

**Kenya 2009 National
Population and Housing Census
Edit Specifications**

18 March 2011

Kenya National Bureau of Statistics

Table of Contents

I. INTRODUCTION.....	1
TYPES OF ERROR.....	1
EDITING TEAM.....	2
BASICS OF EDITING.....	2
RISKS OF OVER-EDITING.....	3
DOCUMENTATION.....	3
II. OVERVIEW OF VARIABLES EDITED	4
III. EDIT SPECIFICATIONS FOR THE HOUSING RECORD	6
MODULE G: ANNUAL BIRTHS.....	6
MODULE H: INFORMATION ABOUT DEATHS IN THE HOUSEHOLD	6
MODULE I: INFORMATION REGARDING LIVESTOCK	8
MODULE J: HOUSING CONDITIONS AND AMENITIES	9
MODULE K: OWNERSHIP OF HOUSEHOLD ASSETS	12
MODULE I: EMIGRANTS	12
IV. EDIT SPECIFICATIONS FOR THE POPULATION RECORD	13
HEAD OF HOUSEHOLD	13
MODULE A: INFORMATION REGARDING ALL PERSONS	17
MODULE B: INFORMATION REGARDING FEMALES AGED 12 YEARS AND ABOVE.....	26
MODULE C: INFORMATION REGARDING PERSONS WITH DISABILITY	31
MODULE D: INFORMATION REGARDING PERSONS AGED 3 YEARS AND ABOVE	32
MODULE E: LABOUR FORCE PARTICULARS	35
MODULE F: INFORMATION REGARDING ICT.....	37
SPECIAL EDIT FOR OVERLAPPING CODES.....	38
SPECIAL EDIT FOR TRIBE CODES	39
V. EDIT SPECIFICATIONS FOR THE SHORT FORMS.....	40
OVERVIEW OF VARIABLES EDITED.....	40
COMMON VARIABLES	40
INSTITUTION-ONLY VARIABLES	42
EMIGRANT-ONLY VARIABLES.....	43
VI. STRUCTURAL EDITS	47
REFORMATTING LONG FORMS.....	47
REFORMATTING SHORT FORMS	49

I. INTRODUCTION

A well-designed census with minimal errors in the final product is an invaluable resource for a nation. To obtain useful census results, data must be free from errors and inconsistencies to the greatest extent possible, especially after the data processing stage. Census editing is the process of detecting errors that were made during and after data collection and capture, and then adjusting individual items to improve data quality.

No census data is ever perfect. As a result, national statistical offices have developed and implemented various approaches for dealing with data gaps and inconsistent responses. The exact editing procedures within countries vary from census to census because of changes in census content, loss of documentation from previous census editing, and revised international recommendations. As a result, countries have had to reinvent the process used in earlier data collection activities when a new census is planned.

This edit specifications document outlines the procedures developed and implemented by the Kenya National Bureau of Statistics (KNBS) to edit data from the Kenya 2009 Population and Housing Census.

Types of Error

Census data suffers from many types of errors that may be classified, generally, as coverage errors and content errors. Coverage errors arise from omissions or duplications of persons or housing units in the census enumeration. The sources of coverage errors include incomplete or inaccurate maps or lists of enumeration areas, failure by enumerators to canvass all of the units in their assignment areas, duplicate counting, omission of persons who are not willing to be enumerated, erroneous treatment of certain categories of persons such as visitors or travellers, and loss or destruction of census records after enumeration. Coverage errors should be resolved, to the greatest extent possible, in the field.

Content errors arise from the incorrect reporting or recording of the characteristics of persons, households, and housing units. Content errors can occur in the field or during data capture and processing. In the field, content errors can arise due to enumerator mistakes such as mistranslating, omitting questions, incorrectly recording responses, or misunderstanding a respondent. Careful training of enumerators minimizes the risk of generating content errors in the field. As with coverage errors, it is best to identify and correct content errors during field operations because enumerators can easily revisit households to obtain correct information. During data capture and processing, content errors can be caused by problems with data capture, coding, data entry, manual and computer editing, and tabulations of results. Data capture procedures and editing programs are carefully documented and reviewed to minimize the risk of generating these errors.

Editing Team

The editing process is the responsibility of an editing team that includes census managers, subject-matter specialists, and data processors. The editing team is responsible for addressing omissions and inconsistencies in the data after the data capture process. The editing team of KNBS subject-matter and data processing specialists was established at the beginning of the census process. The editing team began its work months before census enumeration by reviewing edit specifications from the previous census. After the 2009 Census data was captured, the team worked together to develop editing and imputation rules (described in this document) and the corresponding computer programs.

Basics of Editing

Editing is the systematic inspection and correction (or changing) of responses according to predetermined rules. Editing is done to ensure the validity and consistency of individual records and relationships among records in a household (micro-editing), and to check the reasonableness of the aggregated data (macro-editing). Census publications will contain a certain amount of meaningless data if national statistical offices do not edit the census or survey results. For example, unedited data may contain records that are highly unlikely or impossible, such as a 4 year-old child with a Ph.D. or a male who reported giving birth in the last year. Editing reduces distorted estimates, facilitates processing, and increases user confidence. However, editing cannot correct all census errors, including questionnaire responses that are internally consistent but are in fact instances of misreporting on the part of respondents or misrecording on the part of enumerators.

The editing team must develop “tolerance levels” for each item, and sometimes for combinations of items. Tolerance levels indicate the number of invalid and inconsistent responses allowed before the editing team takes remedial action. For most items in a census, for example, some small percentage of the respondents will not give “acceptable” responses, for whatever reason. For some items, like age and sex, which are used in combination with so many other items for planning, the tolerance level might be quite low. When the percentage of missing or inconsistent responses is low (less than one or two percent), any reasonable editing rules are not likely to affect the use of the data. When the percentage is high (five to ten percent, or more, depending on the situation), simple or even complex editing may distort the census results.

Imputation is the process of resolving problems concerning missing, invalid, or inconsistent responses identified during editing. Imputation works by changing one or more of the responses or missing values in a record or several records being edited to ensure that plausible, internally coherent records result. Imputation values for erroneous or missing items are generated by using other entries from the housing unit, person, or other persons in the household or a comparable group as a guide, always in accordance with specified procedures. Records should satisfy all edits after imputation. The imputation methods used by the KNBS editing team for the 2009 Census are noted in this document.

An imputed record should closely resemble its failed record. Imputing a minimum number of variables is usually best, thereby preserving as much respondent data as possible. The underlying assumption (which is not always true in practice) is that a respondent is more likely to make only one or two errors rather than several. The KNBS editing team retained files with the original data and edited data (including imputed data) in order to evaluate the effects of imputation on the individual and household records and aggregated data. During this comparison process, known as a reasonableness review, the KNBS editing team verified that the distributions of imputed variables did not change drastically from the original distributions.

Risks of Over-Editing

No amount of computer editing can take the place of higher quality census data collection. National statistical offices know that at some point computer editing is not only limited, but becomes counterproductive: the edit adds more errors to the data set than it corrects. Changing a census item is not the same as correcting it. Over-editing has a negative impact on the editing process in several ways, including timeliness, cost, and the distortion of true values. It also gives a false sense of security regarding data quality.

The more editing a national census/statistical office does, the longer the total process will take. The major issue is to determine how much the added time adds to the value of the census product. Often, the returns are so small in terms of the time invested that it is better to have small “glitches” in the data rather than deprive prime users of receiving the information on a timely basis. The KNBS editing team has edited the data to the extent appropriate to ensure data usability and timely dissemination.

Documentation

Editing procedures are documented and made available to data users and future census planners as a guide to understanding the data. The documentation aims to be complete enough for data users and census planners to be able to reconstruct the same processes at a later date to assure compatibility with the census or survey under consideration. This edit specification document has been produced by the KNBS editing team as a reference for data users and future census planners.

Odd results will appear in census tabulations no matter how much editing is done, so documentation is also provided in order to warn data users that small errors may occur. Additionally, if the editing team finds inconsistencies in released data, correction procedures will be followed if necessary and announcement will be made to data users.

