calculation of a non-response adjustment independently for each sampled person.
- 3) The required output from the weighting process is a final Kenya data file with the survey design weights (i.e., for each sampled PSU, household, person) appended to each data record.

4 Files - Weighting Process

4.1 Input Files

The key file for the weighting process is the file '[STEP Kenya - Data for Weighting - March 25 2014_S7.dta](#)' provided by the WB core team. The file '[STEP Kenya - Data for Weighting - March 25 2014_S7.dta](#)' contains a data record for each of the 8,040 sampled households from the Kenya sample frame. An SPSS version of this file, i.e., '[STEP Kenya - Data for Weighting - March 25 2014_S7.sav](#)', was created for the weighting process.

4.1.1 Revisions to Input Data File

The following general revisions were made to the Kenya SPSS data file '[STEP Kenya - Data for Weighting - March 25 2014_S7](#)':

- 1) '[m11_q07](#)', the Result of Interview Code (N.B. the survey firm assigned codes 21 to 88 only)
 - a) added value labels; the value labels for '[m11_q07](#)' are as follows:

Table 4: VALUE LABELS for the variable ' m11_q07 '	
00-No assigned code	71-SP looked through GB-unable to read & write language of GB'
21-Household Refusal to continue Household Module (i.e., begun but not complete)'	72-SP refused to continue GB after starting'
22-Household module terminated due to unusual circumstance'	73-GB interruption too long'
31-No eligible household member aged 15-64'	74-GB terminated due to unusual circumstance'
32-Refusal-Refusal by SP to begin modules 2-7'	75-Attempted GB but no answers marked'
33-Interview termination-SP unusual circumstance'	76-Attempted part or all of GB; less than 3 correct Core answers'
34-Refusal-Other Household Member refused to let SP participate'	77-Marked some items & attempted all items in GB'
35-Unable to contact SP after 3 attempts'	78-Marked all items in GB'
36-SP absent for entire survey period'	81-No attempt at any exercise booklet items / refusal to start'
37-SP deaf or hard of hearing'	82-Interruption too long during exercise booklet administration'
38-SP learning or mental disability'	83-Unusual circumstance, Exercise Booklet administration terminated'
39-SP speech impairment'	84-Attempt at some Exercise Booklet items, no answers marked'
40-SP language problem'	85-Attempt at all Exercise Booklet items, no answers marked'
51-Refusal by SP to continue modules 2-7'	86-Marked some Exercise Booklet items, not all items attempted'
52-Interview termination-SP unusual circumstance'	87-Marked some Exercise Booklet items, all items attempted'
61-SP unable to read language of General Booklet (GB)'	88-Marked all Exercise Booklet items'
62-SP refusal to begin GB'	
63-SP unable to do GB due to unusual circumstance'	
64-SP blindness or visual impairment-unable to do GB'	
65-SP physical disability'	

2) **final_status**: Revised the final status code for sampled cases as follows:

- i) Added value labels; Refer to Table 3 in section 2 for the description of **final_status** codes.
- ii) The data file variables '*final_status*' and '*field_rst* - Fieldwork result code ' were used to revise the variable '*final_status*'.
 - ♦ For , '*field_rst*' the description of the codes is as follows:

Code	Code Description
1	Activated; for an 'initial sample' HH this means that it was contacted, for a 'reserve sample' HH it means that it was activated for an interview attempt.
2	Never activated; applicable to reserve sample HHs, the case was correctly never activated, i.e., the prescribed fieldwork procedures were followed.
3	Wrongly activated; applicable to reserve sample cases, the reserve case was activated, but according to the prescribed fieldwork procedures it should <u>not</u> have been activated (i.e., a reserve sample HH was incorrectly activated).
4	Wrongly never activated; applicable to reserve sample cases, the reserve case was never activated but according to the prescribed fieldwork procedures it should have been activated (i.e., a reserve HH should have been activated based on the provided activation order, but it wasn't).
5	Incorrect eligible HH member selected; during the random selection process, an incorrect household member was selected from the list of <u>eligible</u> members.
6	Ineligible HH member selected; during the random selection process, an <u>ineligible</u> household member was selected.

- iii) The specifications for creating the variable '**final_status**' are as follows;

- | |
|---|
| <ul style="list-style-type: none">▪ IF (final_status ~= 0 & (field_rst = 1 field_rst = 3)) final_status=final_status;▪ IF (field_rst = 2) final_status=95;▪ IF (field_rst = 4) final_status=97;▪ IF (field_rst = 5) final_status=98;▪ IF (field_rst = 6) final_status=99. |
|---|

4.1.1.1 Input File – Revised for Weighting

The SPSS Revised Input File for weighting is ‘[W00-KEN Final Revised Data File\(25-Mar-2014\)](#)’. This file is the primary file for the process of creating the survey weights; it includes all 8040 sampled cases; each case includes a final status code. Table 5 provides the list of key variables for weighting purposes that are included in the file.

Table 5: Key Variables in Primary Input File for Weighting Process		
Variable Name	Source	Description
Country	Sample File	Country abbreviation code
SmpID	Sample File	Sampled case ID#
EACODE	Final Data File	EA Code
psu	Sample File	Sampled PSU ID#
HHsmpOrder	Sample File	Sample activation order
HHsmpSeq	Sample File	Household Sample Sequence #s 01 to 30
HHsmpType	Sample File	HH Sample Indicator
Stratum	Sample Frame	Stratum ID#
Households_A	Sample Frame	Households - # in Sample Frame
numberofhhlisted	Final Data File	Number of listed households in PSU
m1c_a_clean	Final Data File	NUMBER A-# of eligible persons
m11_q07	Final Data File	Result code of interview
field_rst	Final Data File	Fieldwork result code
final_status	Final Data File	Final Result Code
St_PsUsSmpd	Created during weighting	Number of sampled PSUs in Stratum
St_FR_HHs	Created during weighting	Number of Target Population HHs in Stratum
Prim_PSU	Created during weighting	Indicator for Primary PSU Data Record

Table 6 provides the breakdown of the 8040 sample cases by the final result of the interview attempt, i.e., by '*final_status*' code.

