

Департамент по экономическим и социальным вопросам
Статистический отдел

Методологические исследования

Серия F № 82

Руководство по редактированию данных переписи населения и жилого фонда



**Организация Объединенных Наций
Нью-Йорк, 2002 год**

ПРИМЕЧАНИЕ

Используемые в настоящем издании определения и представление материалов не подразумевают выражения какого-либо мнения со стороны Секретариата Организации Объединенных Наций в отношении правового статуса той или иной страны, территории, города или района или их полномочных органов, или же делимитации их границ или установления их пределов.

Используемый в настоящем издании термин «страна» в некоторых случаях относится также к территориям или районам.

Определения «развитые», «развивающиеся» и «наименее развитые», применяемые к странам, районам или регионам, предназначены для статистических целей и не обязательно отражают оценку уровня развития, достигнутого той или иной страной или районом.

Условные обозначения документов Организации Объединенных Наций состоят из прописных букв и цифр.

ST/ESA/STAT/SER.F/82

Издание Организации Объединенных Наций
В продаже под № R.00.XVII.9

Авторское право © Организация Объединенных Наций, 2002 год
Все права сохраняются
Отпечатано Секцией размножения документов
Организации Объединенных Наций

ПРЕДИСЛОВИЕ

На протяжении многих лет Организация Объединенных Наций выпускала серии руководств и технических докладов, предназначенных для оказания помощи странам в планировании и проведении эффективных с точки зрения затрат переписей населения и жилого фонда по усовершенствованной методике. Эти руководства и доклады время от времени пересматриваются и обновляются, с тем чтобы отражать новые тенденции и проблемы, возникающие при проведении переписей. Настоящее издание входит в серию руководств, разработанных с целью оказания помощи странам в подготовке переписей в 2000 году и в дальнейшем. Другие руководства данной серии:

- a) *Руководство по управлению переписями населения и жилого фонда*, Серия F, № 83 (издание Организации Объединенных Наций, в продаже под № R.00.XVII.15);
- b) *Руководство по географическим информационным системам и цифровому картированию*, Серия F, № 79 (издание Организации Объединенных Наций, в продаже под № R.00.XVII.12);
- c) *Guide for the collection of Economic Characteristics in Population Censuses* (готовится к изданию).

В издании *Принципы и рекомендации в отношении переписей населения и жилого фонда, первое пересмотренное издание* (Организация Объединенных Наций, 1998 год), рассматриваются достоинства системы контроля и повышения качества на первоначальном этапе переписи, который является крайне важным с точки зрения успеха всей работы по проведению переписи, и значение плана редактирования, который должен разрабатываться как часть всей программы проведения переписи и объединяться с другими планами и процедурами переписи. Для пользователей *Руководства по редактированию данных переписей населения и жилого фонда* будут особенно полезны *Принципы и рекомендации*, где содержится значительная доля исходной информации относительно процедур редактирования, изложенных в главах III, IV и V.

Цель настоящего издания заключается в том, чтобы дать странам широкое представление о методологии редактирования данных переписей и обследований и предоставить информацию для тех, кто этим занимается, о применении различных подходов к редактированию данных переписи. Еще одна цель состоит в том, чтобы страны сохраняли накопленный ими опыт редактирования, способствовали общению между специалистами,

занимающимися содержанием переписи, и специалистами по обработке данных и документально отражали деятельность, осуществляемую в ходе переписи или обследования с тем, чтобы избежать дублирования в работе во время следующей переписи или обследования.

В *Руководстве* рассматриваются преимущества и недостатки ручного и автоматизированного редактирования. При крупных переписях корректирование вручную почти невозможно с экономической точки зрения. Условия для такого корректирования, как правило, указываются в специально разработанных компьютерных программах, с помощью которых проводится автоматический поиск ошибок и интерполяция на основе других сведений о данном лице или домохозяйстве, или о других лицах или домохозяйствах. В основном *Руководство* посвящено автоматическому исправлению данных.

Редактирование с помощью компьютера играет важную роль в обнаружении и исправлении ошибок. На этапе компьютерного редактирования подробная проверка на согласованность может проводиться в консультации со специалистами, занимающимися содержанием переписи. Обнаруженные ошибки можно исправить или путем ссылки на первоначальные вопросники или автоматически. Автоматическое редактирование ускоряет обработку данных, но вместе с тем требует тщательного контроля за качеством поступающих данных.

Издание состоит из введения и пяти глав. Во введении описывается процесс проведения переписи и различные виды ошибок, которые при этом возникают. Глава I посвящена основам редактирования данных переписи. В главах II–V говорится о процедурах и способах редактирования данных переписи на различных этапах обработки данных. Соображения технического порядка, в частности, те, которые касаются программирования, изложены в приложениях.

Хотя настоящее *Руководство* посвящено, главным образом, редактированию данных переписей населения и жилого фонда, многие понятия и методы применимы также к проведению обследований.

Особая признательность выражается Бюро переписи населения США в лице г-на Майкла Дж. Левина за вклад в подготовку проекта данного *Руководства*. Благодарность выражается также многим специалистам, занимающимся содержанием переписи, и программистам с богатым опытом проведения переписей и обследований из всех регионов земного шара, которые рецензировали и помогли завершить настоящее издание.

Сокращения

ГИС	географические информационные системы
МКДНИ	Международная классификация дефектов, нетрудоспособности и инвалидности
МОТ	Международная организация труда
НМИ	Новая методология интерполяции
ОРМ	оптическое распознавание меток
ОРС	оптическое распознавание символов
ПРО	пострегистрационное обследование
СПОН	статистический пакет для общественных наук
ССА	система статистического анализа
СУ	счетный участок
ЮНЕСКО	Организация Объединенных Наций по вопросам образования, науки и культуры

СОДЕРЖАНИЕ

	<i>Стр.</i>
ПРЕДИСЛОВИЕ	iii
ВВЕДЕНИЕ	1
A. Цель <i>Руководства</i>	1
B. Процесс переписи	1
C. Ошибки в процессе переписи	2
1. Ошибки в охвате	2
2. Ошибки в содержании	2
D. Построение <i>Руководства</i>	4
I. РЕДАКТИРОВАНИЕ ПРИ ПРОВЕДЕНИИ ПЕРЕПИСЕЙ И ОБСЛЕДОВАНИЙ	5
A. Редактирование: историческая справка	5
B. Группа редактирования	6
C. Практика редактирования: отредактированные и неотредактированные данные	6
D. Основы редактирования	9
1. Как чрезмерное редактирование оказывает пагубное воздействие	10
2. Обработка неизвестных величин	11
3. Кажущиеся изменения	11
4. Определение уровня допустимых отклонений	12
5. Извлечение уроков из процесса редактирования	12
6. Гарантия качества	12
7. Затраты на редактирование	12
8. Интерполяция	13
9. Архивирование	13
II. ПРИКЛАДНЫЕ ПРОГРАММЫ РЕДАКТИРОВАНИЯ	15
A. Ручное и автоматизированное корректирование	15
B. Руководящие принципы исправления данных	17
C. Проверка на достоверность и согласованность	20
1. Нисходящий принцип редактирования	21
2. Редактирование по принципу групповых переменных	21
D. Вопросы кодирования	24

	Стр.
E. Методы исправления и интерполяции данных	25
1. Метод статической интерполяции (<i>cold deck</i>)	25
2. Метод динамической интерполяции (<i>hot deck</i>)	25
3. Вопросы динамической интерполяции (<i>hot deck</i>)	28
4. Проверка матриц интерполяции	35
5. Признаки интерполяции	38
F. Другие системы редактирования	39
III. РЕДАКТИРОВАНИЕ СТРУКТУРЫ	41
A. Редактирование географических параметров	41
1. Местоположение жилых помещений (населенный пункт)	41
2. Городское и сельское население	42
B. Проверка охвата	42
1. Регистрация де-факто и де-юре	42
2. Иерархия домохозяйств и жилищных единиц	42
3. Фрагменты переписных листов	42
C. Структура записей о жилом фонде.....	43
D. Соответствие между записями о населении и записями о жилом фонде	43
1. Свободные и занятые жилища	43
2. Двойной учет домохозяйств и жилищных единиц	44
3. Недостающие домохозяйства и жилищные единицы	44
4. Соответствие между числом жильцов и суммарным количеством жильцов	44
5. Соответствие между жильцами и типом строения/домохозяйства	45
E. Двойные записи	45
F. Специальные совокупности	45
1. Коллективное проживание	45
2. Жильцы без жилья	46
G. Определение главы домохозяйства и супруга(и)	47
1. Редактирование переменной, касающейся главы домохозяйства	47
2. Редактирование данных о супруге	50
H. Возраст и дата рождения	50
1. Когда есть дата рождения, но не указан возраст	50
2. Когда возраст и дата рождения не согласуются между собой	50
I. Учет недостоверных данных	50

	<i>Стр.</i>
IV. РЕДАКТИРОВАНИЕ ДАННЫХ ПЕРЕПИСИ НАСЕЛЕНИЯ	51
A. Демографические характеристики	52
1. Родство (P2A)	52
2. Пол (P3A)	54
3. Дата рождения и возраст (P3B)	55
4. Семейное положение (P3C)	59
5. Возраст вступления в первый брак (P4F)	61
6. Рождаемость: дети, родившиеся живыми (P4A) и выжившие дети (P4B)	61
7. Рождаемость: дата рождения последнего ребенка (P4C)	69
8. Рождаемость: возраст при рождении первого ребенка (P4G)	69
9. Смертность (P4D)	70
10. Сиротство в случае смерти матери/отца (P4E) и номер строки матери	70
B. Характеристики миграции	71
1. Место рождения (P1C)	71
2. Гражданство (P3D)	73
3. Продолжительность проживания (P1D)	74
4. Прежнее место жительства (P1E)	76
5. Место жительства в определенный момент в прошлом (P1F)	77
C. Социальные характеристики	77
1. Способность читать и писать (грамотность) (P5A)	77
2. Посещение учебных заведений (P5B)	78
3. Образовательная подготовка (последняя завершенная ступень или уровень) (P5C)	79
4. Область знаний и свидетельства об образовании (P5D)	79
5. Вероисповедание (P3E)	79
6. Язык (P3F)	80
7. Этническая принадлежность (P3G)	81
8. Нетрудоспособность (P8A)	82
9. Дефекты и инвалидность (P8B)	83
10. Причины нетрудоспособности (P8C)	83
D. Экономические характеристики	83
1. Статус экономической активности (P6A)	84
2. Отработанное время (P6B)	87
3. Род занятий (P6C)	87
4. Отрасль (P6D)	87

	<i>Стр.</i>
5. Статус занятости (P6E)	88
6. Доход (P6F)	88
7. Институциональный сектор (P6G)	89
8. Место работы (P6H)	89
V. РЕДАКТИРОВАНИЕ ДАННЫХ О ЖИЛОМ ФОНДЕ	91
A. Основные вопросы	92
1. Строение: описание строения (H01)	92
2. Строение: строительный материал наружных стен (H02)	92
3. Строение: год или период постройки (H03)	93
4. Жилые помещения: местонахождение жилых помещений (H04)	93
5. Жилые помещения: тип жилых помещений (H05)	93
6. Жилые помещения: сведения о занятости (H06)	94
7. Жилые помещения: форма владения (H07)	94
8. Жилые помещения: число комнат (H08)	95
9. Жилые помещения: площадь помещения (H09)	95
10. Жилые помещения: система водоснабжения (H10)	95
11. Жилые помещения: туалет и канализационное оборудование (H11)	96
12. Жилые помещения: ванное оборудование (H12)	97
13. Жилые помещения: кухонное оборудование (H13)	97
14. Жилые помещения: освещение (H14)	98
15. Жилые помещения: удаление мусора (H15)	98
16. Жилые помещения: занятость более чем одним домохозяйством (H16)	98
17. Жилые помещения: число жильцов (H17)	98
18. Жильцы: характеристики главы домохозяйства (H18)	99
19. Жильцы: пользование жильем (H19)	99
20. Жильцы: арендная плата за жилье и жилищные расходы владельца (H20)	99
B. Дополнительные вопросы	99
1. Строение: число жилищ (A01)	99
2. Строение: наличие лифта (A02)	99
3. Строение: сельскохозяйственное назначение (A03)	100
4. Строение: строительный материал крыши (A04a)	100
5. Строение: строительный материал для пола (A046b)	100
6. Строение: эксплуатационное состояние (A05)	100
7. Жилые помещения: число спален (A06)	101
8. Жилые помещения: топливо, используемое для приготовления пищи (A07)	101

	<i>Стр.</i>
9. Жилые помещения: вид отопления и используемая энергия (A08)	102
10. Жилые помещения: наличие горячей воды (A09)	102
11. Жилые помещения: магистральный газ (A10)	102
12. Жилые помещения: наличие телефона (A11)	103
13. Жилые помещения: использование жилищной единицы (A12)	103
14. Жильцы: число автомобилей (A13)	103
15. Жильцы: предметы длительного пользования (A14)	103
16. Жильцы: наличие пространства вне помещения в пользовании домохозяйства (A15)	104
C. Занятые и незанятые жилищные единицы	104
Приложение I. Производные переменные	107
A. Производные переменные для записей о жилом фонде	107
1. Доход домохозяйства	107
2. Семейный доход	107
3. Тип семьи	108
4. Лица, связанные родственными отношениями	108
5. Работники в семье	109
6. Комплексное сантехническое оборудование	109
7. Комплексное кухонное оборудование	109
8. Валовая квартирная плата	110
B. Производные переменные для демографических записей	110
1. Статус экономической активности	110
2. Число субсемей и родственники в субсемье	110
3. Собственные дети	112
4. Родители в доме	112
5. Текущий год в учебном заведении	112
Приложение II. Взаимосвязь между форматом переписного листа и вводом данных с клавиатуры	115
Приложение III. Ввод данных с клавиатуры	119
A. Ввод данных	119
1. Сканирование	119
2. Реверсивный ввод данных	120
3. Интерактивный ввод данных с клавиатуры	121
B. Проверка инструкций по вводу данных с клавиатуры	121

	<i>Стр.</i>
С. Проверка	122
1. Зависимая проверка	122
2. Независимая проверка	122
Приложение IV. Выборочные рабочие блок-схемы	123
Приложение V. Методы интерполяции	129
Приложение VI. Пакеты компьютерного редактирования	133
ГЛОССАРИЙ	135
БИБЛИОГРАФИЯ	141

ТАБЛИЦЫ

	<i>Стр.</i>
Таблица 1. Выборка населения с разбивкой по возрастным группам с интервалом 15 лет и полу с использованием неотредактированных и отредактированных данных	7
Таблица 2. Население и изменения в его составе по возрастным группам с интервалом 15 лет с неизвестными величинами: 1990 и 2000 годы	8
Таблица 3. Население и изменения в его составе по возрастным группам с интервалом 15 лет без неизвестных данных: 1990 и 2000 годы	8