II. Overview of Variables Edited

Variable	Is Don't Know a Valid Value?	Nonconventional Notes
Relationship	No	N/A
Sex	No	Special Edit
Age	No	Special Edit
Line number of mother	No	N/A
Usual member of household	No	N/A
Tribe/nationality	Yes	Special Edit
Religion	Yes	Special Edit
Marital status	No	
Birth place	No	
Previous residence	No	
Duration of residence – MM	No	
Duration of residence – YYYY	No	
Orphanhood – father	Yes	
Orphanhood – mother	Yes	
Children ever born – males	No	
Children ever born – females	No	
Children born alive live here – males	No	
Children born alive live here – females	No	
Children born alive elsewhere – males	No	
Children born alive elsewhere – females	No	
Children born alive then died – males	No	
Children born alive then died – females	No	
Last child born – MM	Yes	
Last child born – YYYY	Yes	
Last birth notified	Yes	
Last birth sex	No	
Last birth still alive	No	
Disabilities	No	
Economic troubles due to disability	Yes	
School attendance	Yes	
Schooling reached	Yes	
Schooling completed	Yes	
Economic activity	Yes	
Employer	No	
Hours worked	No	
Radio service	Yes	Special Edit
TV service	Yes	Special Edit
Mobile phone service	Yes	Special Edit

Variable	Is Don't Know a Valid Value?	Nonconventional Notes
Landline service	Yes	Special Edit
Computer service	Yes	Special Edit
Internet use frequency	Yes	
Internet use place	Yes	
Number household live births	No	N/A
Number household deaths	No	N/A
Death notified	Yes	N/A
Death age	Yes	N/A
Death sex	No	N/A
Death female mortality	Yes	N/A
Livestock (11 questions)	No	N/A
Dwelling units	No	N/A
Habitable	No	N/A
Tenure Status	No	N/A
Roof	No	N/A
Wall	No	N/A
Floor	No	N/A
Water source	No	N/A
Waste disposal	No	N/A
Cooking fuel	No	N/A
Lighting fuel	No	N/A
Household assets (14 questions)	No	N/A
Emigrants	No	N/A

III. Edit Specifications for the Housing Record

Module G: Annual Births

H10. How many live births occurred in this household between 24/08/2008 and 24/08/2009?

Count of Live Births	
Valid (0-10)	Invalid(11+,99), Blank
SawaSawa	Impute H10 by Hotdeck based on # of women aged 12-49 yrs

Module H: Information About Deaths in the Household

H11. How many deaths occurred in this household between 24/08/2008 and 24/08/2009?

Count of Death Records			
0:4		5	
if H11 = Count of Death Records		H11	
Yes	No	in 5:11	Otherwise
SawaSawa	Impute H11 to count of death records	SawaSawa	Impute H11 to count of death records

H13. Was this death notified?

Universe: Households reporting a death in the past 12 months

Death Notification	
Valid (1, 2, 9)	Invalid
SawaSawa	Impute H13 to previous valid H13 value

H14-16. Age, Sex and Maternal Mortality

H14 and H15 Universe: Households reporting a death in the past 12 months

H16 Universe: Deaths to females who died between the ages 12 and 49.

H14	H15	H16	Action
Age at Death	Sex at Death	MMx	
Invalid	Male	1:4,9	Impute AGE to 99, MMx to Blank
		Invalid	Remove Record
		Blank	Impute AGE to 99
	Female	1:4,9	Impute Age to 99
		Invalid	Remove Record
		Blank	Impute Age to 99, MMx to 9
	Invalid	1:4,9	Remove Record
		Invalid	
		Blank	
12:49	Male	1:4,9	Impute MMx to BLANK
		Invalid	
		Blank	
	Female	1:4,9	Sawasawa
		Invalid	Impute MMx to 9
		Blank	
	Invalid	1:4,9	Impute SEX to FEMALE
		Invalid	Remove Record
		Blank	Impute SEX to MALE
99	Male	1:4,9	Impute MMx to BLANK
		Invalid	
		Blank	
	Female	1:4,9	OK
		Invalid	Impute MMx to Don't Know
		Blank	
	Invalid	1:4,9	Impute SEX to FEMALE.
		Invalid	Remove Record
		Blank	Remove Record
Other Ages (ages outside fertility ages)	Male	1:4,9	Impute MMx to BLANK
		Invalid	Impute MMx to BLANK
		Blank	OK
	Female	1:4,9	Impute MMx to BLANK
		Invalid	Impute MMx to BLANK
		Blank	OK
	Invalid	1:4,9	Impute SEX to FEMALE
		Invalid	Remove Record
		Blank	Impute SEX to MALE

Module I: Information Regarding Livestock

H17. Livestock

Livestock			
	Livestock	Blank	Upper Limit (by region) to be determined once tabulations are run
1	Exotic Cattle	Impute to 0	10,000
2	Indigenous Cattle	Impute to 0	10,000
3	Sheep	Impute to 0	10,000
4	Goat	Impute to 0	10,000
5	Camels	Impute to 0	10,000
6	Donkeys	Impute to 0	10,000
7	Pigs	Impute to 0	10,000
8	Indigenous Chicken	Impute to 0	10,000
9	Chicken Commercial	Impute to 0	10,000
10	Bee Hives	Impute to 0	10,000
11	Other	Impute to 0	10,000

If Pigs > 0 and the first person record of a household is listed as Muslim, impute the value of Pigs to 0.

If Camels > 0 and no one in the household has a Main Employer code of 14 or 15 (Pastoralist Employed, Self Pastoralist), impute the value of Camels to 0.

Module J: Housing Conditions and Amenities

H18 - 19. Number of Dwelling Units and Habitable Rooms

Number of Dwelling Units Between 1 and 50			
Yes		No	
Is H19. Rooms in 1:50		Is H19. Rooms in 1:50	
Yes	No	Yes	No
H18. Dwelling Units < = H19. Habitable Rooms		Hot deck Rooms based on HHSIZE, Dwelling Units	Hot deck Dwelling Units based on HHSIZE, Rooms
Yes	No		
Sawasawa	Sawasawa		

H20. Tenure Status of Main Housing Unit

Tenure status in 1:10	
Yes	No
Sawasawa	Impute from previous valid value

H21 – 23. Dominant Construction Material of Main Dwelling Unit

- 1) If any of the three values (H21-H23) are X, make sure that the other values are also X and skip the edit.
- 2) Edit H23 Floor first. If H23 is not valid (1-5), impute from hotdecksas follows:
 - a. Impute H23 from roof and wall if roof and wall are both valid combinations.
 - b. Impute H23 from roof if only roof is valid.
 - c. Impute H23 from wall if only wall is valid.
 - d. Impute H23 from the previous household if neither the roof nor wall is valid.
- 3) After H23 is clean, use the following chart to check H21 and H23.

		Wall									
		Stone	Brick Block	Mud Wood	Mud Cement	Wood Only	Corrugated Iron Sheets	Grass Reeds	Tin	Other	Invalid
Roof	Corrugated Iron Sheets	OK	OK	OK	OK	OK	OK	OK	OK	OK	Hotdeck Wall Based on Roof, Floor
	Tiles	OK	OK	OK	OK	OK	OK	OK	OK	OK	
	Concrete	OK	OK	NO	Impute Wall to STONE	NO	NO	NO	NO	OK	
	Asbestos Sheets	OK	OK	OK	OK	OK	OK	OK	OK	OK	
	Grass	OK	OK	OK	OK	OK	OK	OK	OK	OK	
	Makuti	OK	OK	OK	OK	OK	OK	OK	OK	OK	
	Tin	OK	OK	OK	OK	OK	OK	OK	OK	OK	
	Mud/Dung	NO	NO	OK	OK	NO	NO	NO	NO	OK	
	Other	OK	OK	OK	OK	OK	OK	OK	OK	OK	
	Invalid	Hotdeck Roof Based on Wall, Floor									

Impute all cells marked "NO," imputing Roof using a hotdeck based on Wall and Floor

H24. Main Source of Water

Water Source	
Valid(1:15)	Invalid and Blank
Sawasawa	Impute H24 based on most recent valid value

H25. Main Mode of Human Waste Disposal

Waste Disposal	
Valid(1:9)	Invalid and Blank
Sawasawa	Impute H25 based on most recent valid value

H26. Main Type of Cooking Fuel

Cooking Fuel	
Valid(1:8)	Invalid and Blank
Sawasawa	Impute H26 based on most recent valid value

H27. Main Type of Lighting Fuel

		Lighting								
		Electricity	Pressure Lamp	Lantern	Tin Lamp	Gas Lamp	Fuel Wood	Solar	Other	Invalid
Cooking Fuel	Electricity	OK	OK	NO	NO	OK	NO	OK	OK	Hotdeck Main Source of Lighting Based on Cooking Fuel
	Paraffin	OK	OK	OK	OK	OK	OK	OK	OK	
	LPG	OK	OK	OK	NO	OK	NO	OK	OK	
	Bio Gas	OK	OK	OK	NO	OK	NO	OK	OK	
	Firewood	OK	OK	OK	OK	OK	OK	OK	OK	
	Charcoal	OK	OK	OK	OK	OK	OK	OK	OK	
	Solar	OK	OK	OK	NO	OK	NO	OK	OK	
	Other	OK	OK	OK	OK	OK	OK	OK	OK	
	Impute all cells marked "NO" using hotdeck based on Cooking Fuel									

Module K: Ownership of Household Assets

H28. Ownership of Household Assets

Specification applies to all 14 values in H-28

Ownership Of Household Assets		
Vaild 1:2	Invalid (Blank, 0)	Other
SawaSawa	Impute to 2	Any value > 2 impute to 1

Module I: Emigrants

H29. Emigrants in the last 15 years

Number of Emigrants		
Vaild (0:11)	Invalid (Blank)	Other
SawaSawa	Impute to 0	Any value > 11 report error, no imputation

IV. Edit Specifications for the Population Record

Head of Household

The following edits check the data for the head of household, specifically data that will be referenced by other edits/imputations. These edits do not apply to nonconventional households.