Table 6: Kenya: Number & Percentage of Sampled Dwellings by Final Result Code of Interview '<i>final_status</i>'		
Final_Status	# of Sample Cases	Percent of Sample Cases
01-Refusal-Household	96	1.2%
02-Unusual circumstance-Household	18	0.2%
03-No knowledgeable household member found after 3 visits	62	0.8%
04-Temporarily absent/unavailable for field period	116	1.4%
05-No competent household member (e.g., severe illness, mental disability, etc.)	35	0.4%
21-Household Refusal to continue Household Module (begun,not complete)	4	0.0%
35-Unable to contact SP after 3 attempts	1	0.0%
51-Refusal by SP to continue modules 2-7	8	0.1%
61-SP unable to read language of General Booklet (GB)	159	2.0%
62-SP refusal to begin GB	143	1.8%
63-SP unable to do GB due to unusual circumstance	28	0.3%
64-SP blindness or visual impairment-unable to do GB	10	0.1%
65-SP physical disability	1	0.0%
71-SP looked through GB-unable to read/write language of GB	61	0.8%
72-SP refused to continue GB after starting	221	2.7%
73-GB interruption too long	29	0.4%
74-GB terminated - unusual circumstance	32	0.4%
75-Attempted GB, no answers marked	72	0.9%
76-Attempted part or all of GB; <3 correct Core answers	507	6.3%
81-No attempt at any Exercise Booklet items / refusal to start	184	2.3%
82-Interruption too long during Exercise Booklet administration	27	0.3%
83-Unusual circumstance, Exercise Booklet administration terminated	16	0.2%
84-Attempt at some Exercise Booklet items, no answers marked	139	1.7%
85-Attempt at all Exercise Booklet items, no answers marked	118	1.5%
86-Marked some Exercise Booklet items, not all items attempted	671	8.3%
87-Marked some Exercise Booklet items, all items attempted	780	9.7%
88-Marked all Exercise Booklet items	688	8.6%
95-Household not activated	3798	47.2%
98-Incorrect eligible household member selected	16	0.2%
Total	8040	100.0%

4.2 Final Output Files from Weighting Process

The final output file from the weighting process is,

- 1) **KEN-W01-STEP(HH) Weighted Data-Final Design Weights (25-Mar-2014)**
 - a) This file contains the original data variables from the Kenya data file '**STEP Kenya - Data for Weighting - March 25 2014_S7**' and the applicable PSU weights, household weights, person weights, and other associated variables that were created during the weighting process.
 - b) The file contains a data record for each of the 4242 cases that were activated or should have been activated during the data collection operation.
 - i) 3911 participating households with a final non-zero household weight
 - ♦ For participating households, the weighted sum of the household weights is 2,849,353.
 - ii) 3894 participating selected persons with a final non-zero person weight
 - ♦ For participating individuals, the weighted sum of the person weights is 6,381,235; this is the estimated number of persons aged 15 to 64 in the Kenyan population.
 - c) In addition to the Kenya data file original variables, the weighted data file includes several additional variables that were created during the weighting process. The essential variables that are needed for estimation of survey results are included in the final weighted data file.
 - i) Table 7 provides a list of additional variables from the weighting process that are included in the final weighted data file.
 - ii) Table 8 provides a list of variables from the weighting process that are not included in the final weighted data file. These variables are available from other interim files that were created during the weighting process.
 - iii) Table 9 provides the frequency count of final status codes for the 4242 sampled households that were activated or should have been activated in the final weighted data file '**KEN-W01-STEP(HH) Weighted Data-Final Design Weights (25-Mar-2014)**'.

Table 7: Variables Created During Weighting Process - Added to Kenya Data File

Variable Name	Description	Values
St_PSU_Smpd	Number of sampled PSUs in Stratum	n/a
St_FR_HHs	Number of Target Population HHs in Stratum	n/a
RespCat_PSU	Final Response Category-PSUs	1-Participant PSU 2-Ineligible PSU 3-Non-participant Eligible PSU
RespCat_HH	Final Response Category-Households	1-Participant HH 2-Ineligible HH 3-Non-participant Eligible HH 4-Non-participant Unknown Eligibility HH
RespCat_SP	Final Response Category-Selected Person	1-Participant SP 2-Non-participant SP 3-Not applicable - non-participant HH
W1_PSU	Basic PSU Weight	n/a
W2_FinHHwt	Final Household Weight Use for estimation of Household characteristics	n/a
W3_FinSPwt	Final Person Weight	n/a

Table 8: Variables Created During Weighting Process – Not Added to Kenya Data File

Variable Name	Description	Values
Prim_PSU	Indicator for Primary PSU Record	0-Duplicate PSU Record 1-Primary PSU Record
STpsu_SmpSz	Stratum-PSU Sample Size - HHs	n/a
PSU_RespFreq	Count of 'RespCat_HH' in Stratum-PSU	n/a
W0_PSU	Initial PSU Weight	n/a
PSU_k	# of participant PSUs in Stratum	n/a
PSU_x	# of Ineligible PSUs in Stratum	n/a
PSU_z	# of Non-participant Eligible PSUs in Stratum	n/a
R_adj_PSU	Response adjustment ratio for PSUs	n/a
W0_HH	Initial Household Weight within Stratum-PSU	n/a
W2_HHwt_basic	Basic Household Weight	n/a
HH_k	# of participant HHs in PSU	n/a
HH_x	# of Ineligible HHs in PSU	n/a
HH_z	# of Non-participant Eligible HHs in PSU	n/a
HH_u	# of Non-participant Unknown Eligible HHs in PSU	n/a
Elig_Prop_HH	Proportion of Eligible HHs - known eligibility	n/a
R_adj_HH	Response adjustment ratio for HHs	n/a
W0_SP	Initial Selected Person weight within household	n/a
R_adj_SP	Response adjustment for selected person	n/a
EP_PSU_RCat	Total Eligible Persons in PSU Participant Households	n/a
EP_k	PSU - # of Participant SPs in participant HHs	n/a
EP_z	PSU - # of Non-participant SPs in participant HHs	n/a
EP_PSU	PSU - Total Eligible Sampled Persons	n/a