РИСУНКИ

	<i>Стр.</i>
Рисунок 1. Типичное гипотетическое домохозяйство с указанием родства, пола и данных о рождаемости его членов	18
Рисунок 2. Пример домашнего хозяйства, где глава и супруг одного пола	19
Рисунок 3. Пример домашнего хозяйства с указанием возраста некоторых членов домохозяйства	19
Рисунок 4. Пример домашнего хозяйства с возможной несогласованностью указанных возрастных данных	20
Рисунок 5. Пример правил для редактирования отдельных характеристик населения по принципу групповых переменных	22
Рисунок 6. Пример с главой и супругом одного пола в неотредактированном наборе данных и его разрешение	22
Рисунок 7. Выборочные спецификации редактирования для исправления переменной пола с символическим кодом	23
Рисунок 8. Пример анализа данных об очень молодой вдове с тремя детьми, отредактированных по принципу групповых переменных	23

	<i>Стр.</i>
Рисунок 9. Примеры общих кодов для отдельных элементов	24
Рисунок 10. Выборочные домохозяйства в качестве входных данных для динамической интерполяции	26
Рисунок 11. Исходная статическая матрица для возраста, исчисляемого по полу и степени родства	27
Рисунок 12. Пример матрицы динамической интерполяции после одного изменения	27
Рисунок 13. Пример матрицы динамической интерполяции после многократных изменений ...	28
Рисунок 14. Пример главы домохозяйства и отца главы без указания языка	30
Рисунок 15. Исходные значения для матрицы, предназначенные для динамической интерполяции языка	30
Рисунок 16. Пример членов домохозяйства без указания языка	31
Рисунок 17. Пример главы домохозяйства с детьми без указания возраста ребенка	33
Рисунок 18. Пример главы домохозяйства с ребенком при отсутствии данных о возрасте и уровне образования ребенка	33
Рисунок 19. Выборочный набор величин для статической матрицы и выборочный код интерполяции	35
Рисунок 20. Пример сводного отчета для ряда интерполяций на одну ошибку	36
Рисунок 21. Выборочные сообщения об ошибках в переписном листе	36
Рисунок 22. Пример дополнительной распечатки ошибок по переписному листу, включая множественные переменные	36
Рисунок 23. Выборочные демографические записи с признаками для исчисляемых значений ..	38
Рисунок 24. Пример признака для молодой женщины с пробелом в графе «рождаемость» и добавлением признака	39
Рисунок 25. Пример домохозяйства, где глава домохозяйства зарегистрирован как первое лицо	48
Рисунок 26. Пример домохозяйства, где глава домохозяйства зарегистрирован как пятое лицо ...	48
Рисунок 27. Иллюстрация домохозяйства с информацией о рождаемости	64
Рисунок 28. Первичные значения для определения детей, живущих на момент переписи, когда возраст и количество детей, родившихся живыми, достоверны	65
Рисунок 29. Образец матрицы интерполяции, которую необходимо разработать для пар известных элементов данных	68

РАМКИ

	<i>Стр.</i>
Рамка 1. Для чего нужно редактировать данные переписи	6
Рамка 2. Основные руководящие принципы исправления данных	17
Рамка 3. Руководящие принципы редактирования структуры	41
Рамка 4. Редактирование и исчисление возраста	56

РИСУНКИ ПРИЛОЖЕНИЙ

	<i>Стр.</i>
Рисунок А.І.1. Расширенная семья	111
Рисунок А.І.2. Выборочное домохозяйство с двумя субсемьями	112
Рисунок А.ІІ.1. Выборочный бланк переписного листа с персональными страницами	115
Рисунок А.ІІ.2. Пример потока в переписном листе с персональными страницами	115
Рисунок А.ІІ.3. Выборочный переписной лист, страница, посвященная домохозяйству, со всеми его членами на одной странице	116
Рисунок А.ІІ.4. Пример последовательности информации для переписного листа со страницами, посвященными домохозяйству, где на одной странице содержатся данные о всех членах домохозяйства	117
Рисунок А.ІІ.5. Пример страницы, посвященной домохозяйству, где указаны все лица, без проблем, связанных с вводом данных	118
Рисунок А.ІІ.6. Пример страницы, посвященной домохозяйству, где указаны все лица с возможными проблемами ввода данных	118
Рисунок А.ІV.1. Выборочная блок-схема для определения главы домохозяйства (ГД)	124
Рисунок А.ІV.2. Выборочная блок-схема для определения наличия супруга(и) в домохозяйстве .	125
Рисунок А.ІV.3. Выборочная блок-схема для редактирования переменной для пола главы домохозяйства и супруга(и)	127

ВВЕДЕНИЕ

А. ЦЕЛЬ РУКОВОДСТВА

1. Хорошо спланированная перепись или обследование с минимальными ошибками в конечной продукции является бесценным ресурсом для государства. Для получения точных результатов переписи или обследования, данные должны быть в максимально возможной степени свободны от ошибок и несоответствий, особенно после этапа обработки данных. Процедура обнаружения ошибок в записях данных и между ними во время и после сбора данных и корректирования отдельных элементов известна как редактирование данных переписи населения и жилого фонда.

2. Абсолютно совершенных данных переписи или обследования не бывает. Страны уже давно признали наличие проблем в связи с данными переписей и обследований и приняли различные подходы к тому, как обращаться с пробелами в данных и непоследовательными ответами. Вместе с тем из-за длительных перерывов между переписями, процедуры редактирования данных зачастую должным образом не документируются. В результате при планировании новой переписи или обследования страны вынуждены вновь придумать то, что уже использовалось во время предыдущей работы по сбору данных.

3. *Руководство по редактированию данных переписи населения и жилого фонда* призвано заполнить пробел в знаниях в области методологии редактирования данных переписей и обследований и предоставить информацию соответствующим работникам о применении различных подходов к редактированию данных переписи. Данное *Руководство* преследует также цель содействовать тому, чтобы страны сохраняли накопленный ими опыт редактирования, укрепляли связь между специалистами, занимающимися содержанием переписи, и специалистами по обработке данных и документально отражали то, что делается в ходе текущей переписи населения или обследования во избежание дублирования работы в будущем.

4. Настоящее *Руководство* является справочным пособием как для специалистов, занимающихся содержанием переписи¹, так и для специалистов по обработке данных, поскольку они ведут совместную работу по

¹ По определению, содержащемуся в данном *Руководстве*, специалисты, занимающиеся содержанием переписи, — это те, кто работает в области народонаселения, жилья и других смежных отраслях.

разработке спецификаций редактирования и программ для переписей и обследований. Оно построено по принципу «поваренной книги», которая позволяет странам использовать такие приемы редактирования, которые наиболее приемлемы с точки зрения сложившейся у них статистической ситуации. Настоящее издание предназначено также для того, чтобы содействовать более тесному общению между специалистами в разработке и осуществлении программы редактирования.

5. Во введении описывается процесс переписи, приводятся различные виды ошибок, которые возникают во время переписи, и основы редактирования данных. В последующих главах рассказывается о процедурах и методах редактирования данных переписи на различных этапах обработки. Хотя настоящее *Руководство* посвящено главным образом редактированию результатов переписи населения и жилого фонда, многие из этих понятий и методов можно применять и при проведении обследований.

В. ПРОЦЕСС ПЕРЕПИСИ

6. Перепись населения и/или жилого фонда — это весь процесс сбора, обработки, оценки, анализа и опубликования демографических, экономических, социальных данных и данных о жилом фонде в отношении всех лиц и жилых помещений (Организация Объединенных Наций, 1998 год). Перепись проводится в конкретное время по всей стране или в ее четко обозначенной части. Перепись как таковая отображает положение населения и состояние жилого фонда на определенную дату.

7. Главная цель переписи заключается в том, чтобы предоставить информацию о численности, распределении и характеристиках населения страны. Данные переписи используются при выработке политики, планировании и управлении, а также при ведении и оценке программ в области образования, трудовых ресурсов, планировании семьи, жилья, здравоохранения, транспорта и развития сельских районов. В административных целях данные переписи находят свое основное применение в демаркации избирательных округов и квотировании представительства в органах управления. Перепись является также неоценимым источником для проведения научных исследований, представляя данные для научного анализа состава и распределения населе-

ния и для статистических моделей прогнозирования будущего роста. Перепись обеспечивает деловые круги и промышленность основными данными, необходимыми им для оценки спроса на жилье, учебные заведения, предметы обстановки, продукты питания, одежду, места отдыха и развлечений, предметы медицинского назначения и другие товары и услуги.

8. Основные особенности у всех переписей и обследований одни и те же и включают в себя: *a)* подготовительную работу; *b)* регистрацию; *c)* обработку данных, в том числе ввод данных с клавиатуры, редактирование и табулирование; *d)* создание базы данных или распространение результатов; *e)* оценку результатов; и *f)* анализ результатов.

9. Подготовительная работа содержит в себе множество элементов, таких как определение правовой основы для проведения переписи; составление бюджета; разработка календарного плана; административная организация; картография; составление перечня жилых единиц; разработка программы табулирования; подготовка переписного листа; и разработка планов, и подготовка работников для проведения регистрации, предварительных проверок, обработки данных и их распространения.

10. Процесс регистрации зависит от избранного метода, сроков и длительности периода регистрации, уровня инспектирования и от того, используется ли и как используется выборка. После сбора данных их необходимо закодировать, обработать, отредактировать и свести в таблицы. Во время обработки данных создаются микро и макробазы данных. Национальные бюро переписи/статистические управления используют эти базы данных для составления таблиц, анализа динамических рядов, составления диаграмм и картограмм и в географических информационных системах (ГИС) для тематического картирования и других методов распространения данных. Результаты оцениваются как с точки зрения содержания, так и с точки зрения охвата с использованием целого ряда методов, включая демографический анализ и контрольные обследования после регистрации. И, наконец, результаты анализируются с использованием самых различных методов, включая дескриптивные сводки результатов, тактический анализ результатов переписи и детализированные аналитические исследования одного или более аспектов демографической и социальной ситуации в стране.

С. ОШИБКИ В ПРОЦЕССЕ ПЕРЕПИСИ

11. Данные переписи изобилуют многими ошибками, которые в целом можно классифицировать как ошибки в охвате и ошибки в содержании.

1. Ошибки в охвате

12. Ошибки в охвате возникают в результате пропусков или двойных записей о лицах или жилищных единицах при регистрации. Источниками ошибок в охвате являются, в частности, неполные или неточные карты или списки счетных участков, неспособность счетчиков охватить все единицы на своих участках, двойной учет, пропуск лиц, которые не жалеют регистрироваться, ошибочный учет некоторых категорий лиц, таких как гости или иностранцы из числа нерезидентов и утрата или уничтожение записей переписи после регистрации. Ошибки в охвате следует исправлять, насколько это возможно, на месте. В ходе редактирования в управлении фактические двойные записи устраняются. Однако при этом необходимо определить, зарегистрированы ли эти лица или домохозяйства повторно. Например, о близнецах может быть одинаковая информация за исключением порядкового номера. Следовательно, применяемые в ходе этой работы правила редактирования определяют, когда учитывать, когда исключать якобы дублированную информацию, а когда вносить изменения путем интерполяции.

13. При структурном редактировании, о котором говорится в главе III, проводится проверка домохозяйств на правильность указания числа лиц и порядковых номеров и на наличие дважды зарегистрированных лиц.

2. Ошибки в содержании

14. Ошибки в содержании возникают в результате неправильной отчетности или регистрации характеристик лиц, домохозяйств и жилищных единиц. Ошибки в содержании могут быть вызваны плохо сформулированными вопросами и неудачной последовательностью вопросов или слабым общением между респондентом и счетчиком, а также ошибками в кодировании и вводе данных, ошибками в ручном и компьютерном редактировании и ошибочном табулировании результатов. На каждом этапе данного процесса необходимо должным образом вести и сохранять журнал редактирования (известный так же как контрольный журнал) с тем, чтобы исключить потерю данных. В следующих разделах дается разъяснение каждой из вышеупомянутых ошибок.

а) Ошибки в разработке переписного листа

15. Одним из источников ошибок в содержании являются плохо сформулированные вопросы или инструкции. Вид переписного листа, его формат и точность формулировок или расположения статей вопросника заслуживают самого тщательного рассмотрения, поскольку недостатки слабо составленного переписного

листа нельзя устранить ни во время регистрации, ни после нее. Для того чтобы свести к минимуму ошибки, которые могут возникать из-за плохо составленных вопросов, следует использовать предварительное тестирование. Если, например, прогоняемые элементы указаны нечетко или расположены не на месте, счетчик может по ошибке пропустить разделы вопросника и не собрать всю необходимую информацию.

б) *Ошибки счетчика*

16. Счетчики и респонденты взаимодействуют между собой, если только перепись не проводится с использованием переписных листов, заполняемых самими опрашиваемыми. Счетчик может ошибиться, задавая вопросы и при этом сокращая или изменяя формулировку вопросов или не полностью разъясняя респонденту значение вопросов. Счетчик может также вносить ошибки при записи ответов. Качественный уровень счетчиков и их подготовки оказывает решающее влияние на качество собираемых данных. Счетчики должны быть надлежащим образом подготовлены по всем аспектам проведения переписи. Они должны понимать, почему столь важна их роль в процессе переписи и как регистрация сочетается с другими этапами переписи. Кроме того, поскольку счетчики могут обладать самой разной квалификацией и иметь разные уровни образования, подготовка должна вестись таким образом, чтобы добиться от счетчиков умения задавать вопросы так, чтобы получить правильный ответ.

с) *Ошибки респондентов*

17. Ошибки могут вноситься в данные, если респонденты неправильно понимают некоторые вопросы. Ошибки могут возникать также в результате преднамеренного сообщения неправильных сведений или когда ответы на вопросы даются через доверенное лицо (когда ответы на вопросы дает не тот, кого касается запрашиваемая информация). Качество отдельных ответов можно улучшить, рекламируя перепись и занимаясь подготовкой счетчиков с тем, чтобы разъяснять им цель переписи и основания для различных вопросов. Некоторые страны используют переписные листы, заполняемые опрашиваемыми лицами, поэтому взаимодействия между счетчиком и респондентом не происходит. Если бланки заполняются самими опрашиваемыми лицами, ошибки возникают, если респонденты неправильно понимают вопросы или инструкции.

18. Ошибками респондента и счетчика лучше всего заниматься на этапе регистрации, когда еще доступны и бланки, и респонденты, и счетчики. Инспекторы должны уметь готовить счетчиков. Инспекторы долж-

ны также уметь проверять данные, собранные счетчиками, и делать это регулярно во время регистрации с тем, чтобы счетчики не вносили систематические погрешности в данные. Инспекторы должны исправлять ошибки счетчика и респондента на месте до того, как переписные листы будут направлены в региональные или центральные бюро.

д) *Ошибки кодирования*

19. Ошибки могут возникать в ходе кодирования, поскольку кодировщик может неправильно закодировать информацию. Во время ввода данных ошибки могут вноситься в результате неправильного использования клавиатуры. В целом, отсутствие контроля и проверки на этом этапе задерживает публикацию данных, поскольку на более позднем этапе обнаружить и исправить ошибки становится сложнее. Зачастую до или во время кодирования ведется ручное редактирование.