1. Check that there is one and only one head of household.
2. Check the sex of the head of household
3. Check the age of the head of household
4. Check the religion of the head of household
5. Check the marital status of the head of household
6. Check the ethnicity of the head of household

1. Check that there is one and only one Head of Household

Head of Household Count Is:

0		1		1+		
Age of First Person		Age of Head		Ages of Heads		
<15	15+	<15	15+	None 15+	One 15+	Multiple 15+
Pick first listed oldest person	Set as Head first person listed	Pick first listed oldest person	OK	Pick first listed oldest person, for other heads set relationship to BLANK	Take 15+ as head of household, for other heads set relationship to BLANK	Pick first 15+ head in list, for other heads set relationship to BLANK

2. Check the Sex of the Head of Household

Universe: Relationship is Head of Household

Head of Household's Sex Is:

Valid	Invalid							
OK	Number of Spouses Present							
	0		1			1+		
	Is Fertility for Head Valid (0-18 CEB) for Females?		Sex of Spouse is Known?			Find First Valid Sex of Spouses		
	Yes	No	Yes	No		Male	Female	Other
	Impute Sex to Female	Impute Sex of Head to Sex of Previous Head	Impute Sex of Head to Opposite of Spouse	Is Fertility for Head Valid (0-18 CEB) for Females?		Impute Sex to Female	Impute Sex to Male	Impute Sex to Male
				Yes	No			
				Is Fertility for Spouse Valid (0-18 CEB) for Females?				
				Yes	No			
				Impute Sex of Head to Female	Impute Sex of Head to Male	Impute Sex of Head to Sex of Previous Head		

3. Check the Age of the Head of Household

Universe: Relationship is Head of Household

Head of Household's Age Is:

0-14, 99, Blank					15-95	96-98
Is Spouse Present with Age >= 12						
Yes		No				
Is Child Present with Valid Age (Look for Oldest Child)						
Yes		No				
Is Sibling Present with Age >= 15						
Yes		No				
Is Parent Present with Age >= 27						
Yes		No				
Hotdeck Head's Age on Sex * Spouse Age	Hotdeck Head's Age on Sex * Oldest Child Age	Hotdeck Head's Age on Sex * Sibling Age	Hotdeck Head's Age on Sex * Parent Age	Hotdeck Head's Age on Sex * Size of Household	OK	Impute age to 95

4. Check the Tribe of the Head of Household

Universe: Relationship is Head of Household

If tribe of head is invalid then

If sex is male then

Impute tribe of head as first valid tribe of blood relative (3, 5, 6)

If no such blood relative exists

Set tribe to "not stated" if tribe is blank, "invalid" otherwise

If sex is female then

Impute tribe of head as first valid tribe of blood relative (5, 6, 3)

If no such blood relative exists

Set tribe to "not stated" if tribe is blank, "invalid" otherwise

5. Check the Religion of the Head of Household

Universe: Relationship is Head of Household

If religion of head is invalid (0, blank) then
 Get religion from 1st valid relative in household
 If no valid religion in household then
 Set religion to be “not stated”

6. Check the Marital Status of the Head of Household

Universe: Relationship is Head of Household

Head of Household’s Marital Status is:

Not Married (Never married, Divorced, Separated, Widowed)	Count of Spouses is	0	OK
		1	Impute as Married Monogamous
		2+	Impute married polygamous
Monogamous	Count of Spouses is	0	OK
		1	OK
		2+	Impute married polygamous
Polygamous	Count of Spouses is	0	OK
		1	OK
		2+	OK
Not a valid value	Count of Spouses is	0	Impute from Hot Deck (Age, Sex)
		1	Impute as Married Monogamous
		2+	Impute as Married Polygamous

B. Population Record

The universe for all Population Record edits is all persons, unless otherwise noted.

Module A: Information Regarding All Persons

P-00. What are the names of each person who spent the night of 24th/25th August 2009 in this household?

No edit required here. These data are not captured.

P-10. What is <NAME>'s relationship to the head of this household?

Universe: Non heads of household, conventional households

Relationship value is invalid or DK (99)			
Age is			
Valid			Invalid
Age Difference with Head			Impute relationship to Other Relative
Impute relationship based on hotdeck of Age Difference * Head Sex * Sex; afterwards, if hotdeck introduces excessive parents(or changes head's orphanhood status), grandparents, or spouses (i.e., count of spouses as well as head's marital status), impute relationship to Other Relative			

P-11. What is <NAME>'s sex?

Universe: Non heads of household

CONVENTIONAL HOUSEHOLDS:

Sex value is invalid			
Relationship is			
Spouse		Not Spouse	
Sex is		Is there valid Fertility data	
Valid	Invalid	Yes	NO
Sex of Spouse = Sex of Head		Impute sex to female	Impute sex to male/female/male/female/etc...
Yes	Impute sex of spouse opposite sex of head		
Impute sex of spouse opposite sex of head			
No	OK		

After editing sex, check to make sure that the household only has at most 1 father, 1 mother, 2 grandmothers, and 2 grandfathers. If the number exceeds these limits, set the relationship for just-edited case to other relative.

NONCONVENTIONAL HOUSEHOLDS:

- 1) If sex is invalid, check if valid fertility data (CEB in 0:18) exists for the person. If it exists, set the person to female.
- 2) If no valid fertility data exists, randomly choose one of the previous ten valid values of sex to assign the person.

P-12. How old is <NAME>?

Universe: Non heads of household

CONVENTIONAL HOUSEHOLDS:

1) If age is in 96:98, set age to 95.

Age Difference with Head = Age – Age (HEAD)

<p>Age is VALID, or Relationship is Spouse, Child, Grandchild, Sibling, Parent, Grandparent</p> <p><i>(note that if the hotdeck cannot generate a valid age, for instance a grandchild to a head that is 20-years-old, change relationship to Other Relative)</i></p>	Relationship	Spouse	AGE	0:11	Impute age based on head's age and sex, and relationship
				12+	OK
		Son/Daughter	AgeDiff	> -12 If female, also < -55	Impute age based on head's age and sex, and relationship
				Otherwise	SawaSawa
		Grand Child	AgeDiff	> -24	Impute age based on head's age and sex, and relationship
				Otherwise	SawaSawa
		Brother Sister	AgeDiff	-30 to +30	SawaSawa
				Otherwise	Impute age based on head's age and sex, and relationship
		Father Mother	AgeDiff	Male head: < 12 Female head: < 12 and > 55	Impute age based on head's age and sex, and relationship
				Otherwise	SawaSawa
		Grandparent	AgeDiff	< 24	SawaSawa
				Otherwise	Impute age based on head's age and sex, and

					relationship
		Other	Sawasawa		
INVALID	See Below				

AGE is INVALID		
If female and valid CEB present		
Yes	NO	
Hot Deck based on CEB	If Attending School and Valid Educational Attainment	
	Yes	NO
	Hot deck age based on level of education	Hot Deck Age based on Relationship, Marital Status, and Sex

NONCONVENTIONAL HOUSEHOLDS:

- 1) If age is in 96:98, set it to 95.
- 2) If age is not in 0:95, impute age from a hotdeck based on sex and education reached (P40). If there is no valid cell in that hotdeck, randomly choose one of the previous ten valid values of age to assign the person.