Table 9: Kenya, Frequency/Percent of Final Status Codes For Activated Sample Cases

final_status	# of Cases	% of Total Cases
01-Refusal-Household	96	2.3%
02-Unusual circumstance-Household	18	0.4%
03-No knowledgeable household member found after 3 visits	62	1.5%
04-Temporarily absent/unavailable for field period	116	2.7%
05-No competent household member (e.g., severe illness, mental disability, etc.)	35	0.8%
21-Household Refusal to continue Household Module (begun, not complete)	4	0.1%
35-Unable to contact SP after 3 attempts	1	0.0%
51-Refusal by SP to continue modules 2-7	8	0.2%
61-SP unable to read language of General Booklet (GB)	159	3.7%
62-SP refusal to begin GB	143	3.4%
63-SP unable to do GB due to unusual circumstance	28	0.7%
64-SP blindness or visual impairment-unable to do GB	10	0.2%
65-SP physical disability	1	0.0%
71-SP looked through GB-unable to read/write language of GB	61	1.4%
72-SP refused to continue GB after starting	221	5.2%
73-GB interruption too long	29	0.7%
74-GB terminated - unusual circumstance	32	0.8%
75-Attempted GB, no answers marked	72	1.7%
76-Attempted part or all of GB; <3 correct Core answers	507	12.0%
81-No attempt at any Exercise Booklet items / refusal to start	184	4.3%
82-Interruption too long during Exercise Booklet administration	27	0.6%
83-Unusual circumstance, Exercise Booklet administration terminated	16	0.4%
84-Attempt at some Exercise Booklet items, no answers marked	139	3.3%
85-Attempt at all Exercise Booklet items, no answers marked	118	2.8%
86-Marked some Exercise Booklet items, not all items attempted	671	15.8%
87-Marked some Exercise Booklet items, all items attempted	780	18.4%
88-Marked all Exercise Booklet items	688	16.2%
98-Incorrect eligible household member selected	16	0.4%
Total	4242	100.0%

5 Weight Calculations based on World Bank (WB) Weight Requirement

The input file for the process of calculating the weights based on the World Bank (WB) weight requirement is '[W00-KEN Final Revised Data File\(25-Mar-2014\)](#)'.

5.1 First-stage Weight (PSU Weight)

The basic first stage weight, i.e., the PSU weight, for the Kenya STEP Survey is based on the probability of selection for each of the 268 sampled PSUs.

For the Kenya sample of PSUs, the probability of selection of the j^{th} PSU, say p_j , may be represented as,

$$1) \quad p_{gj} = \frac{y_g * m_{gj}}{M_g}$$

where,

y_g - the number of sampled PSUs in stratum g ,

m_{gj} - the measure of size for the j^{th} PSU in stratum ' g ', in the case of the Kenya STEP Survey the measure of size for the first stage selection of PSUs was the number of households in a PSU,

M_g - the total number of household in stratum ' g ', based on the sample frame information, is represented by

$$1) \quad M_g = \sum_1^{y_g} m_{gj}$$

Table 10 provides the Kenya stratum counts for y_g and M_g .

Table 10: Kenya Overall PSU Sample Allocation by Stratum			
Stratum		Sampled PSUs	# of HHs in Target Pop'n
Area	g	y_g	M_g
1-Nairobi	1	67	984,825
2-Other Large cities (over 100,000 HHs)	2	67	723,075
3- Medium cities (60,000 to 100,000 HHs)	3	67	372,248
4-Other Urban Areas	4	67	612,230
Total		268	2,692,378

The basic first stage weight for the j^{th} PSU, $W1_{gj}$, is calculated as the inverse of its probability of selection:

$$2) \quad W1_{gj} = \frac{1}{p_{gj}} = \frac{M_g}{y_g * m_{gj}}$$

Example 1: First Stage Weight

The basic first stage weight for Kenya PSU # 1 in stratum 1 – '**Nairobi**', with 117 Sample Frame households, is calculated as follows:

$$W1_{11} = \frac{1}{p_{11}} = \frac{M_1}{y_1 * m_{11}} = \frac{984,825}{67 * 117} \approx 125.6315$$

The basic first stage weight is similarly calculated for each of the 268 PSUs.

5.1.1 First-stage Response Adjustment

The first-stage response adjustment was carried out for the PSUs within a Stratum.

For treatment in the weighting process the data collection outcomes for the sampled PSUs may be summarized in three possible categories: participant, ineligible, non-participant with known eligibility.

N.B. There were no non-participating PSUs with unknown eligibility in the Kenya STEP Survey.

Participant - A participating PSU is an eligible sampled PSU that provides sufficient data to be considered a participant in the survey. The sampled PSUs with one or more person-level result codes 32-40, 51-88 or 98-99 would satisfy the requirement for classification of a PSU as a survey participant.

Ineligible – A sampled PSU is categorized as ineligible if there are no eligible household members in the PSU. The PSU # 13 was demolished and thus had no eligible household members in the PSU.

Non-participant - Eligibility known: In the Kenya Step, a non-participating PSU with known eligibility resulted when a sampled PSU was not accessed due to ‘Security issues’. The PSUs 24, 227, 228, 229, 230 and 231 were inaccessible for interviewing due to security concerns; it is assumed that each of these PSUs included some eligible households.

The total number of sampled PSUs may be represented by the following equation.

$$3) \quad y_{gj} = y_{gk} + y_{gx} + y_{gz}$$

where,

y_{gj} - the number of sampled PSUs in Stratum ‘g’, (i.e.,),

y_{gk} - the number of participant PSUs ‘k’ in Stratum ‘g’,

y_{gx} - the number of ineligible PSUs ‘x’ in Stratum ‘g’,

y_{gz} - the number of non-participating PSUs ‘z’ with known eligibility in Stratum ‘g’.

Note: $\sum_{g=1}^{26} y_{gj} = 268$ PSUs in the Kenya STEP Survey.

Table 11 provides the STEP count of sampled PSUs from the Kenya data file by PSU Response Category.

Table 11: Kenya, Total Number of Stratum PSUs by PSU Response Category	
PSU Response Category	# of PSUs
1-Participant PSU ($\sum y_{gk}$)	261
2-Ineligible PSU ($\sum y_{gx}$)	1
3-Non-participant PSU; Known Eligibility ($\sum y_{gz}$)	6
Total ($\sum y_{gj}$)	268

5.1.1.1 Response Adjustment Factor for PSU Non-participation

A PSU response adjustment is necessary to compensate for PSUs that did not participate in the survey. The PSU response adjustment factor for the j^{th} PSU in Stratum 'g', $R_{PSU_{gj}}$, is calculated as follows:

$$4) \quad R_{PSU_{gj}} = \begin{cases} \frac{\text{Total eligible PSUs in Stratum 'g'}}{\text{Total participating PSUs in Stratum 'g'}} = \frac{y_{gk} + y_{gz}}{y_{gk}}, & \text{Participant PSU} \\ 0, & \text{Non-participant PSU} \end{cases}$$

- The PSUs that are categorized as ineligible, y_{gx} , are excluded from the calculation of the PSU response adjustment.