е) *Ошибки при вводе данных*

20. Диапазонные проверки и некоторые основные проверки на согласованность можно встроить в программное обеспечение ввода данных для предотвращения недостоверных вводов. Интеллектуальная система ввода данных гарантирует, что величина каждого поля или элемента данных будет в пределах допустимого диапазона величин для этого элемента. С такой системой увеличиваются шансы на то, что оператор по вводу данных будет набирать допустимые данные и снимать часть нагрузки, связанной с редактированием данных на более поздних этапах процесса подготовки сведений. Вместе с тем эти проверки могут замедлить скорость ввода данных. В связи с этим во время ввода данных необходимо тщательно взвесить объем проверок на согласованность с учетом необходимости поддержания приемлемой скорости ввода данных. Для этого необходимо заранее добиться определенного баланса с тем, чтобы операторы по вводу данных не тратили слишком много времени на эту работу. Проверка набора данных неизбежно улучшает их качество. Набранные бланки можно проверить путем повторного ввода с клавиатуры той же информации, зачастую на выборочной основе.

ф) *Ошибки при компьютерном редактировании*

21. Одним из важнейших этапов обработки данных переписи является редактирование. В процессе редактирования меняются или корректируются недостоверные или несогласующиеся данные путем исчисления отсутствующих ответов или несогласующейся инфор-

мации на основе достоверных данных. Как это ни парадоксально, любая из этих операций в процессе редактирования может повлечь за собой новые ошибки.

г) Ошибки при табулировании

22. Ошибки могут возникать на этапе табулирования из-за ошибок в обработке данных или использования «неизвестной» (не предоставленной) информации. На этом этапе ошибки трудно исправлять, не делая при этом новых ошибок. Отсутствие перекрестной проверки таблиц и ошибки при печатании ведут к ошибкам на этапе публикации. Вместо того чтобы пытаться исправлять сами таблицы, необходимо применять систему обработки таким образом, чтобы дополнительное редактирование производилось при появлении в таблицах несоответствий. Если ошибки пронесются через все этапы работы вплоть до публикации, они станут очевидными и тогда ценность результатов будет сомнительной. Прежде чем выпускать таблицы, необходимо их тщательно проверить, чтобы убедиться в том, что все запланированные таблицы подготовлены для всех намеченных территориальных единиц. Хотя на этапе редактирования диапазонная проверка и проверка на согласованность позволяет устранить большинство ошибок, после табулирования необходимо провести комплексную проверку. Подготовленные и опытные сотрудники должны просмотреть различные таблицы с тем, чтобы проверить, соответствуют ли указанные цифры в различных клетках известной ситуации на месте. В ограниченном количестве случаев при быстром просмотре переписных листов можно обнаружить ошибки, связанные с кодированием. Может оказаться полезным и расчет отдельных отношений и темпов роста, а также сопоставление с показателями предыдущей переписи и другими показателями, опубликованными после выборочных обследований. Однако сравнение с другими показателями, основанными на результатах обследования, можно проводить лишь в том случае, если используемые понятия сопоставимы. В случае обнаружения ошибок в окончательных таблицах исправления следует вносить, прежде всего, в набор данных.

23. Тем, кто занимается обработкой данных, очень важно избегать изменения программ табулирования для устранения проблем в наборе данных. Эти изменения не проявятся на уровне микроданных и поэтому не воспроизведутся при разработке и использовании дру-

гих программ. Сотрудники должны вносить все изменения в набор микроданных, в частности, для того чтобы те, кто занимается обработкой данных в национальных бюро переписи/статистических управлениях, могли строить сопоставимые таблицы. Кроме того, поскольку национальные бюро переписи/статистические управления иногда предоставляют частями файлы микроданных ученым и другим пользователям в государственном и частном секторах, таблицы должны быть воспроизводимы.

24. Как указывалось выше процесс переписи состоит из целого ряда последовательных, взаимосвязанных операций и ошибки могут возникать при каждой операции. Важно не забывать о том, что компьютерное редактирование является частью системы обратной связи и не только направлено на построение таблиц, но и имеет обратную связь в направлении сбора данных и обработки полей. Лучше всего национальное бюро переписи/статистическое управление может избежать проблем, связанных с автоматизированным редактированием, путем извлечения максимальной пользы из редактирования полей. Национальное бюро переписи/статистическое управление должно также быть уверено в том, что кодирование и ввод данных произведены точно, и должно иметь постоянную обратную связь по всем операциям, включая ввод данных, редактирование, и построение таблиц.

D. ПОСТРОЕНИЕ РУКОВОДСТВА

25. В главе I рассматривается роль редактирования в проведении переписей и обследований. Другие главы посвящены конкретным вопросам. В главе II говорится о прикладном назначении редактирования и интерполяции. Глава III посвящена редактированию структуры, которое одновременно охватывает жилищные и демографические единицы, а также определенные процедуры, способствующие редактированию остальных элементов, например, уточнению вопроса о том, есть ли один и только один глава домохозяйства. В главе IV рассматривается вопрос редактирования элементов, касающихся населения, а в главе V — редактирование жилищных единиц. И, наконец, ряд приложений посвящен конкретным вопросам, связанным с редактированием и интерполяцией данных переписей населения и жилого фонда.

I. РЕДАКТИРОВАНИЕ ПРИ ПРОВЕДЕНИИ ПЕРЕПИСЕЙ И ОБСЛЕДОВАНИЙ

A. РЕДАКТИРОВАНИЕ: ИСТОРИЧЕСКАЯ СПРАВКА

26. До появления компьютеров большинство операций при проведении переписи требовали привлечения большого числа полуквалифицированных сотрудников для редактирования индивидуальных бланков. Однако из-за сложностей взаимосвязей между даже небольшим количеством элементов простая проверка не могла охватить все вероятные несоответствия в данных. Разные сотрудники могут по-разному интерпретировать правила, и даже один и тот же сотрудник может быть непоследовательным.

27. С появлением компьютеров редактирование данных переписи заметно изменилось. С помощью компьютера можно выявить намного больше несоответствий, чем при ручном редактировании. Спецификации редактирования становятся все более утонченными и сложными. Стала возможна автоматизированная интерполяция с соответствующими правилами такой работы (Nordbotten, 1963; Naus, 1975). В то же время такой процесс позволяет все чаще общаться с респондентами или, по крайней мере, иметь вопросники, заполненные этими респондентами. Многие группы редактирования стали понимать, что чем больше редакции, тем лучше и что, чем сложнее редактирование, тем точнее результаты. Благодаря этим программам появились тысячи сообщений об ошибках, требующих визуального просмотра первоначальных бланков, а при некоторых обследованиях — и повторного опроса респондентов.

28. С компьютерами становится все легче вносить изменения в набор данных. Иногда эти изменения корректируют записи или элементы данных. Многие записи проходят через компьютер множество раз, с ошибками и несоответствиями, и каждый раз их просматривают разные лица (Boucher 1991; Granquist, 1997).

29. В результате всего этого процесса появился целый ряд обобщенных пакетов редактирования данных переписи, и некоторые из них используются до сих пор. Первоначально, такие пакеты были разработаны для центральных компьютеров; позднее некоторые из них были модифицированы для того, чтобы использовать их в персональных компьютерах. За этот период Фелледжи и Холт (Fellegi and Holt, 1976) разработали новый метод обобщенного редактирования и интерполяции, который не сразу вошел в практику, но стал все шире применяться сегодня по мере того, как нацио-

нальные бюро переписи/статистические управления становятся все более искушенными в редактировании.

30. Крупный шаг вперед в редактировании данных переписи был сделан в 1980 году, когда национальные бюро переписи/статистические управления стали пользоваться персональными компьютерами для ввода, редактирования своих данных и сведения их в таблицы. При этом неожиданно для всех обработчики данных смогли производить редактирование в интерактивном режиме на этапе ввода данных или вскоре после этого. Для обследований появилась возможность разрабатывать программы для выявления ошибок во время сбора или введения данных непосредственно в машину. Компьютерное редактирование позволило иметь больше непрерывных контактов с респондентами для разрешения проблем, возникающих в процессе редактирования (Pierzchala, 1995).

31. В первые годы процесс проведения все более усложненных и тщательных проверок данных переписей и обследований представлялся весьма успешным. Группы редактирования создавали все более сложные спецификации редактирования, а специалисты по обработке данных месяцами занимались разработкой графических схем и программных кодов. Аналитики редко занимались оценкой пакетов. Представлялось, что путем редактирования можно решить любые проблемы, возникающие на более ранних этапах сбора, кодирования и ввода данных с клавиатуры. Но вместе с тем при этом для многих аналитиков стало очевидным, что во многих случаях все эти операции с дополнительным редактированием разрушали данные или, по меньшей мере, задерживали получение результатов или служили причиной погрешностей в этих результатах. Иногда программа делала так много проходов через данные, корректируя сначала один элемент, затем другой элемент, что результаты получались несопоставимыми с первоначальными, необработанными данными.

32. При проведении многих переписей и крупных обследований такое широкое редактирование вызвало значительные задержки в процессе проведения переписи или обследования. Сотрудники тратили много времени на поиски бланков вручную. Специалисты по обработке данных продолжали разрабатывать прикладные программы, которые относились к весьма небольшому количеству случаев. Гранкист (Granquist, 1997) отмечает, что многие исследования показали, что во многих слу-

чаях при такой дополнительной работе качество повышается незначительно, не повышается совсем или даже ухудшается. Многие виды серьезных систематических ошибок путем редактирования выявить нельзя.

33. По мере того как национальные организации по проведению переписей и обследований продолжают совершенствовать методику проведения этой работы, компьютерное редактирование в широких масштабах становится возможным и даже реальным. Следовательно, вопрос, который должен стоять перед каждым национальным бюро переписи/статистическим управлением, заключается в том, какой уровень компьютерного редактирования им подходит для их целей.

В. ГРУППА РЕДАКТИРОВАНИЯ

34. При подготовке к проведению переписи национальные статистические управления должны рассмотреть множество возможностей повышения качества своей работы. Одной из таких возможностей является создание группы редактирования. Ответственность за редактирование должна возлагаться на группу редактирования в составе руководителей, специалистов, занимающихся содержанием переписи, и обработчиков данных. Такую группу следует создавать, как только начинается подготовка к переписи, желательно — во время составления переписного листа. Группа редактирования играет важную роль с самого начала, и такая роль сохраняется за ней в течение всего процесса редактирования. Тщательный подход к подбору группы и к разработке и выполнению правил редактирования и интерполяции обеспечивает более оперативное и более эффективное проведение переписи.

35. Совместные совещания тех, кто занимается переписью, и пользователей по вопросам построения таблиц и других информационных продуктов могут позволить проникнуть в суть редактирования, которое необходимо производить. Зачастую пользователи просят составить конкретную таблицу или тип таблиц, который требует дополнительного редактирования с тем, чтобы устранить возможные несоответствия. Группа редактирования должна планировать построение таких таблиц на первоначальном этапе редактирования вместо того, чтобы составлять их специально после обработки результатов переписи. Разработка правил редактирования и компьютерных программ во время предварительного обследования или генеральной репетиции позволяет проверить сами программы и ускорить время оборота на различных этапах процесса редактирования и интерполяции. Затем группа редактирования устанавливает степень воздействия этих различных процессов и в случае необходимости принимает меры по исправлению положения.

36. Разработкой правил редактирования и интерполяции должны заниматься совместно специалисты конкретного профиля и специалисты по обработке данных. На самом начальном этапе подготовки к переписи группа редактирования составляет план рассмотрения и редактирования ошибок. Группа редактирования данных переписи или обследования составляет в письменном виде свод правил проверки соответствия и исправления ошибок.

37. Кроме разработки правил редактирования и интерполяции специалисты, занимающиеся содержанием переписи, и специалисты по обработке данных должны работать вместе на всех этапах проведения переписи или обследования, в том числе и в период проведения анализа. Риск проведения чрезмерного редактирования столь же велик, как и риск отредактировать слишком мало и получить необработанную или неправильную информацию в наборе данных. В связи с этим обе группы должны взять на себя ответственность за правильное ведение баз метаданных. Группа редактирования должна также эффективно использовать имеющиеся административные источники и регистры обследований, с тем чтобы лучше проводить последующие операции во время переписи или обследования.

38. Когда национальные бюро переписи/статистические управления использовали центральные компьютеры, общение между специалистами, занимающимися содержанием переписи, и специалистами по обработке данных было ограничено. Такое разделение продолжалось и некоторое время после появления микрокомпьютеров, но пакеты компьютерных программ стали более удобными для пользователя, и сейчас многие специалисты, занимающиеся содержанием переписи, могут фактически разрабатывать и испытывать собственные планы редактирования и табулирования. Хотя специалисты по конкретной тематике, как правило, сами не обрабатывают данные, они зачастую понимают, как действуют специалисты по обработке данных.

С. ПРАКТИКА РЕДАКТИРОВАНИЯ: ОТРЕДАКТИРОВАННЫЕ И НЕОТРЕДАКТИРОВАННЫЕ ДАННЫЕ

Рамка 1. Для чего нужно редактировать данные переписи

Редактирование данных переписи преследует следующие цели:

- 1) обеспечить пользователей высококачественными данными переписи;
- 2) выявить типы и источники ошибок;
- 3) предоставить откорректированные результаты переписи.

39. Страны проводят редактирование результатов переписи, с тем чтобы улучшить данные и их представление. В этом разделе *Руководства* освещается проблема, с которой сталкиваются национальные бюро переписи/статистические управления, когда публикуются неотредактированные данные переписи. Для иллюстрации используется гипотетический набор данных.

40. Национальное бюро переписи/статистическое управление вымышленной страны сталкивается с необходимостью выбора возможности постараться угодить многочисленным пользователям. Некоторым пользователям, возможно, нужны неизвестные введенные данные, используемые для анализа или научных исследований, а

другие, возможно, захотят иметь данные с минимальными искажениями (возможными ошибками) для целей планирования или проведения своей политики. Если национальное бюро переписи/статистическое управление распространит необработанную таблицу с такими данными как, например, в левой части таблицы 1, то и аналитики, и директивные органы будут вынуждены исходить из предположений, используя эти данные. В таблице 1 это показано на примере лишь небольшого количества человек. Из нее видно, что в отношении 23 человек в этой стране не зарегистрирован пол, а в отношении 15 человек — возраст. Эти пропуски могли возникнуть в результате отсутствия ответа или ошибок при вводе данных с клавиатуры. В двух из этих случаях не зарегистрирован ни пол, ни возраст.