P-13. Please insert line number of <NAME>'s biological mother

This edit should be done after Relationship, Sex, and Age have been fully edited for the whole form

0				
Relationship				
Head		Son/Daughter		Other
Parent is present		Sex of Head		
yes	no	Male	Female	OK
Sex of parent		OK	Impute P13 of Son/Daughter to line number of head	
male	female			
OK	Impute P13 of head to line number of mother	OK		

Valid (P13 between 1 and total number of persons in HH)												
Relationship												
Head		Son/Daughter		Father/Mother		Other						
Relationship of person identified in P13 is female parent		Relationship of person identified in P13 is Head or Spouse and is Female		If relationship of person referenced in P13 is female grandparent		If sex of person referenced in P13 is female						
Yes	No	Yes	No	Yes	No	Yes	No					
OK	Is there a female parent in the HH		OK	Is there a female head or spouse in the HH		OK	Is there a female grandparent in the HH		Is age difference between "other" and person referenced in P13 >=12 and <= 55	Impute P13 to 0		
	Yes	No		Yes	No		Yes	No			Yes	No
	Impute P13 to the line number of the female parent	Impute P13 to 0		Impute P13 to the line number of the female head/spouse	Impute P13 to 0		Impute P13 to the line number of the female grandparent	Impute P13 to 0			OK	Impute to 0

Invalid (NOT in 0 to number of records)						
Relationship						
Head		Son/Daughter		Other		
Parent is present		Sex of Head		Impute P13 to 0		
Yes	No	Male	Female			
Sex of parent		Impute to 0	Impute to 0			Impute P13 of Son/Daughter to line number of head
Male	Female					
Impute to 0	Impute P13 of head to line number of mother	Impute to 0	Impute to 0	Impute P13 of Son/Daughter to line number of head	Impute P13 to 0	

After this edit make sure that each mother does not have more children associated with her than allowed (in the fertility edits).

P-14. Is <NAME> a usual member of this household?

Usual Member of Household						
Yes	No		0		Otherwise (Invalid)	
Sawasawa	Relationship		Relationship		Relationship	
	Head	Other Relationship	Head	Other Relationship	Head	Other Relationship
	Impute to usual member of household	Sawasawa	Impute to usual member of household	Impute to 2	Impute to usual member of household	Hot Deck Usual Member (Relationship, Duration of Residence)

P-15. What is <NAME>'s tribe or nationality?

CONVENTIONAL HOUSEHOLDS:

Universe: Non heads of household

If tribe is invalid then

 If relative (codes 1-10) then impute tribe from tribe of head

 Else (non-relative) then set tribe code as either Not Stated or Invalid

NONCONVENTIONAL HOUSEHOLDS:

Tribe value is...	
Valid (in value set)	Invalid
OK	Set tribe code as either Not Stated or Invalid

P-16. What is <NAME>'s religion?

CONVENTIONAL HOUSEHOLDS:

Universe: Non heads of household

If religion is invalid, impute religion to religion of head if person is a relative (P10 is 2:10), otherwise set to Not Stated.

NONCONVENTIONAL HOUSEHOLDS:

If religion is invalid, set to Not Stated.

P-17. What is <NAME>'s marital status?

Universe: Non heads of household

If Age < 12 then marital status must be never married

Marital Status of Spouse <i>(Age, Sex, Head data have been edited and age >= 12)</i>	
Marital Status = Head's Marital Status	
Yes	No
OK	Impute to Head's Marital Status

If relationship is not head or spouse (or nonconventional households)

Is MS a valid value (1-6)?	
Yes	No
OK	Hot Deck MS (AGE, SEX)

P-18. Where was <NAME> born?

Place of Birth		
Valid	Invalid	
OK	Is Duration of Residence Valid?	
	Yes	No
	Is the (Date of Census) - (Duration of Residence) = Age	
	Yes	No
	Impute POB as District of enumeration	Impute POB based on urban/rural status and tribe
	impute POB based on urban/rural status and tribe	

If the Urban/Rural Status x Tribe hotdeck does not have a valid cell, impute POB as district of enumeration.

P-19. Where was <NAME> living in August 2008?

- 1) If P12 (age) is 0, then P19 must be 0.
- 2) If Duration of Residence is 8/2008 or earlier, P19 must be district of enumeration.
- 3) If P19 is otherwise not valid:
 - a. If the first person on a questionnaire, use the Urban/Rural Status x Tribe hotdeck. If the cell in that hotdeck is not valid, set P19 to P18.
 - b. If not the first person on a questionnaire, set to the previous person's P19 value that wasn't 0.

P-20. When did <NAME> move to the current district (Month)?

P-21. When did <NAME> move to the current district (Year)?

First impute year, then month:

Any year before 1914, set to 1914

This edit has to be done after line number of mother is edited

Year is...				
Valid (1914-2009 and >= year of birth)	Too early (year < year of birth)	Else		
OK	Impute Year to year of birth	Is person's line number of mother not 0 and does this person's mother have a valid year?		
		Yes		No
		Is mother's year >= year of birth?		
		Yes	No	
		Impute Year to mother's year	Impute Year to year of birth	Impute Year from Province of Birth Place x Sex x Age hotdeck

Month is...		
Valid (1-12 for years before 2009, 1-8 for 2009)	Invalid	
OK	Does mother exist in household and has a valid month?	
	Yes	No
	Impute Month of mother's month	Impute Month randomly using selection proportions determined by valid months

P-22. Is <NAME>'s father alive?

- 1) If P22 = Yes, then OK. Otherwise:
 - a. If relationship is head and there exists in the household a person with relationship parent and sex male, change P22 to Yes.
 - b. If relationship is child and the head is male, change P22 to Yes.
 - c. If P22 is invalid and relationship is sibling, set P22 to be the head's P22.
 - d. If P22 is still invalid, set it to 9 DK.

P-23. Is <NAME>'s mother alive?

Line number of mother must be edited before this edit

- 1) If P23 = Yes, then OK. Otherwise:
 - a. If line number of mother is not 0, change P23 to Yes.
 - b. If P23 is invalid and relationship is sibling, set P23 to be the head's P23.
 - c. If P23 is still invalid, set it to 9 DK.

Module B: Information Regarding Females Aged 12 Years and Above

P-24 – P31. Particulars of All Live Births

Universe: Women aged 12 and above; if sex is male or age < 12, delete fertility information

Maximum Fertility by Age

Age	Max Number of Children
12:15	2
16:19	4
20:24	6
25:29	8
30:39	12
40:49	16
50+	18

Not Valid = not a number between 0 and max of possible number of children per woman's age

Fertility edits for male and female children are the same. Only male edits are listed in this document.

Fertility for Male Children Ever Born.

Count the number of male records in the household that indicate woman's line number as P13 and are listed as usual members of the household. If P26 > the above count then impute P26 to the count.

If P24, P26, P28, and P30 have the same value, then set 26, P28, and P30 to blank so that the edits below apply (P24 value remains unchanged).

Scenario Number	P24. CEB	P26. In HH	P28. Elsewhere	P30. Dead	Action
1	Valid	Valid	Valid	Valid	OK Check total CEB
2	Valid	Valid	Valid	Invalid	Dead = CEB - InHH - Elsewhere If Dead < 0 then impute dead to 0 Check total CEB
3	Valid	Valid	Invalid	Valid	Elsewhere = CEB - InHH - Dead If Elsewhere < 0 then impute Elsewhere to 0 Check total CEB
4	Valid	Valid	Invalid	Invalid	If CEB = Children in the HH, impute P28 and P30 to 0
					If P26 > P24, impute P24 = P26 and P28 and P30 to 0
					If P24 > P26, use hot deck based on age, CEB, and Children in the HH to impute P28, calculate P30 using formula P30 = CEB - InHH - Elsewhere
5	Valid	Invalid	Valid	Valid	In HH = CEB - Elsewhere - Dead if in HH < 0 then impute in HH to 0 Check Total CEB
6	Valid	Invalid	Valid	Invalid	If CEB = Children elsewhere, impute Children in HH and Dead to 0
					If P28 > P24, impute P24 = P28 and P26 and P30 to 0
					If P24 > P28, use hot deck based on age, CEB, and Elsewhere to impute P26, calculate P30 using formula P30 = CEB - InHH - Elsewhere
7	Valid	Invalid	Invalid	Valid	If CEB = Children dead, impute Children in HH and elsewhere to 0
					If P30 > P24, impute P24 = P30 and P26 and P28 to 0
					If P24 > P30, use hot deck based on age, CEB, and Children dead to impute P26, calculate P28 using formula P28 = CEB - InHH - dead
8	Valid	Invalid	Invalid	Invalid	If P24 = 0 impute Child in HH, Child Elsewhere, Child dead to 0