Example 2: Response-adjustment Factor

For Kenya PSU # 1 in stratum 1 – 'Nairobi',

$$R_{PSU_{gj}} = \frac{y_{gk} + y_{gz}}{y_{gk}} = \frac{65 + 1}{65} \\ \approx 1.01538$$

5.1.2 Final PSU Weight

The final PSU weight, i.e., the basic response-adjusted PSU weight, is calculated as the product of the basic PSU weight and the PSU response adjustment factor. Therefore, for PSU 'j' in Stratum 'g', the final PSU weight, $WPSU_{gj}$, is calculated as follows:

$$5) \quad WPSU_{gj} = W1_{gj} * R_{PSU_{gj}}$$

- The final PSU weight is appended to the STEP data file.

Example 3: Final PSU Weight, i.e., Basic Response-adjusted PSU Weight

For Kenya PSU # 1 in stratum 1 – 'Nairobi',

$$WPSU_{gj} = W1_{gj} * R_{PSU_{gj}} \\ \approx 125.6315 \times 1.01538 \\ \approx 127.5642$$

The final PSU weight is similarly calculated for all sampled PSUs.

5.2 Second-stage Weight (Household Weight)

5.2.1 Basic Household Weight

The basic second stage household weight is calculated for each sampled household in the participating PSUs.

Table 12 provides the number of sampled households for participating PSUs by Stratum.

Table 12: Kenya Participant PSUs & Sampled Households by Stratum		
Stratum	Participant PSUs	Sampled HHs
1-Nairobi	65	1950
2-Other Large cities (over 100,000 HHs)	67	2010
3- Medium cities (60,000 to 100,000 HHs)	67	2010
4-Other Urban Areas	62	1860
Total	261	7830

The basic second stage weight for the k^{th} household in PSU ' j ', $W2_{jk}$, is calculated as follows:

$$6) \quad W2_{jk} = \frac{m'_j}{n_j}$$

where,

m'_j - the number of listed households (i.e., weighted data file variable 'numberofhhlisted') in the j^{th} PSU,

n_j - the number of households sampled in the j^{th} PSU.

Example 4: Second Stage Basic Weight

The basic second stage weight for each of the 30 sampled households in Kenya PSU # 1 which contains 125 listed households is calculated as follows:

$$W2_{jk} = \frac{m'_j}{n_j} = \frac{125}{30} \approx 4.1667$$

The basic second stage weight is similarly calculated for each of the 7830 sampled households in participating PSUs.

5.2.2 Second-stage Response-adjusted Household Weight

5.2.2.1 Second-stage Household Response Categories

The second-stage response adjustment was carried out for each sampled PSU, i.e., at the PSU level.

For treatment in the weighting process the data collection outcomes for the sampled households may be summarized in four possible categories: participant, ineligible, non-participant with known eligibility, non-participant with unknown eligibility.

Participant - A participating household 'k' is an eligible sampled household that provides sufficient data to be considered a participant in the survey. The sampled households 'k' with person-level result codes 32-40, 51-88 and 98-99 satisfy the requirement for classification of the household as a survey participant. The result codes and corresponding code descriptions for participating households 'k' are as follows:

I. Household-level Codes – Participant, Eligible Household	
Code	Code Description
Individual modules 2-7 not begun	
32	Refusal by Selected Person to begin modules 2-7
33	Unusual circumstance - Selected Person (e.g., death in family, illness, fire in dwelling, etc.)
34	Refusal - Another household member refused to let selected individual do individual modules
35	Unable to contact Selected Person after three visits during field period
36	Temporarily absence/unavailability of Selected Person for field period (info from household member)
37	Hearing Problem - Selected Person is deaf or so hard-of-hearing, no translator available
38	Learning or mental disability (includes emotional conditions like severe depression) of Selected Person
39	Speech impairment of Selected Person prevented participation, no translator available
40	Language problem – Selected Person's language not understood by survey team, no translator available
Individual modules 2-7 begun but not completed	
51	Refusal by Selected Person to continue modules 2-7
52	Unusual circumstance - Selected Person
General Booklet (module 9) not begun	
61	Unable to read or write in language of General Booklet – Selected Person refused to begin
62	Refusal by Selected Person to begin General Booklet
63	Unusual circumstance - Selected Person
64	Blindness or visual impairment of Selected Person prevented completion of General Booklet
65	Physical disability of Selected Person prevented completion of General Booklet
General booklet (module 9) looked through or attempted but not completed	
71	Selected Person looked at Booklet but unable to read & write the language of Booklet - did not begin
72	Refusal by Selected Person to continue the General Booklet
73	Interruption – unacceptably long
74	Unusual circumstance - Selected Person
75	Attempted General Booklet but no answers marked
76	Marked some items but did not attempt all items in General Booklet
77	Marked some items & attempted all items in General Booklet
78	Marked all items in General Booklet
81	No attempt at any of the exercise booklet items / refusal to start
82	Interruption too long during exercise booklet administration
83	Unusual circumstance, Exercise Booklet administration terminated
84	Attempt at some Exercise Booklet items but no answers marked
85	Attempt at all Exercise Booklet items but no answers marked
86	Marked some Exercise Booklet items, not all items attempted
87	Marked some Exercise Booklet items, all items attempted
88	Marked all Exercise Booklet items

I. Household-level Codes – Participant, Eligible Household

Code	Code Description
98	Sample selection error – an incorrect eligible person was selected by the interviewer
99	Sample selection error – an ineligible person was selected by the interviewer

Ineligible – A sampled household ‘x’ is categorized as ineligible if there is no household member in the eligible 15-64 age range. The result codes for ineligible sampled households ‘x’ are as follows:

II. Household-level Codes – Ineligible Household

Code	Code Description
07	Dwelling could not be found / selected address has no household
08	Vacant dwelling
09	Dwelling not habitable/dwelling destroyed/dwelling converted to commercial use
31	No household member in the eligible range of 15-64

N.B. There were no ineligible households in the Kenya STEP.