Таблица 1. Выборка населения с разбивкой по возрастным группам с интервалом 15 лет и полу с использованием неотредактированных и отредактированных данных

Возрастная группа	Неотредактированные данные				Отредактированные данные		
	Всего	Мужской	Женский	Не зарегистрировано	Всего	Мужской	Женский
ВСЕГО	4 147	2 033	2 091	23	4 147	2 045	2 102
Младше 15 лет	1 639	799	825	15	1 743	855	888
15–29 лет	1 256	612	643	1	1 217	603	614
30–44 года	727	356	369	2	695	338	357
45–59 лет	360	194	166	0	341	182	159
60–74 года	116	54	59	3	114	53	61
75 лет и старше	34	12	22	0	37	14	23
Не зарегистрировано	15	6	7	2			

41. Большинство пользователей сами решают, что им делать с неизвестными величинами. Логический и, возможно, наивный подход заключается в том, чтобы распределить неизвестные величины в той же пропорции, что и известные. Если национальное бюро переписи/статистическое управление примет решение интерполировать неизвестные величины, группа редактирования может решить оставить 12 мужчин и 11 женщин, то есть примерно в одинаковой пропорции, однако этот показатель будет искаженным, потому что при переписи зарегистрировано больше женщин. В этом случае результаты будут соотноситься с отредактированными данными, указанными в правой стороне таблицы 1.

42. Для обработки неизвестных величин есть и другие варианты. Например, группа редактирования может постановить интерполировать данные, исходя только из распределения по полу, игнорируя другую имеющуюся информацию, например, отношения между супругами, зарегистрировано ли лицо неуказанного пола как мать другого лица или дается ли положительный ответ в отношении лица неуказанного пола о количестве детей, родившихся живыми. Альтернативная стратегия интер-

поляции может заключаться в том чтобы учитывать одну или более из этих других переменных величин.

43. Национальное бюро переписи/статистическое управление может еще в качестве альтернативы взять за основу интерполяцию по возрастному распределению. Для выборки населения, показанной в таблице 1, незарегистрированный возраст отмечается в общей сложности в пятнадцати случаях. Эти данные можно также распределить в тех же пропорциях, что и известные величины, то есть, опять в соответствии с логической стратегией интерполяции. При этом группа редактирования может, вероятно, получить более точные результаты, учитывая другие переменные и их комбинации, например, относительный возраст мужа и жены, родителя и ребенка или дедушки или бабушки, внука или внучки, или наличие детей школьного возраста, пенсионеров и лиц из числа самостоятельного населения.

44. В таблице 1 отредактированные данные справа «чище», поскольку неизвестные величины изъяты (см. колонки в графе «отредактированные данные»). В этой части таблицы нет неизвестных величин, поскольку про-

грамма распределяет их по другим ответам. Тем не менее, многие демографы и другие специалисты, занимающиеся содержанием переписи, по традиции хотят, чтобы неизвестные величины указывались в таблицах, как то сделано в графе «неотредактированные данные» в таблице 1. Они считают, что такая процедура позволяет им производить различные виды оценок по показателям, чтобы определить эффективность переписных процедур или оказать содействие в планировании дальнейших переписей и обследований. Обе задачи вполне выполнимы — составить отредактированную таблицу для реальных пользователей и неотредактированную таблицу для оценки — включая в них неизвестные величины и обходясь без них.

45. Еще одна проблема с использованием неизвестных величин в публикуемых таблицах заключается в том, что неизвестные величины могут влиять на результаты анализа тенденций. Новая технология позволяет намного упростить такой анализ по сравнению с при-

нятой практикой. Например, в таблице 2 показано распределение по возрасту из двух проводимых одна за другой переписей. Количество неизвестных величин снизилось в отношении этой небольшой страны с 217 или 6,5 процента зарегистрированных ответов в 1990 году до 15 или менее 1 процента ответов в 2000 году.

46. Здесь национальное бюро переписи/статистическое управление должно учитывать, как несогласующееся число неизвестных величин влияет на отдельную перепись и на изменения между переписями. Например, с показателем 6,5 процента неизвестных величин в переписи 1990 года трудно сопоставить изменения в процентном распределении для возрастных групп с интервалом 15 лет по двум переписям. За десятилетие процент лиц в возрасте от 15 до 29 лет по показателям увеличился с 27 процентов всего лишь до 30, но распределенные неизвестные величины могут изменить результаты анализа.

Таблица 2. Население и изменения в его составе по возрастным группам с интервалом 15 лет с неизвестными величинами: 1990 и 2000 годы

Возрастная группа	Число		Изменение в числовом выражении	Изменение в процентном выражении	Процент	
	2000 год	1990 год			2000 год	1990 год
ВСЕГО	4 147	3 319	828	24,9	100,0	100,0
Младше 15 лет	1 639	1 348	291	21,6	39,5	40,6
15–29 лет	1 256	902	354	39,2	30,3	27,2
30–44 года	727	538	189	35,1	17,5	16,2
45–59 лет	360	200	160	80,0	8,7	6,0
60–74 года	116	89	27	30,3	2,8	2,7
75 лет и старше	34	25	9	36,0	0,8	0,8
Не зарегистрировано	15	217	–202	–93,1	0,4	6,5

47. В пересмотренной таблице 3 показаны неизвестные величины, распределенные или пропорционально, или путем того или иного исчисления. Здесь намного проще увидеть изменения как в числовом, так и в процентном отношении, а также распределение возрастных групп в двух переписях. Несомненно, для

того чтобы получить точные, надежные результаты группа редактирования должна удостовериться в том, что отредактированные данные согласуются между двумя переписями и/или обследованиями, а также внутренне совместимы. Строка «не зарегистрировано» опущена.

Таблица 3. Население и изменения в его составе по возрастным группам с интервалом 15 лет без неизвестных данных: 1990 и 2000 годы

Возрастная группа	Число		Изменение в числовом выражении	Изменение в процентном выражении	Процент	
	2000 год	1990 год			2000 год	1990 год
ВСЕГО	4 147	3 319	828	24,9	100,0	100,0
Младше 15 лет	1 743	1 408	335	23,8	42,0	42,4
15–29 лет	1 217	952	265	27,8	29,3	28,7
30–44 года	695	578	117	20,2	16,8	17,4
45–59 лет	341	230	111	48,3	8,2	6,9
60–74 года	114	109	5	4,6	2,7	3,3
75 лет и старше	37	42	–5	–11,9	0,9	1,3

Д. ОСНОВЫ РЕДАКТИРОВАНИЯ

48. Редактирование — это систематическая проверка и корректировка (или внесение изменений) ответов в соответствии с заранее установленными правилами. Некоторые операции редактирования включают в себя ручную корректировку, что означает корректировку рукой человека. Другие операции редактирования заключаются в электронном редактировании с использованием компьютеров. В публикациях результатов переписи может содержаться какое-то количество незначительной информации, если национальные бюро переписи/статистические управления не редактируют результаты переписи или обследования. Редактирование снижает количество искаженных оценок, облегчает обработку и повышает доверие пользователя. Кроме того, как отмечали Пуллам, Харфам и Ожевер (Pullum, Harpham and Ozsever, 1986), главная задача успешного редактирования или чистки состоит в том, чтобы, в первую очередь, определить, согласуются ли различные ответы один с другим и с основным форматом инструментария обследования.

49. Файлы первичных данных переписи содержат самые разнообразные ошибки. При обработке данных ошибки подразделяются на две категории. Ошибки, которые могут заблокировать дальнейшую обработку данных, и ошибки, которые дают недостоверные или несогласующиеся результаты без нарушения логической последовательности дальнейших операций по обработке данных. Как отмечается в *первом пересмотренном варианте Принципов и рекомендаций в отношении переписей населения и жилого фонда* (ООН, 1998 год, пункт 1.195), исправляться должны все ошибки первого рода, а также как можно больше ошибок второго рода. Таким образом, основная цель редактирования результатов переписи на этапе обработки заключается в том, чтобы выявить как можно больше ошибок и внести изменения в набор данных с тем, чтобы элементы данных были достоверными и согласовывались между собой. Однако в процессе обработки нельзя исправить все ошибки переписи, в том числе ответы на вопросник, которые согласуются в рамках одной переписи, но фактически являются примерами неправильного сообщения со стороны респондентов или неправильной регистрации со стороны счетчиков.

50. Редактирование, как правило, разделяется на две категории: 1) фатальное редактирование, при котором ошибки выявляются с уверенностью, и 2) редактирование под вопросом, которое указывает на сомнительные элементы данных (Granquist and Kovar, 1997, p. 420). При фатальном редактировании выявляются элементы данных, которые определенно даны ошибочно, а при редактировании под вопросом отмечаются данные, которые могут быть недостоверными или несогласующимися. К фатальным ошибкам, то есть ошибкам, выяв-

ляемым при фатальном редактировании, относятся недостоверные или отсутствующие данные, а также ошибки, обусловленные несогласованностью. И, наоборот, при редактировании под вопросом, выявляются элементы данных, которые не подпадают под преимущественно субъективные пределы редактирования, элементы относительно высокие или низкие по своему значению, если сравнивать их с другими данными в том же вопроснике, и другие сомнительные записи. Для того чтобы поддерживать доверие к переписи, особенно когда национальное бюро переписи/статистическое управление принимает решение распространять микроданные, процесс редактирования должен быть направлен на выявление и обработку фатальных ошибок. Сомнительные ошибки исправлять труднее, при этом преимуществ меньше, чем при выявлении и исправлении фатальных ошибок, а стоимость всего процесса становится выше.

51. Поскольку все элементы включаются в перепись исключительно потому, что они нужны планирующим и директивным органам, то во время редактирования и интерполяции данных переписи необходимо исправлять больше сомнительных ошибок, чем при обследованиях. Тем не менее, определяя окончательные результаты редактирования данных переписи, сотрудники, занимающиеся содержанием переписи, должны тщательно изучить приемы редактирования, разработанные для экспериментальных переписей, и методы, выработанные во время обработки данных, с тем чтобы убедиться, что отдельные приемы редактирования соответствуют ожидаемым затратам и результатам. Такое исследование должно проводиться в рамках оценки результатов переписи. Как отмечают Гранкист и Ковар (Granquist and Kovar, 1997, p. 422) данные по коэффициентам совпадений, которые входят в число признаков, ведущих к изменениям в первоначальных данных, редко сообщаются при оценке или изучении процессов редактирования.

52. Еще один набор методов и терминологии относится к микроредактированию и к макроредактированию. Как уже отмечалось, при редактировании данных переписи и обследования выявляются ошибки в записях данных и между ними. В настоящем *Руководстве* описывается процесс микроредактирования, который направлен на то, чтобы обеспечить достоверность и согласованность отдельных информационных записей и взаимосвязей между записями в домохозяйстве. Еще один метод, макроредактирование, заключается в том, чтобы проверять агрегированные данные, с тем чтобы убедиться, что они также обоснованы. Например, в стране может быть очень большой процент лиц, которые не сообщают свой возраст. После исчисления возраста, для того чтобы получить полный набор данных, проверки на макроуровне или агрегированном уровне могут показать, что выборочные данные, не сообщенные лицами старшего возраста, не искажают исчисленные

величины. Группа редактирования в зависимости от результатов анализа может предпочесть принять меры к тому, чтобы снизить риск потенциального искажения.

53. При редактировании следует по мере возможности сохранять первоначальные данные. Группа редактирования должна работать с достоверными данными высокого качества, но при этом сохранять те данные, которые организация собирает на месте. Первоначальные данные необходимо сохранять на всех этапах машинной обработки, если группа редактирования решит, что ей необходимо пересмотреть сам процесс редактирования. Иногда первоначальные данные приходится пересматривать, если группа обнаружит, что в процессе редактирования появилась систематическая ошибка. Иногда пересмотр обусловлен тем, что часть показателей в наборе данных отсутствует или дублируется и тогда набор данных приходится пересматривать и снова редактировать.

54. Иногда источник ошибок находится за пределами того учреждения, где ведется обработка. Банистер (Banister, 1980, р. 2) отмечает, что, если известно, что большая доля какой-то подгруппы не ответила на конкретный вопрос переписи, это означает, что они или не поняли вопроса, или не хотят отвечать на него и содействовать проведению переписи. При этом она утверждает, что коэффициент неполучения ответов для подгрупп следует записывать на носители данных переписи и в публикуемые таблицы. Сейчас национальные бюро переписи/статистические управления больше предпочитают хранить эти данные на компакт-дисках или других носителях для исследователей.

55. Появляется все больше подтверждений тому, что никакое компьютерное редактирование не может заменить собой сбор данных переписей более высокого качества. Национальные бюро переписи/статистические управления знают, что в какой-то момент автоматизированное редактирование не только ограничено, но и приводит к обратным результатам: редактирование добавляет больше ошибок в набор данных, чем их исправляет. Замена элемента переписи — это не то же самое, что его корректировка. В связи с этим группа редактирования должна вся вместе определить начало, середину и конец процесса редактирования.

56. Редактирование и интерполяция могут улучшить, а могут и не улучшить качество данных, но чистый набор данных намного облегчает проведение анализа. Процесс начинается с разработки переписного листа. Демографы и другие специалисты, занимающиеся содержанием переписи, как правило, определяют, что нужно включать в него, зачастую консультируясь с группами пользователей. В конечном итоге, данные переписи выпускаются в первую очередь не конкретно для демографов, а для намного более широкой аудито-

рии ученых, тех кто занимается выработкой политики, и простых людей (Banister, 1980, р. 17). Однако улучшение результатов переписи без недостоверных и несогласующихся данных необходимо, если на карту поставлено доверие к переписи и репутация национального бюро переписи/статистического управления. Как отмечает Банистер, занимающиеся переписью организации могут сослаться на примеры журналистов, пишущих юмористические статьи, или граждан, с негодованием пишущих официальным лицам, занимающимся переписью, об опубликованных таблицах, где указаны трехлетние дедушки и пассажиры, пользующихся несуществующими поездами.

57. Проблема заключается в том, чтобы определить, как далеко можно заходить, чтобы получить набор данных хорошего качества. Как отмечалось выше, появление компьютеров, сначала центральных компьютеров, а затем микрокомпьютеров, позволило практически полностью автоматизировать процесс редактирования. Во многих национальных бюро переписи/статистических управления специалисты, занимающиеся содержанием переписи, стали по существу энтузиастами редактирования. В связи с этим сейчас бюро проводят много проверок на согласованность, что раньше трудно было сделать, в частности, перекрестные проверки записей и перекрестные проверки домохозяйств. К сожалению, это свойство микрокомпьютеров вызвало и множество проблем, самой большой из которых является чрезмерное редактирование.

1. *Как чрезмерное редактирование оказывает пагубное воздействие*

58. Чрезмерное редактирование оказывает отрицательное воздействие на весь процесс редактирования по-разному, в том числе, удлиняя сроки, повышая затраты и искажая истинные значения. Оно также дает ложное чувство уверенности в качестве данных. Обо всем этом говорится ниже.