Scenario Number	P24. CEB	P26. In HH	P28. Elsewhere	P30. Dead	Action
					Else use Hot deck base on age, and CEB to impute P26; then follow Scenario 4
9	Invalid	Valid	Valid	Valid	Impute CEB to InHH+Elsewhere+Dead
10	Invalid	Valid	Valid	Invalid	Use Hot deck based on age, Child in HH and Children Elsewhere to impute Children dead, calculate CEB = Dead + InHH + Elsewhere
11	Invalid	Valid	Invalid	Valid	Use Hot deck based on age, Child in HH and Children dead to impute Children elsewhere, calculate CEB = Dead + InHH + Elsewhere
12	Invalid	Valid	Invalid	Invalid	Use Hot deck based on age, and Child in HH to impute CEB; then follow Scenario 4
13	Invalid	Invalid	Valid	Valid	Use hotdeck to calculate P26 based on Age, P28, P30; CEB = P26+P28+P30
14	Invalid	Invalid	Valid	Invalid	Use Hot deck based on age, and Children Elsewhere to impute CEB; then follow Scenario 6
15	Invalid	Invalid	Invalid	Valid	Use Hot deck based on age, and Children Dead to impute CEB; then follow Scenario 7
16	Invalid	Invalid	Invalid	Invalid	Use hotdeck to calculate CEB based on age, then follow scenario 8

After processing the above table, check if $P24 > 0$ and $P26=P28=P30=0$; if so, set P28 to P24.

Check total CEB	
CEB = InHH + Elsewhere + Dead?	
Yes	No
OK	Impute CEB to sum InHH + Elsewhere + Dead

If CEB (male and female combined) > max number of children for age:

- Subtract by one randomly from P26-P31 until CEB = max number of children for age.
- Maintain that Children in Household \geq # with line number of mother who are listed as usual members of the household.
- Maintain that Children Born – Children Dead \geq # with line number of mother.

P-32-33. Particulars of Last Live Births

Universe: Women aged 12 and above

If P24 and P25 = 0 then impute P32 and P33 to Blank

Valid years criteria:

- Woman can give birth between 12 and 54 years old;
- Year of last birth cannot be before the year of birth of the woman's youngest child in household

- 1) If there is a 0-year-old, set year of birth to 2008 if month is 9:12 and 2009 if month is 1:12, or impute randomly if month isn't set
- 2) If CEB = children in household and that number of children are on the form, set year of birth to year of youngest child's birth; regard above for month rules
- 3) If P33 is still not valid (see above "valid years criteria"), set P33 to 9999

When was your last child born?							
Year							
Valid				DK			
Month				Month			
Valid (1:8 if 2009, 1:12 otherwise)	Invalid				99	Else	
SawaSawa	Less than 2008	Equal 2008		Equal to 2009		Sawasawa	Impute to 99
	Impute month Randomly (1:12)	If 0 year old and reports case line number as P13,		If 0 year old and reports case line number as P13,			
		Yes	No	Yes	No		
		Impute month randomly (8:12)	Impute month randomly (1:12)	Impute month randomly (1:8)	Impute month randomly (1:12)		

P-34. Was this last birth notified?

Universe: Women aged 12 and above

Particulars of last live birth		
If no births	If there are births	
Impute P34 to Blank	Was this birth Notified? (P34)	
	Valid (1, 2, 9)	Invalid
	SawaSawa	Use Hot deck based on age and Education of mother and EA type

P-35. Was this child male or female?

Universe: Women aged 12 and above

(First check if P36 = Both Twins; if so, and P35 = Male, change to Male Twins; if P35 = Female,

change to female Twins)

If P35 = 3 (male twins) and P24 < 2, change P35 to 1 (male) if P24 = 1, to 2 (female) otherwise

If P35 = 4 (female twins) and P25 < 2, change P35 to 2 (female) if P25 = 1, to 1 (male) otherwise

If P35 = 5 (male/female twins) and either P24 or P25 is 0, change P35 to the non-zero sex

If P35 = 6 (multiple) and P24+P25 < 3, change P35 to 3 if P24=2, to 4 if P25=2, to 5 if P24&P25=1, if not any of the previous, to 1 if P24=1, to 2 if P25=1

If child listed in P32/P33 is identified in household, P35 must match sex of child (making sure to check for all children of given birth year [for cases of twins or multiples])

If child is not listed in household, sex in P35 must be accounted for in either P24 or P25

If sex is still undetermined, pick sex proportionally based on numbers in P24+P25

P-36. Is this child still alive?

Is this child still alive?		
# of Births for this Woman (P24 + P25)		
0		# of Births > 0
		Valid (1:7,DK)
Is P35 Blank?		Invalid
Yes	No	
OK	Impute P35 to Blank	
		Refer to Sheet Table P36/35

Table P36/35:

1 Yes	If CEB = #dead, change to No; if P35 in 3:5, change P36 to 4; if P35=6, change to 7
2 No	If dead = 0, change to Yes
3 One of the twins	P35 in 3:5, if dead = 0, change to 4 (both twins)
4 Both twins	P35 in 3:5, if num. kids still alive = 1, change to 3, if = 0, change to 2
5 one of the multiples	P35 = 6, if dead = 0, change to 7
6 two of the multiples	P35 = 6, if dead = 0, change to 7, if dead = 1, change to 5
7 all of the multiples	P35 = 6, if num. kids still alive = 2, change to 6, if num kids still alive = 1, change to 5
Else	If can find kid(s) in household, set to Yes; otherwise hotdeck on P35,CEB,dead

Module C: Information Regarding Persons with Disability

P37 – P38. Disability

To edit P37:

1. Impute all blanks to “8”
2. Impute all invalid to “8”
3. Remove duplicate disabilities
4. If Age < 5 and person reported Self Care, impute that value to “8”
5. Shift all valid disability codes to the left

To edit P38:

1. If Age < 5 then P38 should be NA
2. If P37 = Self Care (only) then P38 should be NA
3. Otherwise use the following chart:

Error #	P37A	P37B	P37C	P38	Action
1001	no Disability	no Disability	no Disability	1	Impute P38 to 3
1002	no Disability	no Disability	no Disability	2	Impute P38 <-3
1003	no Disability	no Disability	no Disability	3	OK
1004	no Disability	no Disability	no Disability	9	Impute P38 <-3
1005	no Disability	no Disability	no Disability	Invalid	Impute P38 <-3
1006	with Disability	no Disability	no Disability	1	OK
1007	with Disability	no Disability	no Disability	2	OK
1008	with Disability	no Disability	no Disability	3	Impute P38 <-9 (for age >= 5)
1009	with Disability	no Disability	no Disability	9	OK
1010	with Disability	no Disability	no Disability	Invalid	Impute P38 <-9 (for age >= 5)
1011	with Disability	with Disability	no Disability	1	OK
1012	with Disability	with Disability	no Disability	2	OK
1013	with Disability	with Disability	no Disability	3	Impute P38 <-9 (for age >= 5)
1014	with Disability	with Disability	no Disability	9	OK
1015	with Disability	with Disability	no Disability	Invalid	Impute P38 <-9 (for age >= 5)
1016	with Disability	with Disability	with Disability	1	OK
1017	with Disability	with Disability	with Disability	2	OK
1018	with Disability	with Disability	with Disability	3	Impute P38 <-9 (for age >= 5)
1019	with Disability	with Disability	with Disability	9	OK
1020	with Disability	with Disability	with Disability	Invalid	Impute P38 <-9 (for age >= 5)

Module D: Information Regarding Persons Aged 3 Years and Above

Universe: Persons 3 years and above

Education Age Reference Table	Code	Minimum age for highest level reached	Minimum age for highest level completed	Maximum age for attending school
Standard 1 (incomplete)	0	4	4	45
Standard 1	1	4	5	46
Standard 2	2	5	6	47
Standard 3	3	6	7	48
Standard 4	4	7	8	49
Standard 5	5	8	9	50
Standard 6	6	9	10	51
Standard 7	7	10	11	52
Standard 8	8	11	12	53
Form 1	9	12	13	
Form 2	10	13	14	
Form 3	11	14	15	
Form 4	12	15	16	
Form 5	13	16	17	
Form 6	14	17	18	
Incomplete Tertiary College	15	16	16	
Completed College	16	18	18	
Incomplete Undergraduate	17	16	16	
Completed Undergraduate	18	20	20	
Incomplete Masters/PHD	19	21	21	
Completed Masters/PHD	20	23	23	
Incomplete Basic/Post Literacy	21	5	5	
Completed Basic/Post Literacy	22	8	8	
Incomplete Polytechnic	23	12	12	
Completed Polytechnic	24	13	13	
Incomplete Madrassa	25	3	3	
Completed Madrassa	26	12	12	
Pre-primary	96	3	4	24
Never Attended	97	0	0	
Not Stated/DK	99	0	0	