Non-participant - Eligibility Known: A non-participating household ‘z’ results when an eligible household does not participate due to refusal, unusual circumstance, or the household was selected in error. The result codes for sampled households ‘z’ with known eligibility are as follows:

III. Household-level Codes – Non-participants, Eligible Household

Code	Code Description
21	Household Refusal to continue Household Module (i.e., begun but not complete)
22	Household unusual circumstance (e.g., death in family, illness, fire in dwelling, etc.) prevented continuance
97	Sample selection error – a wrong household was selected in the field

Non-participant - Eligibility Unknown: A non-participating household ‘u’ results when a household does not participate due to non-contact with a responsible household member, inconclusive information about the eligibility status of the sampled household, or sample case was not activated (this usually pertains to the reserve sample cases). The result codes for non-participating households ‘u’ with unknown eligibility are as follows:

IV. Household-level Codes – Non-participants, Unknown Eligibility Household

Code	Code Description
01	Household Refusal
02	Household unusual circumstance (e.g., death in family, illness, fire in dwelling, etc.)
03	No knowledgeable household member found after 3 visits (only a child, non-competent adult, etc.)
04	Temporarily absent/unavailable for field period (information from other persons)
05	No competent household member to interview (because of severe illness, mental disability, etc.)
06	Language problem – no household member spoke a language understandable by survey team, no translator available
95	Sample case not activated
96	General non-response; Reason unspecified by survey firm

The total number of sampled households in each PSU may be represented by the following equation.

$$7) \quad n_j = n_{jk} + n_{jx} + n_{jz} + n_{ju}$$

where,

n_j - the number of households sampled in the j^{th} PSU,

n_{jk} - the number of households 'k' that participate in the j^{th} PSU,

n_{jx} - the number of ineligible households 'x' in the j^{th} PSU,

n_{jz} - the number of non-participating households 'z' with known eligibility in the j^{th} PSU,

n_{ju} - the number of non-participating households 'u' with unknown eligibility in the j^{th} PSU.

Table 13 provides the STEP counts of sampled households in participating PSUs from the Kenya data file by the outcome categories described above.

Table 13: Kenya, Count of Sampled Households in Participating PSUs by Final Status Sample & Final Response Category

final_status	Final Response Category-Households			# of Cases	% of Total Cases
	Participant HH	Non-participant HH: Eligible	Non-participant HH: Unknown Eligibility		
01-Refusal-Household			96	96	1.2%
02-Unusual circumstance-Household			18	18	0.2%
03-No knowledgeable household member found			62	62	0.8%
04-Temporarily absent			116	116	1.5%
05-No competent household member			35	35	0.4%
21-Household Refusal to continue HH Module		4		4	0.1%
35-Unable to contact SP after 3 attempts	1			1	0.0%
51-Refusal by SP to continue modules 2-7	8			8	0.1%
61-SP unable to read language of General Booklet (GB)	159			159	2.0%
62-SP refusal to begin GB	143			143	1.8%
63-SP unable to do GB - unusual circumstance	28			28	0.4%
64-SP blindness or visual impairment	10			10	0.1%
65-SP physical disability	1			1	0.0%
71-SP unable to read/write language of GB	61			61	0.8%
72-SP refused to continue GB after starting	221			221	2.8%
73-GB interruption too long	29			29	0.4%
74-GB terminated - unusual circumstance	32			32	0.4%
75-Attempted GB, no answers marked	72			72	0.9%
76- less than 3 correct Core answers	507			507	6.5%
81-No attempt at any EB items/refusal to start	184			184	2.3%
82-Interruption too long during EB administration	27			27	0.3%
83-Unusual circumstance, EB terminated	16			16	0.2%
84-Attempt at some EB items, no answers marked	139			139	1.8%
85-Attempt at all EB items, no answers marked	118			118	1.5%
86-Marked some EB items, not all items attempted	671			671	8.6%
87-Marked some EB items, all items attempted	780			780	10.0%
88-Marked all EB items	688			688	8.8%
95-Household not activated			3588	3588	45.8%
98-Incorrect eligible HH member selected	16			16	0.2%
Total	3911	4	3915	7830	100.0%

5.2.2.2 Eligible Proportion of Households

The STEP assumes that the households 'u' with unknown eligibility in the j^{th} PSU are comprised of some proportion, say ε_j , that are eligible households and the complementary proportion, $1 - \varepsilon_j$, that are ineligible households.

Amongst the n_{ju} non-participating households 'u' with unknown eligibility in PSU 'j', the proportion of households with known eligibility is estimated to be

$$8) \quad \varepsilon_j = \frac{\text{Total \# of known eligible cases}}{\text{Total \# of known eligible \& known ineligible cases}} = \frac{n_{jk} + n_{jz}}{n_{jk} + n_{jx} + n_{jz}}$$

➤ **Note: Since there are no ineligible households in the Kenya STEP, the eligible proportion is the same, and equals 1, for all eligible participating households.**

Example 5: Eligible Proportion of Households

For PSU #253, the following household results were obtained:

n_j	n_{jk}	n_{jx}	n_{jz}	n_{ju}
30	14	0	1	15

For PSU # 253, the proportion of households with known eligibility is calculated as,

$$\begin{aligned} \varepsilon_{101} &= \frac{\text{Total \# of known eligible cases}}{\text{Total \# of known eligible \& known ineligible cases}} \\ &= \frac{n_{jk} + n_{jz}}{n_{jk} + n_{jx} + n_{jz}} \\ &= \frac{14 + 1}{14 + 0 + 1} = \frac{15}{15} = 1.0 \end{aligned}$$

5.2.2.3 Response Adjustment Factor for Household Non-participation

A household-level response adjustment is necessary to compensate for households that did not participate in the survey. The household-level response adjustment factor for the j^{th} PSU, R_{HHj} , is calculated as follows:

$$9) \quad R_{HHj} = \frac{\text{Total eligible households in PSU } 'j'}{\text{Total participating households in PSU } 'j'} = \frac{n_{jk} + n_{jz} + (\epsilon_j * n_{ju})}{n_{jk}}$$

- The response adjustment ratio, R_{HHj} , incorporates the adjustment for unknown eligibility, ϵ_j .
- The households that are categorized as ineligible, n_{jx} , are excluded from the calculation of the household-level response adjustment.

Example 6: Second Stage Response Adjustment Factor for sampled Households

For example, the following results were obtained for PSU #1

n_j	n_{jk}	n_{jx}	n_{jz}	n_{ju}
30	15	0	0	15

Therefore,

$$\begin{aligned}
 R_{102} &= \frac{\text{Total eligible households in PSU \#253}}{\text{Total participating households in PSU \#253}} \\
 &= \frac{n_{jk} + n_{jz} + (\epsilon_{253} * n_{ju})}{n_{jk}} \\
 &= \frac{15 + 0 + (1.0 * 15)}{15} \approx 2.0
 \end{aligned}$$

The response adjustment ratio, R_{HHj} , is similarly calculated for all 261 participant PSUs.

5.2.2.4 Response-adjusted Household Weight

The second stage response-adjusted household weight is calculated as the product of the basic household weight and the response adjustment factor for non-participant households. Therefore, for household 'k' in PSU 'j', the final second stage household weight, $W2_{Radj}$, is calculated as follows:

$$10) \quad W2_{Radj} = W2_{jk} * R_{HH_j}$$

- The final second stage household weight is appended to the STEP data file.