а) *Своевременность*

59. Чем больше национальное бюро переписи/статистическое управление занимается редактированием, тем больше времени занимает весь процесс. Главный вопрос заключается в том, чтобы определить, сколько добавленное время прибавляет к ценности продукта переписи. Каждая группа редактирования должна оценивать на постоянной основе и постфактум чистую выгоду от добавленного времени и ресурсы для всего продукта переписи. Зачастую выгоды с точки зрения затраченного времени настолько ничтожны, что лучше иметь небольшие проблемы с данными, чем лишать ос-

новых пользователей возможности получать своевременную информацию.

b) *Финансы*

60. Аналогичным образом с увеличением времени увеличиваются и затраты на проведение переписи. Каждое национальное бюро переписи/статистическое управление должно решить, если оно увеличивает объем и сложность редактирования данных, стоит ли увеличение затрат дополнительных усилий и может ли оно себе позволить эти дополнительные затраты.

c) *Искажение истинных значений*

61. Хотя в целом процесс редактирования направлен на повышение качества данных, увеличение объема и сложности редактирования может повлечь за собой и отрицательные последствия. Иногда группы редактирования ошибочно меняют элементы данных по самым разным причинам. Это может происходить и из-за недостаточного общения между специалистами по обработке данных и специалистами, занимающимися содержанием переписи; ошибок в очень сложной и насыщенной программе; или многократной обработки того или иного элемента переписи во время редактирования. Национальные бюро переписи/статистические управления намерены всячески избегать такого рода проблем. Например, Гранкист и Ковар (Granquist and Kovar, 1997) указывают, что исчисление возраста мужа или жены по установленной возрастной разнице между ними может быть полезным, но может искусственно исказить данные, если таких случаев много.

d) *Ложное чувство уверенности*

62. Чрезмерное редактирование дает сотрудникам национального бюро переписи/статистического управления и другим пользователям ложное чувство уверенности, особенно когда бюро не осуществляют и не документируют меры по обеспечению гарантии качества. Кроме того, в таблицах с данными переписи появляются случайные результаты независимо от того, какой объем работы делает группа редактирования, поэтому необходимо предупреждать пользователей о том, что могут возникать небольшие ошибки. Особенно это актуально сейчас, когда многие страны публикуют выборочные микроданные. Национальные бюро переписи/статистические управления в принципе не имеют желания публиковать данные в ущерб процессу планирования, поэтому необходимо тщательно следить за тем, чтобы все наиболее важные переменные редактировались должным образом и могли быть использованы при планировании. Например, ни одно национальное бюро переписи/статистическое управление не хотело бы

публиковать микроданные или таблицы с неизвестными показателями пола или возраста. Вместе с тем, такими переменными, как нетрудоспособность или грамотность, можно оперировать, особо не подвергая их редактированию. В сводных таблицах могут появляться некоторые несоответствия, поскольку национальные бюро переписи/статистические управления не могут отредактировать все пары переменных, и группы редактирования должны проверять наиболее важные комбинации. Если при редактировании обнаруживаются несогласующиеся величины, то в группе редактирования должна быть предусмотрена процедура корректирования.

2. *Обработка неизвестных величин*

63. Группа редактирования должна решить на самом раннем этапе планирования переписи, как обращаться с «неуказанными» или неизвестными величинами. Как отмечалось выше, колонки или ряды неизвестных величин в таблицах не несут никакой информации и бесполезны, поэтому планирующие органы в большинстве стран предпочитают, чтобы эти данные интерполировались. Без обработки неизвестных величин многие пользователи распределяют их в выпускаемых таблицах в той же пропорции, что и известные данные, таким образом интерполируя неизвестные величины постфактум. Группа редактирования должна решить, по какой системе обрабатывать неизвестные величины.

3. *Кажущиеся изменения*

64. Национальные бюро переписи/статистические управления, как правило, не работают с моделями, если они разрабатывают собственные правила редактирования. Группы редактирования должны разрабатывать правила, которые соответствовали бы фактическим характеристикам населения или жилищ. Все данные должны редактироваться по правилам. Например, свод правил может предусматривать, чтобы ребенок главы домохозяйства был, по меньшей мере, на 15 лет младше главы. Однако ребенок главы домохозяйства может быть фактически социальным, а не биологическим ребенком: он или она может быть биологическим ребенком супруга, который при этом не является главой домохозяйства. Отсюда разница в возрасте может быть меньше 15 лет. Поскольку планирующие органы в большинстве стран не проводят различий между детьми и пасынками или падчерицами, то при вышеупомянутых обстоятельствах, если по правилам редактирования меняется возраст ребенка, могут возникать несоответствия в уровне образования, показателях самостоятельности населения и других областях. С учетом этого, такое правило следует проверить прежде, чем применять его

в полной мере, чтобы посмотреть, какие будут результаты.

4. *Определение уровня допустимых отклонений*

65. Группа редактирования должна разработать «уровни допустимых отклонений» для каждого элемента, а иногда для сочетания элементов. Уровень допустимых отклонений показывает допустимое количество недостоверных и несогласующихся ответов, прежде чем группа редактирования приступит к устранению ошибок. Например, в отношении большинства элементов переписи какой-то небольшой процент респондентов не даст «приемлемых» ответов по той или иной причине. Для некоторых элементов, например, таких как возраст и пол, которые используются в сочетании со многими другими элементами в целях планирования, уровень допустимых отклонений может быть весьма низким. В случае низкой доли недостающих или несогласующихся ответов (менее одного или двух процентов), любые приемлемые правила редактирования вряд ли повлияют на использование данных. Если эта доля высокая (5–10 процентов или более в зависимости от ситуации) в простых или даже в сложных случаях, интерполяция может исказить результаты переписи.

66. Чтобы сократить до минимума количество недостающих ответов, национальные бюро переписи/статистические управления должны добиваться того, чтобы те, кто проводит перепись, прилагали все усилия для получения информации на месте. Если та или иная страна решит, что ей не нужна столь большая точность в отношении некоторых элементов, таких как грамотность или нетрудоспособность, то уровень допустимых отклонений для таких элементов может быть намного выше. Иногда группы редактирования могут корректировать элементы, которые содержат слишком много ошибок, вновь направляя счетчиков на места, проводя повторный опрос по телефону или используя свои знания в той или иной области. Однако зачастую повторное направление счетчиков на места или проведение каких-либо других последующих операций обходится слишком дорого и национальные бюро переписи/статистические управления могут принять решение или не использовать данный элемент, или использовать его, сопровождая предупредительными примечаниями.

67. При этом возникает вопрос, кто должен определять уровень допустимых отклонений для того или иного элемента. Установить уровень допустимых отклонений может команда редактирования, включая специалистов, занимающихся содержанием переписи, и специалистов по обработке данных. Сотрудники, занимающиеся содержанием переписи, должны использовать элементы в течение долгого времени и поэтому

профессионально заинтересованы в том, чтобы получить данные самого высокого качества. Однако специалисты по обработке данных могут признать, что они фактически не в состоянии разработать соответствующие программы редактирования, чтобы сократить уровень допустимых отклонений до приемлемых размеров и что сами данные могут препятствовать успешному применению программы с учетом уровня допустимых отклонений.

5. *Извлечение уроков из процесса редактирования*

68. По мере редактирования данных переписи необходимо проводить подробный документальный анализ положительной и отрицательной обратной связи в целях повышения качества результатов проводимой переписи или обследования с перспективой для дальнейших переписей и обследований. Группа редактирования должна постоянно работать над тем, чтобы определять, что работает должным образом, а что не работает. Она также должна определить, можно ли улучшить и рационализировать те аспекты данного процесса, которые работают должным образом с тем, чтобы эти данные доходили до пользователя еще быстрее. Чем раньше национальные бюро переписи/статистические управления обнаружат в ходе проведения переписи ошибки, тем больше вероятность их исправления.

6. *Гарантия качества*

69. На протяжении всей переписи важное значение имеет гарантия качества. Исходя из этого, несомненно, должны действовать официальные механизмы обеспечения качества с тем, чтобы наблюдать за ходом компьютерного редактирования и этапом интерполяции. Для анализа качества редактирования и скорости обработки крайне важное значение имеют контрольный журнал, критерии качества и диагностическая статистика (Granquist and Kovar, 1997; Statistics Canada, 1998).

7. *Затраты на редактирование*

70. Данное *Руководство* может помочь странам сократить большие затраты времени и ресурсов, необходимых для редактирования и интерполяции данных переписи или обследования. Как отмечают Гранкист и Ковар (1997, р. 418), даже в 90-х годах редактирование по существу обходится столько же дорого, как и в 70-х годах, несмотря на то, что данный процесс в значительной степени рационализирован благодаря постоянному использованию технологических достижений. В большинстве стран редактирование отнимает несораз-

мерно много времени и средств, поэтому каждая страна должна определить доходность инвестирования в эту работу. Согласно Гранкисту и Ковару (Granquist and Kovar, 1997), затраты на редактирование данных обследований домашних хозяйств в начале 90-х годов составляли около 20 процентов всех бюджетных средств в мире, выделяемых на проведение переписей.

71. Чрезмерное редактирование может задержать получение результатов переписи. Хотя сотрудники, занимающиеся у себя в стране переписями/обследованиями, могут иметь лишь эпизодические свидетельства этого в связи с переписями, исследование Пуллама, Харфама и Ожевера (Pullum, Harpham and Ozsever, 1986) показало, что в результате машинного редактирования результатов всемирного обследования фертильности публикация результатов задержалась почти на один год. Национальные бюро переписи/статистические управления могли бы лучше распоряжаться своими средствами, чтобы во время переписи или обследования проводить, прежде всего, регистрацию с более высоким качеством.

8. Интерполяция

72. Интерполяция — это процесс разрешения проблем, касающихся недостающих, недостоверных или несогласующихся ответов, выявленных во время редактирования. Интерполяция проводится путем замены одного или более ответов или недостающих величин в записи или нескольких записях, которые редактируются с тем, чтобы получить достоверные логически согласованные учетные данные. Контакт с респондентом или изучение вопросника с помощью ручных приемов устраняет некоторые проблемы в самом начале процесса. Однако решить все проблемы на этих ранних этапах, как правило, невозможно из-за большого количества ответов, затрат и установленных сроков. Таким образом, путем интерполяции обрабатывается то, что не удалось сделать при редактировании, поскольку желательно получить полный и согласованный файл с интерполированными данными. Лучше всего интерполяцию проводят те члены группы, которые имеют полный доступ к микроданным и владеют хорошей дополнительной информацией.

- a) Интерполированная запись должна быть очень похожа на запись, которую не удалось отредактировать. Лучше всего интерполировать минимальное число переменных, сохраняя, таким образом, как можно больше данных, представленных респондентом. При этом исходят из предположения (которое на практике не всегда верно) о том, что у респондента больше вероятность сделать всего одну или две ошибки, а не несколько;
- b) интерполированная запись должна согласовываться со всеми отредактированными данными;
- c) группы редактирования должны пометать интерполированные величины и четко указывать методы и источники интерполяции. Группа редактирования должна сохранять неинтерполированные и интерполированные величины в полях записи с тем, чтобы можно было оценить степень и последствия интерполяции.

9. Архивирование

73. Процедура обеспечения качества данных переписи или обследования заключается, в частности, в том, чтобы документировать все операции, а затем эту документацию архивировать. Национальные бюро переписи/статистические управления должны сохранять как файлы с отредактированными данными, так и файлы с неотредактированными данными для последующего анализа. Некоторые процедуры, например сканирование, автоматически сохраняют первоначальное отображение. Кроме того, сразу после ввода с клавиатуры пакетов данные следует сцепить и сохранить для возможного анализа.

74. Документация должна быть достаточно полной, для того чтобы органы, планирующие проведение переписи или обследования, могли восстановить те же процессы позднее с тем, чтобы обеспечить сопоставимость с проводимой переписью или обследованием. И процессы и результаты должны быть воспроизводимы. И, наконец, неотредактированные, также как и отредактированные данные, необходимо хранить в нескольких местах, приняв надлежащие меры по обеспечению их постоянной доступности с течением времени.

II. ПРИКЛАДНЫЕ ПРОГРАММЫ РЕДАКТИРОВАНИЯ

75. В главе II дается общий обзор прикладных программ для редактирования и интерполяции. В ней приводятся отдельные примеры, показывающие, какого рода проблемы могут возникнуть у пользователя в связи с необработанными данными и почему отредактированные данные более полезны. В ней рассматриваются вопросы введения данных с клавиатуры и кодирования в качестве предварительной части процесса редактирования. В данной главе представлены также общие вопросы компьютерного редактирования наряду с руководящими принципами по таким вопросам, как проверка на достоверность и согласованность. Два характерных для машинного редактирования способа, статическая интерполяция (*cold deck*) и динамическая интерполяция (*hot deck*), рассматриваются подробно.

76. Цель редактирования данных переписей и обследований заключается в обнаружении пропусков и несоответствий в записях данных, а для их устранения применяется метод интерполяции. Для обработки пропусков и различных видов неприемлемых элементов вводимых данных применяются конкретные процедуры редактирования. При интерполяции меняются недостоверные данные и устраняются несоответствия, обнаруженные в наборе данных. Результатом этой работы является отредактированный файл микроданных для сведения их в таблицы с указанием приемлемых и согласующихся величин по всем применяемым элементам данных для каждой зарегистрированной жилищной единицы или лица.

77. Опять же необходимо отметить, что никакое редактирование не может заменить собой высококачественную регистрацию. Процесс редактирования хорошо срабатывает, когда метод интерполяции применяется по отношению к произвольным пропускам и несоответствиям. Однако если во время сбора данных возникают систематические ошибки, качество данных нельзя повысить путем редактирования, какими бы усложненными ни были процедуры. С точки зрения качества получаемых данных основополагающее значение имеет выбор вопросов для исследования. Во время опроса респонденты должны проявлять желание и суметь предоставить нужную информацию. Таким образом, видимо, необходимо избегать тем, которые могут вызвать опасения, пробудить предрассудки или суеверия, а также вопросов слишком сложных и трудных для того,

чтобы средний респондент мог легко ответить на них в контексте переписи населения. Точная формулировка каждого вопроса, которая необходима для получения наиболее достоверного ответа, неизбежно будет зависеть от национальных условий и должна быть проверена до проведения переписи. В связи с этим крайне важно, чтобы национальные бюро переписи/статистические управления выделяли достаточно ресурсов для получения данных переписи высокого качества.

78. Для проведения этапа компьютерного редактирования группа редактирования готовит письменные инструкции или спецификации по редактированию, таблицы решений, графические схемы и символический код². Графические схемы помогают специалистам, занимающимся содержанием переписи, понимать различные взаимосвязи между переменными и облегчают написание инструкций по редактированию. Выборочные графические схемы показаны в приложении IV. Специалисты, занимающиеся содержанием переписи, готовят инструкции по редактированию в сотрудничестве со специалистами по компьютерам, расписывая, что нужно делать по каждому элементу данных. Инструкции по редактированию должны быть ясными, лаконичными и четко выраженными, поскольку они служат основой для всего пакета программ редактирования.

79. Вся группа, которая занимается редактированием данных переписи, и демографы, и обработчики данных, должна обладать обширными познаниями в деле обработки и анализа демографических данных. Неквалифицированный персонал может непреднамеренно внести дополнительные ошибки и погрешности в результаты переписи.