P39. Schooling Attendance

P39. School Attendance	1. Currently	Is the age too high for P40 Level Reached?	Yes	Impute P39 to Previously				
			No	Sawa Sawa				
	2. Previously	Sawa Sawa						
	3. Never	Is P40 = 97? Is P41 = 97?	Yes	Sawa Sawa				
			No	Impute P40 to 97 Impute P41 to 97				
	9. Don't Know	Is P40 = 97	Yes	Is P41 = 97	Yes	Sawa Sawa		
					No	Impute P41 to 97		
			No	Is P41 = 97	Yes	Impute P40 to 97		
					No	Is P40 a valid value for the age and it is not 0?	Yes	Hot deck school attendance based on P40, Age
							No	Impute P40 and P41 to 97
	Invalid	Is P40 = 97	Yes	Is P41 = 97	Yes	Impute P39 to Don't Know		
No					Impute P41 to 97 and P39 to Don't Know			
No			Is P41 = 97	Yes	Impute P40 to 97 and P39 to Don't Know			
				No	Is P40 a valid value for the age and it is not 0?	Yes	Hot deck school attendance based on P40, Age	
						No	Impute P39 to DK, P40 and P41 to 97	

P40 – P41. Education Levels Reached / Completed

1) If P39 = 9, set P40 and P41 to 98.

2) Check to see if P40 is valid for the age, using the age boundaries found in the Education Age Reference Table. If P40 is valid for the age, check the combination with P41 in the table below. If the combination is BLK, impute P41 to blank. Otherwise, if the combination is not GOOD, impute P41 from a hotdeck of P40, age, and sex.

	P41. Level Completed																													
	Standard 1 (incomplete)	Standard 1	Standard 2	Standard 3	Standard 4	Standard 5	Standard 6	Standard 7	Standard 8	Form 1	Form 2	Form 3	Form 4	Form 5	Form 6	Incomplete Tertiary College	Completed College	Incomplete Undergraduate	Completed Undergraduate	Incomplete Masters/PHD	Completed Masters/PHD	Incomplete Basic/Post Literacy	Completed Basic/Post Literacy	Incomplete Polytechnic	Completed Polytechnic	Incomplete Madrassa	Completed Madrassa	Pre-primary	Never Attended	Blank
P40. Level Reached	Standard 1 (incomplete)																											GO OD	BLK	GO OD
	Standard 1	GO OD																										GO OD	BLK	GO OD
	Standard 2	GO OD	GO OD																											
	Standard 3		GO OD	GO OD																										
	Standard 4			GO OD	GO OD																									
	Standard 5				GO OD	GO OD																								
	Standard 6					GO OD	GO OD																							
	Standard 7						GO OD	GO OD																						
	Standard 8							GO OD	GO OD																					
	Form 1								GO OD	GO OD																				
	Form 2									GO OD	GO OD																			
	Form 3										GO OD	GO OD																		
	Form 4											GO OD	GO OD																	
	Form 5												GO OD	GO OD																
	Form 6													GO OD	GO OD															
	Incomplete Tertiary												GO OD	GO OD																
	Completed College																GO OD													
	Incomplete Undergraduate												GO OD	GO OD																
	Completed Undergraduate																	GO OD												
	Incomplete Masters/PHD																		GO OD											
	Completed Masters/PHD																				GO OD									
	Incomplete Basic/Post																											BLK	GO OD	
	Completed Basic/Post																					GO OD								
	Incomplete Polytechnic								GO OD	GO OD	GO OD	GO OD	GO OD	GO OD																
	Completed Polytechnic																							GO OD						
	Incomplete Madrassa																											BLK	GO OD	
	Completed Madrassa																									GO OD				
	Pre-primary																										GO OD	BLK	GO OD	

3) If here, then P40 was not valid. Check if P41 is valid for the age and if it not an "incomplete" code. If it is valid, impute P40 from a hotdeck of P41, age, and sex.

4) If still here, then neither P40 nor P41 was valid. Impute P40 from a hotdeck of age and sex. Then impute P41 from a hotdeck of P40, age, and sex.

Module E: Labour Force Particulars

Universe: Persons 5 years and above

P42 – P42. Economic Activity and Employer

1) First employ these “special edits”:

- a) If P42 is economically active (1:7), P43 is economically active (1:17), and P44 is 99 (DK), impute P44 to 98.
- b) If P42 is 15 (other), P44 >= 1, and P43 in 12:15, hotdeck P42 by P43, sex, and age
- c) If P42 is 13 (full time student) and P39 does not equal 1 (currently attending school), hotdeck P42 by P43, P44 and sex.
- d) If P42 is 15 (other), P43 is 17 (other), and P39 is 1 (currently attending school), impute P42 to 13 (full time student), P43 to 99 and P44 to 99.
- e) If P42 is 99, P43 is 99, and P39 = 1 (currently attending school), impute P42 to 13 (full time student).

2) Check if the values of P42 or P43 are valid for the age using the table below:

P42. Economic Activity	Minimum Age	P43. Main Employer	Minimum Age
Worked for Pay	5	Private Sector	5
On Leave	16	Local Authorities	18
Sick Leave	16	Central Government	18
Own - Family Business	5	TSC	18
Own - Family Agriculture Holding	5	State Owned Enterprise	18
Intern/Apprentice	14	International NGO	18
Volunteer	18	Local NGO	18
Seeking Work (Action Taken)	15	Faith Based Organisation	18
Seeking Work (No Action Taken)	15	Self Modern	5
No Work Available	15	Informal Sector ('Jua Kali')	5
Retired	50	Self Employed – Informal	5
Homemaker	15	Small Scale Agriculture	5
Full Time Student	5	Self Small Scale Agriculture	5
Incapacitated	5	Pastoralist Employed	5
Other	5	Self Pastoralist	5
		Private Household	5
		Other	5

3) If the combination of P42 and P43 in the table below is bad, impute P43 based on age, sex, and P42. Other imputations are listed in the table.

		P43. Employer																		
		Private Sector	Local Authorities	Central Government	TSC	State Owned Enterprise	International NGO	Local NGO	Faith Based Organisation	Self Modern	Informal Sector ('Jua Kali')	Self Employed - Informal	Small Scale Agriculture	Self Small Scale Agriculture	Pastoralist Employed	Self Pastoralist	Private Household	Other	Not Applicable	Invalid
P42 Main Activity	Worked for Pay	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	BAD	OK	BAD	OK	OK		
	On Leave	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK		
	Sick Leave	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK		
	Own - Family Business	OK	BAD	BAD	BAD	BAD	BAD	BAD	OK	OK	OK	OK	OK	OK	BAD	BAD	BAD	OK		Impute Employer based on AGE, SEX, Main Activity
	Own - Family Agriculture Holding	OK	BAD	BAD	BAD	BAD	BAD	BAD	OK	OK	BAD	BAD	OK	OK	OK	OK	OK	OK		Impute Employer based on AGE, SEX, Main Activity
	Intern/Apprentice	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	BAD	BAD	BAD	BAD	BAD	OK	OK		
	Volunteer	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	BAD	BAD	BAD	BAD	BAD	OK	OK		
	Seeking Work (Action Taken)	If P44 is in 1:98, then impute Main Activity based on AGE, SEX, Employer; otherwise impute Employer to BLANK																		
	Seeking Work (No Action Taken)																			
	No Work Available																			
	Retired																			
	Homemaker																			
Full Time Student																				
Incapacitated																				
Other	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK		Impute Employer based on AGE, SEX, Main Activity
99	<i>Imputation 2: Impute Main Activity based on AGE, SEX, Employer</i>																	Impute P43 to BLANK	Impute P43 to BLANK	
Invalid	<i>Imputation 2: Impute Main Activity based on AGE, SEX, Employer</i>																	Impute P42 to 99 Impute P43 to BLANK	Impute P42 to 99 Impute P43 to BLANK	