Example 7: Second Stage Response-adjusted Household Weight

For PSU #1,

$$\begin{aligned} W2_{Radj} &= W2_{jk} * R_{HH_j} \\ &= 4.1667 \times 2.0 \\ &\approx 8.3333 \end{aligned}$$

The response-adjusted household weight is similarly calculated for all sampled households.

5.2.3 Second-stage Final Household Weight

The final household weight is the weight used for producing household-level estimates, such as the proportion of households with a specific attribute. This weight is appended to the STEP data file.

The final second stage household weight is calculated as product of the response-adjusted household weight and the final first-stage PSU weight. Therefore, for household 'k' in PSU 'j', the final second stage household weight, $W2_{HH_k}$, is calculated as follows:

$$11) \quad W2_{HH_k} = W2_{Radj} * WPSU_{gj}$$

- The final second stage household weight is appended to the STEP data file.

Example 8: Final Second Stage Household Weight

For PSU #1,

$$\begin{aligned} W2_{HH_k} &= W2_{Radj} * WPSU_{gj} \\ &= 8.3333 \times 127.5642 \\ &\approx 1063.0354 \end{aligned}$$

The final second stage household weight is similarly calculated for all sampled households.

5.2.4 Weighted Household Counts

The STEP data collections found that some of the sampled PSUs are ineligible, or were not accessible for security reasons.

Table 14 provides,

- i. the weighted count of households based on the listed households in the 268 sampled PSUs;
- ii. the weighted count of households based on the final status of the cases in the Kenya data file.

Table 14: Kenya – Weighted Number of Households – Sample Frame vs. STEP Data File		
Source of PSU Household Count	# of PSUs	Weighted Sum of Households
1. STEP Sample Frame [all sampled PSUs , i.e., participants, non-participants, ineligible], based on # of Sample Frame PSU HHs	268	2,692,378
2. STEP Data File ' STEP Kenya - Data for Weighting - March 25 2014_S7 ' [eligible PSUs only], based on # of PSU <u>Listed</u> HHs	261	2,849,353

As shown in the second row of Table 14, when ineligible PSUs are excluded, the weighted number of households based on the number of listed households, is **2,849,353**.

For participating selected households, Table 15 provides a summary of the central tendency for the participating households.

Table 15: Summary - Household Weights	
Statistic	Final Household Weight
No. of Cases	3911
Mean	728.55
Std. Error of Mean	5.18
Median	726.89
Minimum	332.08
Maximum	3165.91
Sum of Weights	2,849,353

5.3 Third-stage Weight (Person Weight)

5.3.1 Basic Third-stage Person Weight

The basic third stage person weight is calculated for each sampled person in the participating households; one person aged 15 to 64 was selected in each participating household.

Table 16 provides the number of sampled persons for participating households by Stratum.

Table 16: Kenya Sampled Persons by Stratum	
Stratum	Sampled Persons
1-Nairobi	975
2-Other Large cities (over 100,000 HHs)	1005
3- Medium cities (60,000 to 100,000 HHs)	1005
4-Other Urban Areas	926
Total	3911

The basic third stage weight for the i^{th} person, $W3_{jki}$, is calculated as follows:

$$12) \quad W3_{jki} = \frac{a_{jk}}{b_{jk}} = \frac{a_{jk}}{1} = a_{jk}$$

where,

a_{jk} - the number of eligible persons in the k^{th} household in the j^{th} PSU,

b_{jk} - the number of persons sampled in the k^{th} household in the j^{th} PSU, (i.e., $n_{jk}=1$ person in each Kenya sampled household).

N.B. $A_{jk} = \sum_{k=1}^{n_{jk}} a_{jk} = \text{total \# of selected persons in the } j^{\text{th}} \text{ PSU}$
 where,

n_{jk} - the number of households 'k' that participate in the j^{th} PSU

Example 9: Third Stage – Basic Person Weight within sampled Household

For PSU #1, selected household # 05 has 3 eligible persons. Therefore,

$$W3_{jki} = \frac{a_{jk}}{b_{jk}} = \frac{3}{1} = 3$$

5.3.2 Third-stage Response-adjusted Person Weight

5.3.2.1 Third-stage Person-level Response Categories

For eligible participating households, there are three possible person-level outcomes, a selected eligible person participates in the survey, a selected eligible person does not participate, a selected person is ineligible.

Participant SP - A participating selected person 'i' is an eligible sampled person that provides sufficient data to be considered a participant in the survey. The sampled persons 'i' with result codes 51-88 satisfy the requirement for classification of the person as a survey participant. The result codes for participating persons 'i' are as follows:

I. Person-level Codes – Participant in Eligible Selected Household	
Code	Code Description
Individual modules 2-7 begun but not completed	
51	Refusal by Selected Person to continue modules 2-7
52	Unusual circumstance - Selected Person
General Booklet (module 9) not begun	
61	Unable to read or write in language of General Booklet – Selected Person refused to begin
62	Refusal by Selected Person to begin General Booklet
63	Unusual circumstance - Selected Person
64	Blindness or visual impairment of Selected Person prevented completion of General Booklet
65	Physical disability of Selected Person prevented completion of General Booklet
General booklet (module 9) looked through or attempted but not completed	
71	Selected Person looked at Booklet but unable to read & write the language of Booklet - did not begin
72	Refusal by Selected Person to continue the General Booklet
73	Interruption – unacceptably long
74	Unusual circumstance - Selected Person
75	Attempted General Booklet but no answers marked
76	Marked some items but did not attempt all items in General Booklet
77	Marked some items & attempted all items in General Booklet
78	Marked all items in General Booklet
81	No attempt at any of the exercise booklet items / refusal to start
82	Interruption too long during exercise booklet administration
83	Unusual circumstance, Exercise Booklet administration terminated
84	Attempt at some Exercise Booklet items but no answers marked
85	Attempt at all Exercise Booklet items but no answers marked
86	Marked some Exercise Booklet items, not all items attempted
87	Marked some Exercise Booklet items, all items attempted
88	Marked all Exercise Booklet items

Non-participant SP - A non-participating person ‘q’ results when an eligible person does not participate due to refusal, unusual circumstance, non-contact, language problem, physical disability (hearing, vision, other) or the person was selected in error. The result codes for non-participating persons ‘q’ are as follows:

II. Person-level Codes – Non-participant in Eligible Selected Household	
Code	Code Description
Individual modules 2-7 <u>not</u> begun	
32	Refusal by Selected Person to begin modules 2-7
33	Unusual circumstance - Selected Person (e.g., death in family, illness, fire in dwelling, etc.)
34	Refusal - Another household member refused to let selected individual do individual modules
35	Unable to contact Selected Person after three visits during field period
36	Temporarily absence/unavailability of Selected Person for field period (info from household member)
37	Hearing Problem - Selected Person is deaf or so hard-of-hearing, no translator available
38	Learning or mental disability (includes emotional conditions like severe depression) of Selected Person
39	Speech impairment of Selected Person prevented participation, no translator available
40	Language problem – Selected Person’s language not understood by survey team, no translator available
98	Sample selection error – an incorrect eligible person was selected by the interviewer

Ineligible SP - A sampled person ‘f’ is categorized as ineligible if the person is not a member of the target population, for example, the person’s age is outside the eligible 15-64 age range. Ineligible sampled persons ‘f’ are as assigned the result code 99.