А. РУЧНОЕ И АВТОМАТИЗИРОВАННОЕ КОРРЕКТИРОВАНИЕ

80. Ручное редактирование данных переписи может занять месяцы или годы, предоставляя множество возможностей для внесения ошибок по вине человека. Ручное редактирование является слабой альтернативой автоматизированному редактированию отчасти потому, что нельзя создать или воссоздать журнал редактиро-

² Символический код — это набор письменных инструкций или спецификаций по редактированию, как показано на рисунке 7.

вания для ручной корректировки. Компьютерное, или автоматизированное редактирование, сокращает время, необходимое для обработки данных, и количество ошибок по вине человека. Как при компьютерном так и при ручном редактировании производится проверка достоверности той или иной величины путем поиска приемлемого значения, но с помощью компьютерных программ производится также проверка значения, по сравнению со смежными величинами с точки зрения согласованности. И, наконец, что наиболее важно, автоматизированное редактирование позволяет создать журнал редактирования и поэтому его можно воспроизвести, а ручное редактирование не позволяет делать этого.

81. Когда при переписях и обследованиях набираются огромные тома данных, сотрудники не всегда могут обращаться к исходным документам для исправления ошибок. Даже если имеются исходные вопросники, записанные в них данные иногда могут быть неправильными или не согласовываться между собой. При компьютерном редактировании и системе интерполяции можно незамедлительно исправить или изменить ошибочные данные и составить отчет по всем обнаруженным ошибкам и всем произведенным изменениям. Компьютерное редактирование следует тщательно планировать с тем, чтобы сэкономить сотрудникам время для других операций по обработке данных. Хотя прогон большого количества данных через компьютерную систему может потребовать много времени, это все же не отнимает столько времени, сколько при ручном редактировании.

82. Ручное корректирование имеет несколько форм. Рассмотрим простой пример ошибки в ответе о поле: инспектор проверяет работу счетчика и находит очевидную ошибку, например, «мужской» приписан кому-то по имени «Мэри». При изменении пола на «женский» инспектор производит ручное редактирование. Если инспектор не исправляет вопросник, а вместо этого направляет его в отделение на месте, то сотрудники того отделения могут увидеть проблему и вручную исправить ошибку. В центральном отделении во время кодирования кодировщики могут заметить несоответствие между именем и полом и тоже произвести ручную корректировку. Или же кодировщики могут и не заметить эту проблему, но несогласованность между именем и полом могут заметить операторы, вводя данные вопросников с клавиатуры, и в этом случае они могут произвести ручное корректирование до ввода данных.

83. Однако если ошибка не замечена, и оператор вводит код для ответа «мужской», здесь можно поступить по-разному. В отношении элементов, связанных с полом, например, в блоке «рождаемость», редактор может отметить тот факт, что это мужчина с информацией о рождаемости и сообщить об этом, ко-

гда оператор вводит данные. Оператор может после этого посмотреть на вопросник, увидеть, что это действительно женщина, и произвести исправление вручную. Другой вариант состоит в том, что, если национальное бюро переписи/статистическое управление пользуется программой-редактором, независимой от клавиатуры, компьютерная программа может пометить этого человека как мужчину с информацией о рождаемости. Затем, используя географическую информацию, работники управления могут найти в корзине исходный вопросник, достать его и определить, что респондент по имени «Мэри» был ошибочно зарегистрирован не как женщина, а как мужчина. В этом случае сотрудник управления может снова передать информацию оператору, а тот может вытащить запись и произвести ручную корректировку.

84. Этот пример показывает как преимущества, так и недостатки ручного редактирования. На любом из вышеуказанных этапов те, кто занимаются переписью, могут обнаружить ошибку — несоответствие между именем и полом — и исправить ее. Однако в национальных бюро переписи/статистических управлениях, использующих ручное редактирование, возможно, есть сотрудники, проверяющие эту взаимосвязь на каждом этапе. На эту работу уходит огромное количество энергии, а результаты, возможно, мало отличаются, особенно в агрегированной форме, от тех, которые получаются, если сотрудникам дано указание не производить ручное редактирование.

85. До недавнего времени единственным способом внесения исправлений в набор данных был ручной. Многие страны до сих пор испытывают неудобства, используя автоматизированное корректирование, и поэтому они делают это вручную на одном из этапов, описанных выше. Если набор данных небольшой, сроки приемлемы и работники действуют интенсивно, то во многих случаях ручное корректирование срабатывает. Преимущество здесь заключается в том, что если информация в вопроснике представлена полно и точно, а несогласованность можно фактически устранить, глядя на заполненный бланк, то качество переписи или обследования, вероятно, в самой малой степени улучшится (группа редактирования должна исходить из того, например, что «Мэри» — это не «Гари»; что если речь идет о рождаемости, то данные, видимо, собраны в отношении этого лица, а не по ошибке). Фактически, процедуры редактирования и интерполяции редко повышают качество собранных данных. Они только меняют некоторые элементы.

86. Иногда просматривать переписной лист в целях ручного корректирования бесполезно. Информация там по какой бы то ни было причине отсутствует. Иногда человек не хочет указывать свой возраст, поэтому этот

элемент в вопроснике отсутствует. В этом случае изучение вопросника не разрешит проблему. Тогда группа редактирования должна решить, как выйти из этой ситуации. Для ручного корректирования национальное бюро переписи/статистическое управление должно вписать «не известен» или применить какой-то набор величин, чтобы установить позицию возраста.

87. Если не обращаться к респонденту, ручное корректирование неизбежно снижает качество и согласованность. Оно требует больше времени и затрат. Компьютеры не устают, поэтому они быстрее; у них нет личных проблем, которые могли бы помешать сохранению качества или согласованности; и в большинстве случаев они снижают себестоимость обработки. Большинство стран используют сейчас тот или иной вид автоматизированного корректирования.

88. Недостающие и несогласующиеся ответы снижают качество данных и затрудняют представление данных переписи в легко понимаемой форме. Некоторые пользователи предпочитают указывать в таблицах недостающие или несогласующиеся ответы в категории «не сообщается», в то время как другие предпочитают распределять их пропорционально между сообщенными согласующимися данными. Еще одна группа пользователей рекомендует устанавливать правила интерполяции «вероятных» ответов для недостающих или несогласующихся данных. Использование компьютеров позволяет эффективно интерполировать ответы на основе другой информации в вопроснике или указанной информации в отношении лица или жилищной единицы с аналогичными характеристиками.

89. Поскольку компьютер позволяет увидеть многие характеристики, из этого следует извлекать пользу в

процессе редактирования. Таким образом, редактирование с использованием многих смежных характеристик может позволить интерполировать более приемлемые ответы, чем при простом редактировании. Вместе с тем слабо разработанная процедура редактирования может привести к получению данных переписи плохого качества. В группу редактирования должны входить опытные специалисты, занимающиеся содержанием переписи, из самых разных имеющих к этому отношение областей, а также обработчики данных. Члены группы редактирования должны тщательно отобрать переменные для их проверки на согласованность с тем, чтобы определить спецификации редактирования и интерполяции. Программная продукция должна содержать долю ответов, которые были изменены или интерполированы. В этом случае аналитикам легче будет судить о качестве данных; например, большая доля интерполированных ответов будет служить предупреждением к тому, чтобы использовать эти данные с осторожностью.

90. Журнал редактирования, или контрольный журнал фиксирует изменения каждой переменной. Журнал используется для того, чтобы проследить историю ответов, начиная с получения данных до процесса редактирования и интерполяции.

В. РУКОВОДЯЩИЕ ПРИНЦИПЫ ИСПРАВЛЕНИЯ ДАННЫХ

91. Редактирование, будь-то ручное или автоматизированное, должно быть направлено на то, чтобы данные как можно ближе отображали реальную ситуацию путем устранения пропусков и неправильных ответов или путем изменения несогласующихся данных.

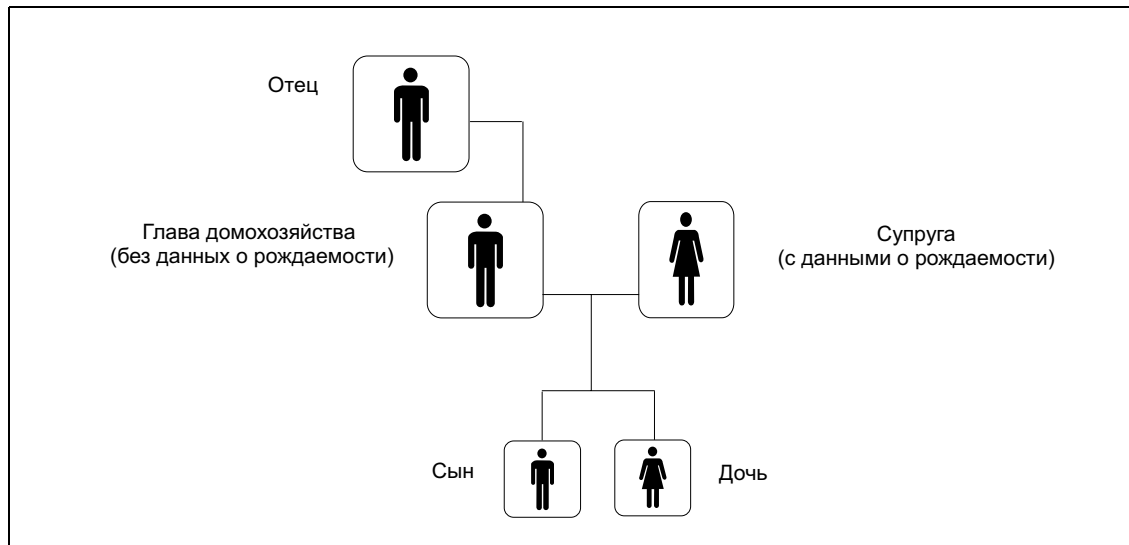
Рамка 2. Основные руководящие принципы исправления данных

После определения процедуры редактирования, возможно, полезно учитывать следующие рекомендации по исправлению данных:

- 1) вносить как можно меньше изменений в первоначальные зарегистрированные данные;
- 2) устранять очевидные несоответствия между ответами;
- 3) вводить данные для ошибочных или недостающих элементов, используя другие введенные данные о жилищной единице, лице или других лицах в домохозяйстве или в сопоставимой группе в качестве руководства, всегда соблюдая при этом конкретные процедуры. В некоторых случаях для определенных элементов подходит категория «не сообщается».

92. Рассмотрим следующую диаграмму (рисунок 1) для конкретного домохозяйства. Диаграмма показывает домохозяйство с согласующимися взаимосвязями и данными о поле. Глава домохозяйства — мужчина, и по нему нет информации о рождаемости; его супруга — женщина, и о ней есть соответствующая информация по рождаемости.

Рисунок 1. Типичное гипотетическое домохозяйство с указанием родства, пола и данных о рождаемости его членов



93. Однако во многих случаях информация непоследовательна. И тогда возникают следующие вопросы: каким должен быть процесс редактирования в отношении домохозяйства с несогласующимися данными? Как должна группа редактирования проводить свою работу, если глава домашнего хозяйства и его супруга зарегистрированы как лица мужского пола, как на рисунке 2? Раньше типичное правило редактирования содержало предпосылку о том, что первым лицом в паре является мужчина, особенно если это лицо является главой домохозяйства, а второе лицо, или супруга, является женщиной.

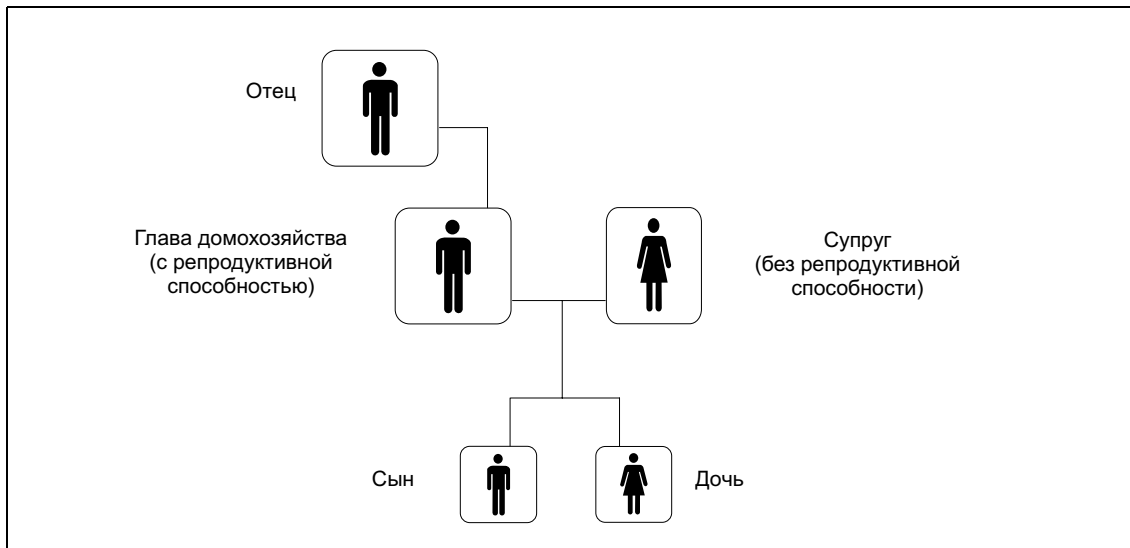
94. При этом если главой домохозяйства будет не муж, а жена, то принятое правило редактирования будет неправильным, и национальное бюро переписи/статистическое управление, совершит в конечном итоге четыре ошибки:

- пол главы домохозяйства будет указан неправильно;
- пол супруга будет указан неправильно;
- глава домохозяйства потеряет информацию о рождаемости;
- супругу мужского пола будет по ошибке приписана информация о рождаемости.

Таким образом, совершенно очевидно, что подобная процедура редактирования не является удовлетворительной.

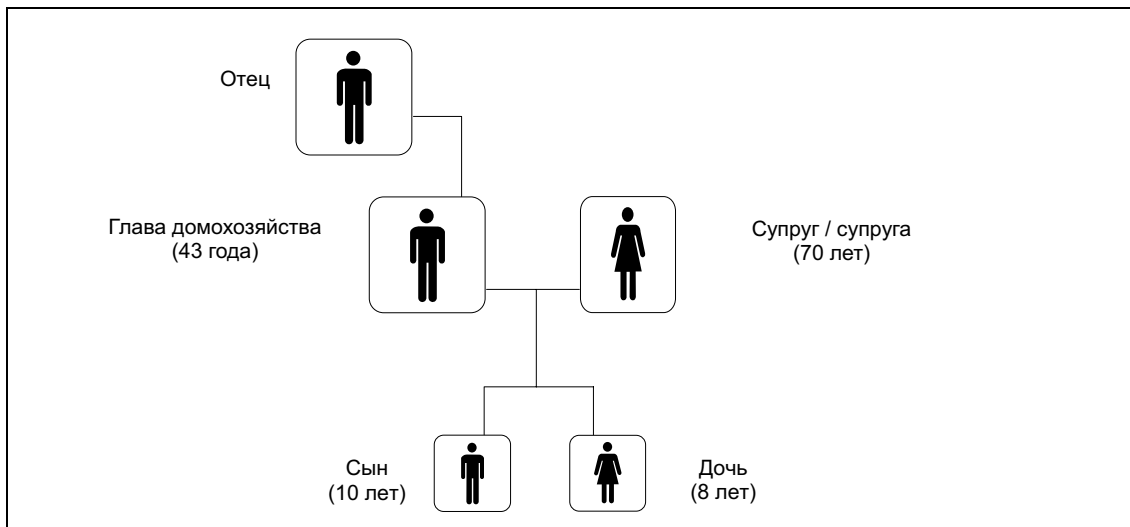
95. И наоборот, когда при удовлетворительной процедуре редактирования обнаруживается, что и глава и супруг одного пола, тогда производится проверка обоих лиц на рождаемость. Поскольку репродуктивной способностью обладает только глава домохозяйства, то глава рассматривается как лицо женского пола. В этом случае правила редактирования таких элементов удовлетворительны.