P44. Hours worked last 7 days

P42 in 1:7,15			
Yes		No	
Hours Worked		Hours Worked	
Valid (0:98)	otherwise (BLANK, 99)	Blank	Other
Sawa sawa	Hot Deck based on age, employer	Sawa Sawa	Impute to BLANK

Module F: Information Regarding ICT**P-45 – P49. ICT**

1. If age is less than 3 then impute all values to blank
2. If P45-49 = 0 then impute all values to 2
3. If P45-49 does not equal 1, 2, 9 then:
 - a. If conventional household, check the corresponding value for H28 and hotdeck on the corresponding H28 value by age (1 year intervals until 70+)
 - b. If nonconventional household, impute to 9

P50 – P51. Internet use

If age is less than 3 then impute all values to blank

Internet Use									
0		1:4			5,9		Otherwise (6, 7, 8, BLANK)		
Where Used?		Where Used?			Where Used?		Where Used?		
Non-Blank	Other (NR)	1 to 8	9	Other (NR/Invalid)	Non-Blank	Other (NR)	0	1 to 9	Other (NR/Invalid)
Impute P50 to 5 Impute P51 to BLANK	Impute P50 to 5	OK	OK	Impute P51 using hotdeck based on P50 and age	Impute P51 to Blank	OK	Impute P50 to 9 Impute P51 BLANK	Hotdeck P50 based on AGE, P51	Impute P50 to 9

Special Edit for Overlapping Codes

Birth Place (P18) and Previous Residence (P19) have 20 overlapping codes (801 to 820) where values correspond to the districts in Western Province or 20 countries in Africa. All African countries (codes 800 to 851) have been reassigned to previously unused codes (20 to 71).

P18. Birth Place

Is tribe/nationality a foreign nationality code?						
No, P15 < 800 or P15 = 999		Yes, P15 in 800 to 990				
Is birth place in 821 to 851?		Is the district of enumeration in Western Province?				
Yes	No	Yes, Province = 8		No, Province ≠ 8		
Recode P18 in 821:851 to new country codes 41:71	OK	Is P18 in 801 in 820?		Is P15 = P18?		
		Yes	No	Yes	No	
		Recode P15 to INVALID (P15 = 0)	Recode P18 in 821:851 to new country codes 41:71	Recode P18 in 800:851 to new country codes 20:71	Is P18 < 821?	
					Yes	No
					OK	Recode P18 in 821:851 to new country codes 41:71

P15. Tribe/Nationality

Is tribe/nationality a foreign nationality code?	
Yes, P15 in 800:990	No, P15 < 800 or P15=999
Recode P15 in 800:851 to new country codes 20:71	OK

P19. Previous Residence

Is P19 different from the district of enumeration (DoE)?				
No, P19 = DoE or 0	Yes, P19 ≠ DoE			
OK	Is P19 in 801:820?			
	No	Yes		
	Recode P19 in 800, 821:851 to new country codes 20, 41:71	Is P19 equal to nationality (P15)?		
		No	Yes	
		OK	Recode P19 in 801:820 to new country codes 21:40	

Special Edit for Tribe Codes

A post-censal analysis of the tribe data revealed some problems in the values that some enumerators (or respondents) gave for the tribe question. There are three special edits for tribe data:

- 1) In some cases, enumerators accidentally wrote place of birth codes (which are also district of enumeration codes) for the tribe response, perhaps mixing up the two questions. This led to an abundance of certain tribes in areas where the tribes would not be expected, for instance Djiboutians in Mt. Elgon.

Edit: If tribe is the same value as birthplace, and is also the same value as the district of enumeration, change the code to blank so that it is later edited as a not stated value.

- 2) Some enumerators wrote 3s that were interpreted as 8s. This led to more foreigners (800+ codes) than would be expected in certain areas.

Edit: If there is only one code in the 800 to 817 series, referred to as 8xx, and there is a matching 3xx code in the household, change the 8xx code to 3xx. Otherwise, if there is only one 8xx code in the household and there are more than two 3xx (300 to 317) codes, change the 8xx code to one of the 3xx codes.

- 3) Some enumerators switched the digits of tribe codes so that expected codes for an area, such as 170 (Luo) in Western Province, would be written as 107.

Edit: For the following codes, if the previous household head has the correct code, or another member of the household has the correct code, change the incorrect code (not a valid tribe code for any tribe in Kenya) to the correct one:

Incorrect Code	Correct Code	Tribe
103	130	Kamba
104	140	Kikuyu
105	150	Kisii
107	170	Luo
201	210	Meru
203	230	Pokot
209	140	Kikuyu

(The 209 district of enumerator is named Kikuyu.)

V. EDIT SPECIFICATIONS FOR THE SHORT FORMS

Overview of Variables Edited

Variable	Is Don't Know a Valid Value?	Applies to Which Short Forms
Sex	No	All four
Age	No	All four
Home District / Country	No	All but emigrants
Tribe / Nationality	Yes	All but emigrants
Education Level	No	All but travellers
Duration of Stay (MM/YY)	Yes	Institutions
Professional Training	Yes	Emigrants
Destination	Yes	Emigrants
Current Residence	Yes	Emigrants
Year of Departure	Yes	Emigrants
Status/Reason	Yes	Emigrants
Remittances Given	Yes	Emigrants
Remittances Amount	No	Emigrants
Remittances Utilization	No	Emigrants

Common Variables

Sex (Emigrants: E03, Institutions: SQ03, Travellers: SQ03, Vagrants: SQ03)

<i>Sex value is...</i>	
Valid (1:2)	Invalid
OK	Impute sex by selecting randomly from the previous 10 valid sex values

Age (Emigrants: E04, Institutions: SQ04, Travellers: SQ04, Vagrants: SQ04)

<i>Age value is...</i>		
Valid (0:95)	Invalid (96:98)	Invalid (else)
OK	Impute age to 95	Impute age by selecting randomly from the previous 10 valid age values

Home District / Country (Institutions: SQ07, Travellers: SQ05, Vagrants: SQ05)

<i>Home District value is...</i>		
Valid (in value set)	Invalid	
OK	<i>Is this a travellers short form or an institution short form with either hotel or prison checked?</i>	
	Yes	No
	Impute home district by selecting randomly from the previous 10 valid home district values	Impute home district by using the province and district codes of enumeration

Tribe / Nationality (Institutions: SQ08, Travellers: SQ06, Vagrants: SQ06)

<i>Tribe value is...</i>	
Valid (in value set)	Invalid
OK	Impute to DK (999)

Education (Emigrants: E05, Institutions: SQ06, Vagrants: SQ07)

If age is less than 3, set the education variable to blank.

<i>Education value is...</i>		
Valid (1:6)		Invalid (9, else)
<i>Is age >= minimum age for level (see table below)?</i>		Impute education using an age-sex hotdeck
Yes	No	
OK	Impute education using an age-sex hotdeck	

Code	Education Level Completed	Minimum age for highest level completed
1	None	0
2	Primary	12
3	Secondary	16
4	Completed Undergraduate	20
5	University Completed Masters/PHD	23
6	Other Tertiary Colleges	18

Recode the education variable to the long form variable P41 using the following codes:

Short Form Education Variable	Long Form (P41) Value
1	97
2	8
3	12
4	18
5	20
6	16

Institution-Only Variables

Duration of Stay (Institutions: SQ05)

- 1) If year is 0 and month is between 12 and 48, set year to month divided by 12 and month to the remainder of month divided by 12.

<i>Duration of stay year is...</i>			
Valid (between 0 and age)		Invalid	
<i>Duration of stay month is...</i>		<i>Duration of stay month is...</i>	
Valid (between 0 and 11)	Invalid	Valid (between 1 and 11)	Invalid
OK	Impute month to 0	Impute year to 0	Impute year and month to 99 if not already

Emigrant-Only Variables**Professional Training (Emigrants: E06)**

- 1) If age is less than 18, then E06 cannot be 1:3,5:8. Set such values to 99 (DK).
- 2) Use the table below to see if the Education x Professional Training combination is valid. If not, set the value to 99 (DK).