III. Person-level Codes – Non-participant in Eligible Selected Household	
Code	Code Description
99	Sample selection error – an ineligible person was selected by the interviewer

Note: there were no ineligible selected persons in the Kenya STEP Survey.

The third stage person weight is only relevant for sampled households in the final response category RespCat_HH=1, i.e., participating households. Table 17 provides the pertinent frequency counts of the Final Status Codes for the participant households.

Table 17: Kenya, Final Status Code for ‘Participating’ Households (i.e., Households with RespCat_HH=1)

Final Status	# of Cases	% of Total Cases
35-Unable to contact SP after 3 attempts	1	0.0%
51-Refusal by SP to continue modules 2-7	8	0.2%
61-SP unable to read language of General Booklet (GB)	159	4.1%
62-SP refusal to begin GB	143	3.7%
63-SP unable to do GB due to unusual circumstance	28	0.7%
64-SP blindness or visual impairment-unable to do GB	10	0.3%
65-SP physical disability	1	0.0%
71-SP looked through GB-unable to read/write language of GB	61	1.6%
72-SP refused to continue GB after starting	221	5.7%
73-GB interruption too long	29	0.7%
74-GB terminated - unusual circumstance	32	0.8%
75-Attempted GB, no answers marked	72	1.8%
76-Attempted part or all of GB; <3 correct Core answers	507	13.0%
81-No attempt at any Exercise Booklet items / refusal to start	184	4.7%
82-Interruption too long during Exercise Booklet administration	27	0.7%
83-Unusual circumstance, Exercise Booklet administration terminated	16	0.4%
84-Attempt at some Exercise Booklet items, no answers marked	139	3.6%
85-Attempt at all Exercise Booklet items, no answers marked	118	3.0%
86-Marked some Exercise Booklet items, not all items attempted	671	17.2%
87-Marked some Exercise Booklet items, all items attempted	780	19.9%
88-Marked all Exercise Booklet items	688	17.6%
98-Incorrect eligible household member selected	16	0.4%
Total	3911	100.0%

5.3.2.2 Response Adjustment Factor for Selected Person Non-participation

In each participating household in the STEP Survey, one eligible person was selected to be interviewed. Therefore, for each participating household the basic person weight, W_{3jki} , would also be the final person weight within a sampled PSU if all persons in the PSU participated. However, when there is one or more non-participating selected persons within a sampled PSU, the basic person-weights for the PSU require a response adjustment so that the final person weight for each participating person in the PSU is also representative of the non-participating person(s) in the sampled households in the PSU. In other words, the total number of participating persons in a sampled PSU must represent the total number of eligible persons in the sampled households in the PSU.

Therefore, within each sampled PSU, a person-level response adjustment to the basic person weight is necessary to compensate for persons that did not participate in the survey.

The total number of eligible persons in each PSU may be represented by the following equation:

$$13) \quad A_{jk} = A_{jki} + A_{jkq}$$

where,

$$\begin{aligned} A_{jk} &= \sum_{k=1}^{n_{jk}} a_{jk} && \text{the total number of eligible persons for all } n_{jk} \text{ participating households in the } j^{\text{th}} \text{ PSU,} \\ A_{jki} &= \sum_{k=1}^{n_{jk}} a_{jki} && \text{the total number of eligible persons 'i' that participate for all } n_{jk} \text{ participating households in the } j^{\text{th}} \text{ PSU,} \\ A_{jkq} &= \sum_{k=1}^{n_{jk}} a_{jkq} && \text{the total number of non-participating eligible persons 'q' for all } n_{jk} \text{ participating households in the } j^{\text{th}} \text{ PSU,} \end{aligned}$$

Therefore, the person-level response adjustment for the j^{th} PSU, R_{SP_j} , is calculated as follows:

$$14) \quad R_{SP_j} = \frac{\text{Total eligible persons in all participating sampled households in PSU 'j'}}{\text{Total participating persons in all participating sampled households in PSU 'j'}} = \frac{A_{jk}}{A_{jki}}$$

Example 10: Third Stage – Basic Person Weight Response Adjustment

- i) For PSU #1, there are 35 eligible persons across all participating households; also, all selected persons participated. Therefore,

$$A_{jk} = \sum_{k=1}^{n_{jk}} a_{jk} = 35 ; \quad A_{jki} = \sum_{k=1}^{n_{jk}} a_{jki} = 35$$

$$R_{SP_j} = \frac{A_{jk}}{A_{jki}} = \frac{35}{35} = 1.00$$

- ii) For PSU # 6, the total number of eligible persons across all participating households is 25. For PSU # 6, there are 3 non-participating selected persons. Therefore,

$$A_{jk} = \sum_{k=1}^{n_{jk}} a_{jk} = 25 ; \quad A_{jki} = \sum_{k=1}^{n_{jk}} a_{jki} = 22$$

$$R_{SP_j} = \frac{A_{jk}}{A_{jki}} = \frac{25}{22} \approx 1.1364$$

Table 18 provides the sample counts of the Final Response Category for selected persons.

Table 18: Kenya, Sample Counts by Selected Person Final Response Category for Participating HHs		
For Participating Households		
Final Response Category-SP	Frequency	Percent
1-Participant SP	3894	99.6
2-Non-participant, Eligible	17	0.4
Total	3911	100.0

The above table shows that there are 3894 participating selected persons for which a final non-zero Person-weight was calculated. The sampled persons in category 2 have a zero-value final person weight.