Рисунок 2. Пример домашнего хозяйства, где глава и супруг одного пола



96. Еще один пример, также иллюстрирующий этот вопрос, показан на рисунке 3. В большинстве стран считается, что детородный возраст составляет от 15 до 49 лет. Предположим, что женщина сообщает о рождении ребенка в 52 года, что подтверждается прямым свидетельством через номер строки, указанной для матери ребенка или расчетной разницей в возрасте (разница в возрасте между матерью и ребенком не может превышать 50 лет). Группа редактирования должна решить, является ли указанная разница в возрасте приемлемой, или ее надо заменить, поменяв при редактировании один или другой возраст. Если возрастной диапазон для того, чтобы иметь детей при редактировании увеличивается, и другие женщины заявляют о рождении детей в более позднем возрасте, то в набор данных могут быть внесены другие отклонения, если сам возраст сообщен неправильно. В этом случае группа редактирования снова должна решить вопрос о пригодности сообщаемых возрастов для конкретных переменных.

Рисунок 3. Пример домашнего хозяйства с указанием возраста некоторых членов домохозяйства

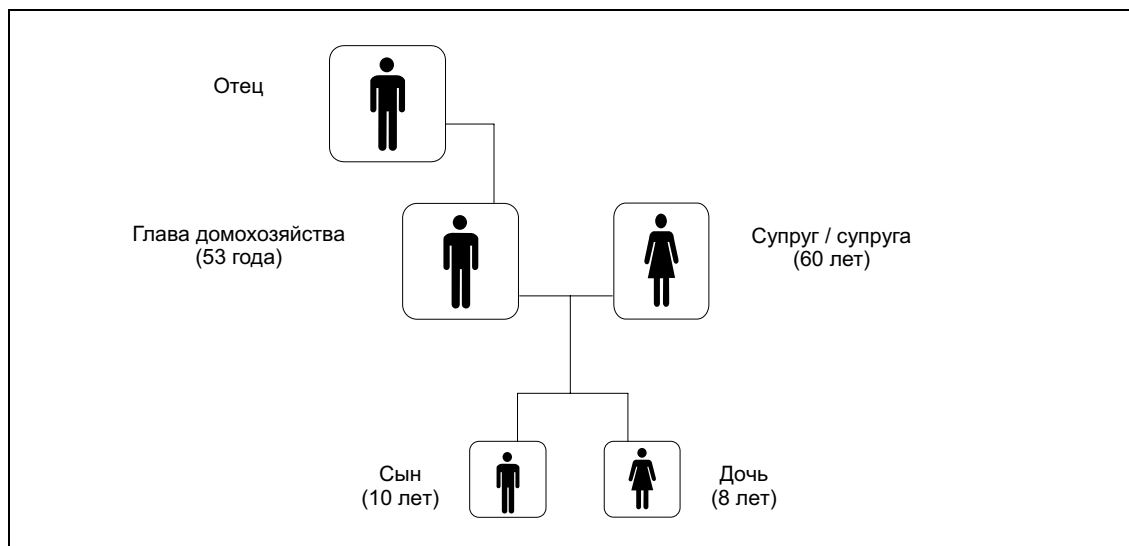


97. На рисунке 4 предлагается еще один возможный сценарий. Предположим, что при редактировании обнаруживается, что женщина в возрасте 70 лет имеет детей 10 и 8 лет, как на рисунке 4. Такая ситуация возможна потому, что муж мог бы иметь детей от предыдущей жены. При таких обстоятельствах дети относятся к главе домохозяйства, а не к супруге как таковой, даже если более вероятно, что операторы сделали ошибку, набрав «7» вместо «4», имея в виду 40 лет. Независимо от причины предположим, что специалисты, занимающиеся содержанием переписи, требуют от обработчика данных заменить возраст матери или ребенка, если разница между ними составляет более 50 лет. Такое требование ведет к еще более сложному редактированию. Если женщине 70 лет, а первому ребенку 10, то группа ре-

дактирования должна решить, чей возраст заменить. Группа редактирования может решить заменить возраст первого ребенка на 20 лет, и этим проблема для этого первого ребенка разрешится, или она может заменить возраст супруги и тогда остается проблема с возрастом второго ребенка, который также требует корректировки.

98. Если учитывать только возраст матери и одного ребенка, то при интерполяции возраст будет приписываться произвольно и это будет правильно наполовину. Однако если редактор посмотрит еще и на возраст мужа, то более вероятно, что он изменит возраст супруги, руководствуясь этой дополнительной информацией. Одно это изменение сделает возрастные показатели всей семьи более совместимыми.

Рисунок 4. Пример домашнего хозяйства с возможной несогласованностью указанных возрастных данных



С. ПРОВЕРКА НА ДОСТОВЕРНОСТЬ И СОГЛАСОВАННОСТЬ

99. Одно из основных требований к редактированию заключается в том, чтобы ни в одном элементе не было недостоверных величин. Кроме того, ответы по всем смежным элементам в записях и между ними должны согласовываться между собой.

100. При интерполяции нужно максимально учитывать всю информацию о смежных переменных одновременно, а не обязательно последовательно по отношению к смежным величинам. Однако в некоторых случаях редактор может провести проверку на согласованность до определения достоверности указанной величины. Если

интерполяция дает величину, исходя из проверки на согласованность, то необходимо сравнить эту величину с исходной записью, для того чтобы определить, есть ли действительное изменение. Если изменения нет, то исходная запись остается в прежнем виде.

101. Например, при редактировании данных о семейном положении, прежде всего, проверяется степень родства и если указано «супруг», если таковой имеется, и он или она не указали, что женат или замужем, то семейное положение указывается как «женат»/«замужем». До присвоения кода состоящим в браке программа проверяет, каким был исходный ответ. Если код для состоящих в браке уже есть, то программа не меняет указанное значение и никакой ошибки не возникает.

1. Нисходящий принцип редактирования

102. Эта процедура начинается с редакции первого элемента («верх»), который является первой переменной в вопроснике, с последовательным переходом к другим элементам вплоть до завершения редактирования всех данных. Хотя нисходящий принцип не отслеживает взаимосвязи между элементами данных, он тем не менее обеспечивает приемлемые рамки для завершения редактирования.

103. В процессе редактирования некоторые поправки изменяют значение того или иного элемента больше одного раза. При такой процедуре в набор данных могут быть внесены одна или более ошибок. Интерполированное значение может не согласовываться с другими данными. Даже если переменные обрабатываются последовательно, конкретную переменную следует по возможности редактировать одновременно с другими величинами. Например, возраст ребенка, исчисленный по возрасту матери, может не согласовываться с указанными школьными годами или годами проживания ребенка на данном счетном участке. В этом случае возраст будет исчисляться до тех пор, пока не исчезнет противоречие. Исчисленный возраст будет промежуточной переменной до окончательного закрепления. При редактировании интерполированные промежуточные переменные не следует регистрировать как изменения до окончательного закрепления.

104. Хотя в отношении некоторых элементов и условий программа редактирования может воспринять незаполненные графы или запись «не сообщается», смежная информация поможет воспроизвести большинство неуказанных данных или данных, указанных ошибочно. Полученные таким образом показатели могут быть правильными или неправильными в зависимости от случая. Однако широкие возможности и скорость компьютера при сопоставлении самых различных хранящихся данных позволяет определить замеченные значения, которые достаточно четко характеризуют ситуацию. Полученные в результате этого таблицы в большинстве случаев будут содержать меньше противоречий, чем те, которые составлены по неотредактированным записям или записям, в которых все неприемлемые значения интерполируются в «не сообщается».

105. Программа редактирования может производить также структурную проверку (см. главу III). Она должна проверять элементы, относящиеся к населению (см. главу IV), и элементы, касающиеся жилья (см. главу V). Кроме того, при редактировании, вероятно, следует создавать одну или несколько повторно закодированных переменных по отдельной записи, необходимой для табулирования, как отмечается в приложении I.

106. Чрезвычайно важно избегать кругового редактирования, сначала меняя один или несколько элементов, а затем на каком-то более позднем этапе возвращая их к прежнему значению. В данном руководстве отмечается, что сотрудники должны делать несколько прогонов с тем, чтобы быть уверенными в том, что все элементы полностью отредактированы. Можно выработать критерии редактирования для изменения данных в ходе первого прогона, но при этом если их применять к измененным данным при втором прогоне, то произойдет возвращение к первоначальной конфигурации. Такая процедура может продолжаться в течение многих прогонов. Группа редактирования должна избегать введения таких критериев в процесс редактирования.

2. Редактирование по принципу групповых переменных

107. Нисходящий принцип редактирования данных переписи и обследования, представленный в разделе 1, выше, не всегда дает лучшие результаты, то есть те, которые ближе всего соответствуют реальному распределению переменных. Как уже говорилось, нисходящий принцип применяется без особой предосторожности и часто вызывает проблемы при редактировании.

108. Другой принцип заключается в редактировании групповых переменных и основан на системе Фелледжи-Холта. Этот принцип требует больше специальных знаний и более высокой производительности компьютера, но может принести результаты более близкие к «реальности». Различные виды редактирования групповых переменных показаны в приложении V «Методы интерполяции». В системе редактирования групповых переменных необходимо определить набор позитивных представлений и взаимосвязь между переменными. Затем редактор проверяет каждое представление по отношению к данным о домашнем хозяйстве, чтобы посмотреть, все ли представления верны. По каждому ложному представлению редактор отследит элемент за элементом все недостоверные данные или несоответствия. После всех проверок система редактирования и интерполяции должна оценить, как лучше всего изменить запись с тем, чтобы она прошла по всем редактируемым данным. Группы редактирования, как правило, применяют принцип наименьших изменений и меняют минимально возможное количество переменных для получения приемлемой записи.

109. Одиннадцать декларативных представлений на рисунке 5 служат примером правил, которые можно применять при редактировании отдельных характеристик населения по групповому принципу. В этом примере глава домашнего хозяйства должен быть на 15 лет старше. Для обобщенных данных было бы лучше использовать «X» лет, где X — установленный минимум

для данной страны. Такие представления в этом примере, как степень родства, пол, возраст, семейное положение и рождаемость, сосредоточены на других важных первичных переменных. Эти переменные тесно взаимосвязаны и поэтому группы редактирования должны

рассматривать их вместе, с тем чтобы отредактировать данные наиболее эффективным образом. Здесь следует отметить, что хотя все переменные имеют большое значение, некоторые из них важнее для представления данных, чем другие.

Рисунок 5. Пример правил для редактирования отдельных характеристик населения по принципу групповых переменных

Номер	Правило	Степень родства	Пол	Возраст	Семейное положение	Рождаемость
1	Глава домохозяйства должен быть в возрасте 15 лет или старше					
2	Супруг(а) должен быть в возрасте 15 лет или старше					
3	Супруг(а) должен состоять в браке					
4	Если есть супруг(а), глава домохозяйства должен состоять в браке					
5	Если есть супруг(а), глава домохозяйства и супруг(а) должны быть противоположного пола	1	1			
6	Лицо младше 15 лет никогда не должно состоять в браке					
7	Лицо мужского пола не должно иметь рождаемости		1			1
8	Лицо женского пола младше 15 лет не должно иметь рождаемости					
9	Для лица женского пола в возрасте 15 лет или старше графа «рождаемость» должна быть заполнена					
10	Ребенок должен быть младше главы домохозяйства					
11	Родитель должен быть старше главы домохозяйства ВСЕГО	1	2			1

110. В примере, показанном на рисунке 6, оба супруга относятся к той же совокупности, что и на рисунке 5. Оба зарегистрированы как лица мужского пола. Здесь процедура редактирования простая и открытая. Переменная с наибольшим количеством дублированных ошибок должна быть исправлена в первую очередь. На рисунке 6 программа-редактор осуществляет процедуру интерполяции для признака «пол», поскольку, если основываться на данных рисунка 5, эта переменная больше всего содержит ошибок по

отношению к 1) родству и полу и 2) рождаемости и полу. Когда программа-редактор проверяет рождаемость и обнаруживает, что в отношении главы домохозяйства информация о рождаемости присутствует, а в отношении супруга нет, то методом интерполяции признак «женский» присваивается главе домохозяйства. И, наконец, когда группа редактирования вновь проверяет ряды повторений и все позитивные представления верны, дальнейшего редактирования не требуется.

Рисунок 6. Пример с главой и супругом одного пола в неотредактированном наборе данных и его разрешение

Лицо	Степень родства	Пол	Дети, родившиеся живыми
<i>Неотредактированные данные</i>			
1	Глава домохозяйства	Мужской	03
2	Супруг	Мужской	Не заполнено
<i>Данные после редактирования графы «пол»</i>			
1	Глава домохозяйства	Женский	03
2	Супруг	Мужской	Не заполнено

111. Спецификации редактирования для этой программы можно записать, как показано на рисунке 7. Если признак рождаемости заполнен для обоих супругов, то редактор сработает. Однако очевидно, что редактирование будет неполным, поскольку программа просматривает только те случаи, где точно указана рождаемость, как для главы домохозяйства, так и для супруга.