		E05: Level of Education Completed					
		None	Primary	Secondary	Under graduate Completed	University postgraduate completed	Other Tertiary Colleges
E06: Professional Training	Doctor	BAD	BAD	BAD	OK	OK	BAD
	Teacher	BAD	BAD	OK	OK	OK	OK
	Nurse	BAD	BAD	BAD	OK	OK	OK
	Artisan	OK	OK	OK	BAD	BAD	OK
	Lecturer	BAD	BAD	BAD	OK	OK	OK
	Engineer	BAD	BAD	BAD	OK	OK	BAD
	Economist	BAD	BAD	BAD	OK	OK	BAD
	Statistician	BAD	BAD	BAD	OK	OK	OK
	Other	OK	OK	OK	OK	OK	OK
	None	OK	OK	OK	BAD	BAD	BAD
	DK	OK	OK	OK	OK	OK	OK
	Else	BAD	BAD	BAD	BAD	BAD	BAD

Destination (Emigrants: E07)

- 1) If E07 and E08 are both within 0:799 (not foreigners), remove the record

<i>Destination value is...</i>		
Valid (in value set with a code >= 800, or 999)	Invalid	
OK	<i>Current residence value is...</i>	
	Valid (in value set with a code >= 800)	Invalid
	Impute destination to current residence	Impute destination to DK (999)

Current Residence (Emigrants: E08)

<i>Current residence value is...</i>		
Valid (in value set with a code >= 800, or 999)	Invalid	
OK	<i>Destination value is...</i>	
	Valid (in value set with a code >= 800)	Invalid
	Impute current residence to destination	Impute current residence to DK (999)

Year of Departure (Emigrants: E09)

<i>Step 1: Year of departure value is...</i>			
95:99	0:9	< 1995	Else
Add 1900 to the year	Add 2000 to the year	Set year to 1995	OK
<i>Step 2: Year of departure value is...</i>			
< birth year	>= birth year and <= 2009	DK (9999)	Else
Impute year to DK (9999)	OK	OK	Impute year to DK (9999)

Status/Reason (Emigrants: E10)

<i>Status value is...</i>	
Valid (1:7,9)	Invalid
OK	Impute status to DK (9)

Remittances Given (Emigrants: E11)

<i>Remittances given is...</i>					
1	2,3		Else		
OK	<i>Is remittances amount greater than 0?</i>		<i>What is the value of remittances amount?</i>		
	Yes	No	> 0	0	Else
	Impute remittances given to Yes (1)	OK	Impute remittances given to Yes (1)	Impute remittances given to No (2)	Impute remittances given to DK (3)

Remittances Amount (Emigrants: E12)

<i>Remittances amount is...</i>		
A Number		Blank
<i>Does remittances given = Yes (1)?</i>		
Yes	No	
OK	Impute remittances amount to blank	

Remittances Utilization (Emigrants: E13)

<i>Remittances given is...</i>			
1		2,3	
<i>Remittance utilization is...</i>		<i>Remittance utilization is...</i>	
Valid (1:5)	Else	Blank	Else
OK	Impute utilization based on hotdeck of remittances amount	OK	Impute utilization to blank

VI. Structural Edits

Reformatting Long Forms

For the iCADE dictionary, use as an ID field the geographic IDs (province, district, division, location, sub location, EA number), EA type and status, household number and type, political units (constituency and ward), structure number, form bar code, and TIF number. This ensures that when CSPro reads in a case, the data read in is only one scanned page.

EDIT 1: Determining When to Combine Pages to Create Continuation Households

1. If Province, District, Division, Location, Sub Location, or EA Number changes, then the form is not a continuation form
2. If Household Number changes and is not blank, then the form is not a continuation form
3. If Batch Number changes, then the form is not a continuation form
4. If fewer than 3 control fields (items at the top of the form in green) are filled out and the serial number is not sequential with the previous form's serial number, then this form is invalid (probably a practice) form and should be removed from the output
5. If (1) through (3) are all false, perform the following checks:
 - a. Is a household head missing on the multiple page and at least one relationship field is not blank?
 - b. Is the maximum line number on the multiple page greater than 10 and the maximum of the line numbers already read in?
 - c. Does the total number of continuation pages on the multiple page match the value on the first page, and is the continuation page number different from the one on the first page?
 - d. Are 75% of sections I through L blank on the multiple page?
 - e. Is the total population value marked on the first page greater than the number of records currently read in?

With true/false values for tests (a) through (e), apply the following weights to the values:

Edit	Conventional	Nonconventional
a. No household head	30%	0%
b. Increasing line numbers	35%	50%
c. Continuation page markers	15%	20%
d. Blank housing information	20%	0%
e. Marked total population not yet read	10%	40%

In each case the weighted total adds up to 110%. For the multiple page, if the weights equal or exceed 50%, consider the multiple page to be a continuation form, otherwise treat the page as

a new household.

EDIT 2: Retrieving Household Information from Multiple Forms

When combining multiple pages on continuation forms for conventional households, initially take the housing data from the first page. Check if any housing data is set on any following page, and if it is, and that information was not already set (on the first or following forms), take the housing data from the following page (but only the housing data for items currently blank). For example, if section I is defined on page 1 and section J is defined on page 2 but not on page 1, take section I from page 1 and section J from page 2.

EDIT 3: Household Type

If Household Type = Conventional or Refugee Camp:

1. If fewer than 25% of the values expected for P10, P13, and P14, and for sections J and K have no values, then set to nonconventional

If Household Type = Nonconventional:

1. If 25% or more of the values expected for P10, P13, and P14, and for sections J and K have values; and 50% or more of the P10, P13, and P14 variables have values; and if at least some of the relationships (other than the head's) are not 11, then set to conventional

If Household Type = Invalid or Missing:

1. If 25% or more of the values expected for P10, P13, and P14, and for sections J and K have values, then set to conventional; otherwise set to nonconventional

If Edited Household Type = Nonconventional:

1. Set all fields of P10, P13, P14, sections G, H, J, K (and L) to blank

EDIT 4: Dropping Deleted Person Records

When processing person records, check whether or not there was an X written across the line number on either the front or back page of any given form. If an X was written on the front page, drop that person record. If an X was written on the back page but not the front page, keep the record and assign 0 as the line number.

EDIT 5: Deleting Person Records Due to Lack of Information

If the number of records on the form equals the number listed in Total Household Population, skip this edit. Otherwise, delete any person record if none of the values for relationship, sex or age is defined. If only one of the three variables is defined, check whether for defined values for tribe, religion, marital status, and birth place. If two or more of these variables is defined, keep the person record; otherwise delete it.

EDIT 6: Deleting Death Records Due to Lack of Information

Keep only death records where two of the four possible values (death notification, age, sex, maternal mortality) are defined.

EDIT 7: Deleting Empty Households

If a case has no person records (usually due to the deletion of Xs or lack of information), delete that case from the output file.

LATER EDIT: Geographic IDs

A program run after the structural edits will replace the constituency, ward, EA type, and EA status values with accurate values taken from the master database of geographic information. At this time an area size variable will also be inserted into the data file.

Reformatting Short Forms

A lot of the complexity with the long form edit occurs when multiple forms are concatenated to form one household. KNBS decided that for the purposes of the short forms each “case” would constitute the data on only one questionnaire page (the front and back sides), rather than implementing code to concatenate large institutions, emigrants, vagrants, or travelers into one case. (A case is the lowest level of data that CSPro processes, and constitutes a collection of person records. In the case of conventional households, each case is a different household.) This approach will not affect population figures.

The downside of this approach occurs if an analyst decides to tabulate data on a case level. If an analyst wants to know, for example, the average number of emigrants per household, or residents per hospital, this figure will not be accurate because continuation forms are ignored in these structural edits. For example, a hospital with 50 patients was enumerated on three Institution sheets, with 20, 20, and 10 person records on each form. Instead of creating one case with 50 patients, these structural edits will create three hospitals, two with 20 patients and one with 10 patients. KNBS decided that such tabulations are unlikely and do not justify added coding complexity.

The CSPro dictionary for the short forms is structured for compatibility with the long form. Variables that are common across the five forms, such as sex and age, are positioned in the short form data file so that they line up with those same variables in the long form data file. This means that tabulations for the entire country (for those key variables) can be performed on one large data file that contains the data from both the short and long forms. The IDs of the short forms must thus be made compatible with the IDs for the long forms. The long forms have more ID fields than the short forms (save for the emigrants form), so a blank household number is added to the short forms, and the household type is set to specify the kind of short form: vagrants (6), travelers (7), and institutions (8).

EDIT 1: Dropping Deleted (“X”) Person Records

If an X is written across the line number for a given person record, examine the line number of the previous record. If the previous line number is not 99, delete the current person record. This exception comes about because some enumerators wrote X0 for a line number because there were not enough spaces to write 100.

EDIT 2: Deleting Person Records Due to Lack of Information

If neither sex nor age is defined for a person record, delete that person record.

EDIT 3: Deleting Empty Households

If a case has no person records (usually due to the deletion of Xs or lack of information), delete that case from the output file.