5.3.2.3 Response-adjusted Person Weight

The response-adjusted person weight is calculated as the product of the basic third stage person weight, and the response adjustment for the basic person weight. Therefore, for the selected person 'i' in household 'k' in PSU 'j', the response-adjusted person weight, W_{Radj} , is calculated as follows:

$$15) \quad W3_{Radj} = W3_{jki} * R_{SP_j}$$

- The final third stage person weight is appended to the STEP data file.

Example 11: Third Stage – Response-adjusted Person Weight

- i) For PSU #1, selected household # 05 with 3 eligible persons, the selected person was a participant. All selected persons participated in PSU #1, therefore, $R_{SP_j} = 1$

$$W3_{jki} = 3 * 1 = 3$$

$$\begin{aligned} W3_{Radj} &= W3_{jki} * R_{SP_j} \\ &= 3 * 1 = 3 \end{aligned}$$

- ii) For PSU #6, in participant household # 14 with 2 eligible persons, the selected person was a participant. For this household,

$$W3_{jki} = 2 ; \quad R_{SP_j} \approx 1.13636$$

$$\begin{aligned} W3_{Radj} &= W3_{jki} * R_{SP_j} \\ &= 2 * 1.13636 \approx 2.2727 \end{aligned}$$

The response-adjusted person weight is similarly calculated for all sampled persons.

5.3.3 Third-stage Final Person Weight

The final person weight is the weight used for producing person-level estimates, such as the proportion of persons with a specific attribute. This weight is appended to the STEP data file.

The final person weight is calculated as the product of the final household weight (which incorporates the PSU weight and the response adjustment for the basic household weight) and the response adjusted basic person weight. Therefore, for the selected person 'i' in household 'k' in PSU 'j', the final person weight, $W3_{SP}$ (i.e., weighted data file variable 'W_FinSPwt'), is calculated as follows:

$$16) \quad W3_{SP} = W2_{HH_k} * W3_{Radj}$$

➤ The final third stage person weight is appended to the STEP data file.

Example 12: Third Stage – Final Person Weight

- iii) For PSU #1, selected household # 05 with 3 eligible persons, the selected person was a participant.

$$W2_{HH_k} = 1063.0354; \quad W3_{Radj} = 3$$

$$W3_{SP} = W2_{HH_k} * W3_{Radj}$$

$$W3_{SP} = 1063.0354 * 3 \approx 3189.1062$$

- iv) For PSU #6, in participant household # 1 with 2 eligible persons, the selected person was a participant. For this household,

$$W2_{HH_k} \approx 1029.3115; \quad W3_{Radj} = 2.2727$$

$$W3_{SP} = W2_{HH_k} * W3_{Radj}$$

$$W3_{SP} = 1029.3115 * 2.2727 \approx 2339.3444$$

The final person weight is similarly calculated for all sampled persons. Note that the final person weight will be zero for all non-participating selected persons.

5.3.4 Weighted Person Counts

The weighted count of the non-institutionalized persons 15 to 64 years of age (inclusive) living in private dwellings in urban areas of **Kenya** at the time of data collection is provided in Table 19. Based on the 3824 participating selected persons, the sum of the final person weights from the STEP survey is 1,346,787.

Table 19: Sampled Persons & Weighted # of Persons by Response Category		
Response Category (SP)	Sampled Cases	Total # of Eligible Persons in Target Population
		(based on Final Person weight)
1-Participant SP	2,996	6,381,235
2-Non-participant, Eligible	36	0
Total	3,032	6,381,235

For participating selected persons, Table 20 provides a summary of the central tendency of the number of eligible persons in the participating households.

Table 20: Summary - Eligible Persons, Person Weights		
Statistic	Eligible Persons	Final Person Weight
No. of Cases	3911	3894
Mean	2.28	1638.74
Std. Error of Mean	0.020	17.72
Median	2.00	1445.70
Minimum	1	332.08
Maximum	9	14706.18
Sum	8899	6,381,235

6 Response Rate

The STEP response rate, R, is calculated as follows:

$$17) \quad R = \frac{P}{n - X} * 100\%$$

where,

P is the number of participating persons aged 15 to 64,

n is the number of activated sample cases,

X is the number of ineligible sample cases.

An overall response rate of 91.8% was achieved in the Kenya STEP Survey. Table 21 provides the detailed percentage distribution by final status code.

Table 21: Response Rate: Kenya, Activated HHs by Final Status Code

Case Final Status Code	Frequency	Percent
Ineligible HHs (N.B. % is based on total # of activated HHs, i.e., 3911 HHs)	0	0.0%
01-Refusal-Household	96	2.3%
02-Unusual circumstance-Household	18	0.4%
03-No knowledgeable household member found after 3 visits	62	1.5%
04-Temporarily absent/unavailable for field period	116	2.7%
05-No competent household member (e.g., severe illness, mental disability, etc.)	35	0.8%
21-Household Refusal to continue Household Module (begun, not complete)	4	0.1%
35-Unable to contact SP after 3 attempts	1	0.0%
98-Incorrect eligible household member selected	16	0.4%
Non-respondents (Eligible) (N.B. % is based on total # of Eligible HHs, i.e., 3911 HHs)	348	8.2%
51-Refusal by SP to continue modules 2-7	8	0.2%
61-SP unable to read language of General Booklet (GB)	159	3.7%
62-SP refusal to begin GB	143	3.4%
63-SP unable to do GB due to unusual circumstance	28	0.7%
64-SP blindness or visual impairment-unable to do GB	10	0.2%
65-SP physical disability	1	0.0%
71-SP looked through GB-unable to read/write language of GB	61	1.4%
72-SP refused to continue GB after starting	221	5.2%
73-GB interruption too long	29	0.7%
74-GB terminated - unusual circumstance	32	0.8%
75-Attempted GB, no answers marked	72	1.7%
76-Attempted part or all of GB; <3 correct Core answers	507	12.0%
81-No attempt at any Exercise Booklet items / refusal to start	184	4.3%
82-Interruption too long during Exercise Booklet administration	27	0.6%
83-Unusual circumstance, Exercise Booklet administration terminated	16	0.4%
84-Attempt at some Exercise Booklet items, no answers marked	139	3.3%
85-Attempt at all Exercise Booklet items, no answers marked	118	2.8%
86-Marked some Exercise Booklet items, not all items attempted	671	15.8%
87-Marked some Exercise Booklet items, all items attempted	780	18.4%
88-Marked all Exercise Booklet items	688	16.2%
Respondents (N.B. % is based on total # of Eligible HHs, i.e., 4786 HHs)	3894	91.8%
Total Eligible Households	4242	100.0%
Total Activated Households	4242	100.0%