Рисунок 7. Выборочные спецификации редактирования для исправления переменной пола с символическим кодом

Если ПОЛ ГЛАВЫ ДОМОХОЗЯЙСТВА = ПОЛ СУПРУГА
 Если «РОЖДАЕМОСТЬ» ГЛАВЫ ДОМОХОЗЯЙСТВА заполнена
 Если «РОЖДАЕМОСТЬ» СУПРУГА не заполнена
 (если ПОЛ главы домохозяйства уже не женский) сделать ПОЛ = женский
 (если ПОЛ супруга уже не мужской) сделать ПОЛ = мужской
 Еще *Сделать что-нибудь еще, потому что у них один и тот же пол и у обоих указана рождаемость!!!*
 Закончить, если
 Еще *Это тот случай, когда графа «рождаемость» главы домохозяйства не заполнена*
 Если «РОЖДАЕМОСТЬ» СУПРУГА заполнена
 (если ПОЛ главы домохозяйства уже не мужской) сделать ПОЛ = мужской
 (если ПОЛ супруга уже не женский) сделать ПОЛ = женский
 Еще *Сделать что-нибудь еще, если у ОБОИХ не указана рождаемость!!!*
 Закончить, если
 Закончить, если

112. На рисунке 8 показан пример, когда при редактировании согласно введенной с клавиатуры информации получается, что главе домашнего хозяйства женского пола 13 лет, она вдова, но с тремя детьми. Когда программа прочитывает правила редактирования, то получается следующее:

Рисунок 8. Пример анализа данных об очень молодой вдове с тремя детьми, отредактированных по принципу групповых переменных

Номер	Правила	Степень родства	Пол	Возраст	Семейное положение	Рождаемость
1	Глава домохозяйства должен быть в возрасте 15 лет или старше	1		1		
2	Супруг(а) должен быть в возрасте 15 лет или старше					
3	Супруг(а) должен состоять в браке					
4	Если есть супруг(а), глава домохозяйства должен состоять в браке				1	
5	Если есть супруг(а), глава домохозяйства и супруг(а) должны быть противоположного пола	1				
6	Лицо младше 15 лет никогда не должно состоять в браке			1	1	
7	Лицо мужского пола не должно иметь рождаемости					1
8	Лицо женского пола младше 15 лет не должно иметь рождаемости		1	1		
9	Для лица женского пола в возрасте 15 лет или старше графа «рождаемость» должна быть заполнена					
10	Ребенок должен быть младше главы домохозяйства					
11	Родитель должен быть старше главы домохозяйства					
	ВСЕГО	2	1	3	2	1

113. Если переменная возраста основана на ряде позитивных представлений, то она в большинстве случаев ошибочна, и должна быть исправлена в первую очередь. После исправления возраста редактирование будет закончено, если это устраняет все случаи несогласованности. В противном случае программа будет редактировать эту переменную при последующем наибольшем количестве несоответствий.

D. ВОПРОСЫ КОДИРОВАНИЯ

114. Когда национальные бюро переписи/статистические управления составляют перечень кодов для программы редактирования для последующего составления таблиц, они могут по своему усмотрению установить общие коды для нескольких элементов. Например, во многих странах коды места (место рождения, место рождения родителей, прежнее местожительство, место работы), языка, этнической/расовой принадлежности и гражданства очень похожи. В связи с этим можно было бы разработать общую схему кодирования в виде трехцифрового кода, где первая цифра означала бы континент, вторая — регион и третья — конкретную страну. Национальные бюро переписи/статистические управления могут также использовать цифровые коды страны, разработанные международными организациями, например, статистическим отделом Организации Объединенных Наций (Организация Объединенных Наций, 1999 год). Набор общих кодов для тесно взаимосвязанных переменных может сократить ошибки кодирования и помочь обработчикам данных во время редактирования. Общие коды позволяют также обработчикам данных перемещать в необходимых случаях значения из одного элемента в другой.

115. Структура списков кодирования может облегчить процесс кодирования, а также дальнейшую обработку данных во время редактирования, составления таблиц и проведения анализа. Для больших стран, где много иммигрантов или этнических групп, коды составленные по континентам, регионам и странам с разными кодами или цифрами по каждому показателю для простого перечня предпочтительнее.

116. На рисунке 9 показаны примеры общих кодов для таких элементов, как место рождения, гражданство, язык и этническая принадлежность. Для Филиппин коды говорящих на тагальском и илоканском языках отличаются от общего кода филиппинских языков в зависимости от конкретной ситуации в стране и, кроме того, эти коды могут различаться и между собой. Хотя английский язык имеет один код, на нем говорит не одна группа населения. Коды для места рождения, гражданства и этнической принадлежности в Канаде и Соединенных Штатах отличаются незначительно. Для лиц, родившихся во Франции, имеющих французское гражданство, говорящих на французском языке и являющихся французами по этнической принадлежности используется один и тот же код. Таким образом, если не достает одного из этих элементов и если группа редактирования сочтет такое решение приемлемым, обработчик данных может взять код в одной из других граф.

117. Если в группе вопросов в переписном листе они зависят один от другого, то вероятно сотрудники, занимающиеся переписью/обследованием не должны задавать все эти вопросы. Группа редактирования должна решить в каждом конкретном случае, когда использовать элементы непосредственно для присвоения, а когда использовать другие имеющиеся переменные.

Рисунок 9. Примеры общих кодов для отдельных элементов

<i>Группа</i>	<i>Место рождения</i>	<i>Гражданство</i>	<i>Язык</i>	<i>Этническая принадлежность</i>
Франция/французский	10	10	10	10
Испания/испанский	20	20	20	20
Латинская Америка	25	25	20	25
Филиппины/филиппино	30	30	30	
илоканский			32	
тагальский			32	
Англия/английский	40	40	40	40
Канада	50	50	40	50
США	52	52	40	52

118. Еще одна проблема возникает, когда определения между переписями (или между переписью и обследованием) отличаются для таких переменных, как работа или этническая принадлежность. Национальное бюро переписи/статистическое управление должно решить, как учитывать такие изменения как в отношении редактируемых данных данной переписи, так и в отношении наборов данных из предыдущей переписи для того, чтобы показать тенденции. Если есть исходные, неотредактированные данные, обработчики данных могут вносить изменения в соответствующие элементы и сделать для всех повторный прогон.

119. Например, европейская страна может использовать единый код страны происхождения для всех южно-азиатских стран, если отмечено только несколько случаев. Однако в связи с изменениями характера миграции при следующем обследовании или переписи могут потребоваться отдельные коды для Индии, Бангладеш, Пакистана, Шри-Ланки и других южно-азиатских стран на протяжении всего процесса обработки.

Е. МЕТОДЫ ИСПРАВЛЕНИЯ И ИНТЕРПОЛЯЦИИ ДАННЫХ

120. Как отмечалось выше, незаполненные графы в записях данных таких как «не сообщается», «не известно» или иная недостающая информация встречаются во всех переписях и обследованиях. Недостоверные данные вносятся также респондентом, счетчиком или возникают в результате ошибок при вводе данных. Методы корректирования бывают самыми разными и зависят от элемента. В большинстве случаев элементам данных можно присвоить достоверные коды с достаточной уверенностью, что они правильные, используя при этом ответы из других элементов данных в записях о данном лице или домохозяйстве или в записях о других домохозяйствах или лицах.

121. В данном *Руководстве* представлены два автоматизированных метода исправления ошибочных данных. Один из них — метод статической интерполяции, или метод *cold deck* — используется главным образом для недостающих или неизвестных элементов. Другой метод — метод динамической интерполяции, или метод *hot deck*³ — можно применять для недостающих данных, а также для несогласующихся или недостоверных элементов.

³ Результаты ответов на вопросники, полученные статистическим отделом Организации Объединенных Наций примерно из 130 стран и регионов, показывают, что термины *hot deck* и *cold deck*, как правило, неизвестны. Тем не менее, некоторые страны применяют эти принципы для редактирования в своих данных.

1. Метод статической интерполяции (*cold deck*)

122. При статической интерполяции программа редактирования присваивает конкретный ответ по недостающему элементу из предварительно установленного набора или ответ исчисляется пропорционально из распределения достоверных ответов. При такой методике программа не обновляет первоначальный набор переменных. Значения не меняются по сравнению с теми, которые содержатся в исходной статической матрице после обработки записей для первого, второго, десятого или любого иного лица. Исходные значения позволяют интерполировать любые недостающие данные.

123. Статическая интерполяция так же, как и динамическая интерполяция, представляет собой стохастический метод, но значения со временем не меняются. Этот принцип описан в приложении V.

124. Иногда при статической интерполяции используется метод соотношений, когда ответы присваиваются исходя из заранее установленных пропорций. В качестве примера пропорционального распределения ответов представим себе таблицу достоверных данных, то есть данных, заполненных в противоположность недостающим элементам о времени работы в неделю мужчин в возрасте 33 лет, которые были заняты в сельском хозяйстве, где показано, что 25% работали 50 часов в неделю; 40% — 60 часов в неделю; и 35% — 70 часов в неделю. Недостающие или недостоверные ответы о времени работы мужчин в возрасте 33 лет в сельском хозяйстве с заменой 25% времени на 50 часов, 40% времени на 60 часов и 35% времени на 70 часов. Однако если нет достоверных данных из предыдущих переписей, обследований или других источников, этот метод требует предтабулирования достоверных ответов из проводимой переписи, что может быть нереально с экономической и оперативной точек зрения.

2. Метод динамической интерполяции (*hot deck*)

125. Еще один способ избавления набора данных от неизвестных величин заключается в динамической интерполяции, который используется для распределения значений для отсутствующих, неизвестных, неправильных или несогласующихся позиций. Первоначально этот метод был разработан бюро переписи США, но в последствии другие агентства его усовершенствовали. При динамической интерполяции используется одна или более переменных для оценки вероятного ответа, когда в наборе данных появляется неизвестная величина (или в некоторых случаях несколько неизвестных величин). Динамическая интерполяция становится все более популярной при редактировании данных переписи, поскольку ее легко использовать и получать при

этом чистые, воспроизводимые результаты. Кроме того, устраняя неизвестные величины, легче проследить тенденции между переписями и обследованиями, поскольку аналитику не надо иметь дело с неизвестными величинами в каждом конкретном случае.

126. При динамической интерполяции известные данные об отдельных лицах с аналогичными характеристиками дают наиболее приемлемую информацию, которую можно использовать, если неизвестна какая-то часть информации о другом человеке. К этим характеристикам относятся пол, возраст, степень родства с главой домохозяйства, экономический статус и образование. Сама матрица интерполяции представляет собой набор величин, аналогичный колоде перфокарт. На этих матрицах хранится, а затем выдается информация, используемая в тех случаях, когда встречаются неизвестные величины. Эта колода постоянно меняется путем обновления данных и/или логического «тасования колоды» с тем, чтобы менять интерполированные величины во время обработки данных; отсюда и термин «hot deck» — «горячая колода».

127. Это можно просто проиллюстрировать тем, что единственная величина может храниться в виде колоды. Например, если пол лица по какой-то причине указан неправильно, колоде присваивается произвольное исходное значение (мужской или женский), определяя, таким образом, исходную величину. Начальное значение становится показателем пола первого, кто встретится с пометкой «пол не известен». Однако если пол первого лица указан правильно, то пол первого лица заменяет случайное значение. Если неизвестен пол второго лица, то тогда матрица интерполяции присваивает пол, который на ней хранится. В этом случае исчисленный пол — это пол первого лица. В сущности, если программа-редактор находит приемлемую величину для того или иного элемента, она закладывает ее в матрицу интерполяции. При обнаружении неприемлемой величины она заменяется путем интерполяции достоверным значением из матрицы интерполяции.

128. Одна из проблем, которая встречается при описываемой здесь динамической интерполяции (*hot deck*) заключается в том, что если два различных элемента имеют неизвестные значения, то один и тот же «донор» не может быть использован для присвоения достоверных ответов. Каждая величина может исходить от «реального» лица, но это могут быть разные люди. Лучше всего было бы присваивать обе переменные одновременно от одного и того же лица. Однако программирование этих сложных матриц может представлять некоторые трудности.

129. Приводимые ниже данные (рисунок 10) показывают домохозяйства или набор из десяти индивидов. Числа 9 и 99 для пола и возраста показывают недос-

тающую информацию. Хотя для использования при интерполяции есть и другие переменные, такие как образование и род занятий, они в этот маленький пример не включены.

Рисунок 10. Выборочные домохозяйства в качестве входных данных для динамической интерполяции

Идентификационный номер	Степень родства	Пол	Возраст
1	1	1	39
2	2	2	35
3	3	1	13
4	3	9	10
5	4	2	40
6	4	1	99
7	4	2	13
8	5	9	99
9	5	1	44
10	5	2	36

ПРИМЕЧАНИЕ: 9, 99 = недостающая информация.

130. Если исходное значение для матрицы интерполяции SEXARRAY «мужской» (код =1) матрица интерполяции будет выглядеть примерно следующим образом: SEX = 1.

131. После обработки информации о лице 1 значение останется 1. Однако после обработки информации о втором лице значение изменится на 2, поскольку это лицо женского пола. Тогда переменная будет выглядеть следующим образом: SEXARRAY = 2.

132. Для каждого правильного значения пола того индивида, данные о котором обрабатываются в данный момент, код для пола этого лица заменяет значение в матрице интерполяции. Когда обрабатываются данные о третьем лице, значение при интерполяции снова меняется на 1 или «мужской».

133. Пол четвертого лица не известен, поэтому редактор смотрит на значение матрицы интерполяции и заменяет неизвестное значение значением с матрицы интерполяции. Лицо 5 женского рода, поэтому редактор заменяет предыдущее значение в матрице интерполяции по данным лица 3 (мужской). Такой процесс продолжается вплоть до лица 8.

134. При редактировании опять используется метод интерполяции, и лицо 8 становится лицом женского пола, поскольку значение в матрице интерполяции в отношении лица 7 указывает на женский пол. Матрица интерполяции используется при редактировании для получения величин дважды: один раз для получения значения «мужской» и один раз для получения значе-

ния «женский». Поскольку показатели пола появляются примерно с одинаковой частотой, то во время длинного прогона каждый пол используется при интерполяции в пропорции примерно 50 на 50. После обработки данных по всем десяти индивидам переменная будет выглядеть следующим образом: SEXARRAY = 2.

135. Матрица интерполяции приписывает пол именно таким образом, но при этом существуют и другие, более сложные способы применения такой процедуры. Например, программа-редактор может использовать отношение к главе домохозяйства и пол для того, чтобы было легче определить возраст индивида. Рассмотрим следующий частичный список кодов отношений:

1 = Глава домохозяйства

2 = Супруг/а

3 = Ребенок

4 = Другой родственник

5 = Не родственник

136. Обработчик данных может создать исходные значения возраста, которые могут приближаться к реальной ситуации, касающейся отношений по признаку пола. Эти значения не столь важны, поскольку редактор почти наверняка заменит их, прежде чем использовать. Редакция требует интерполяции многих значений, поэтому лишь немногие исходные величины влияют на окончательные таблицы. Эти величины могут быть такими, как показано на рисунке 11.

Рисунок 11. Исходная статическая матрица для возраста, исчисляемого по полу и степени родства

	Степень родства				
	Глава домохозяйства (1)	Супруг(а) (2)	Сын/дочь (3)	Другой родственник (4)	Не родственник (5)
Мужской (1)	35	35	12	40	40
Женский (2)	32	32	12	37	37

137. Опять рассмотрим десять индивидов, представленных на рисунке 10. Поскольку первое лицо в нашей выборке записано как глава домохозяйства (код = 1) и он мужчина (код = 1) его возраст (39) заменяет при интерполяции первый элемент (координаты 1,1). При этом в колоде находятся величины, указанные на рисунке 12.

Рисунок 12. Пример матрицы динамической интерполяции после одного изменения

	Степень родства				
	Глава домохозяйства (1)	Супруг(а) (2)	Сын/дочь (3)	Другой родственник (4)	Не родственник (5)
Мужской (1)	39	35	12	40	40
Женский (2)	32	32	12	37	37

138. Второе лицо является супругой (код = 2) и женщиной (код = 2), поэтому ее возраст (35) заменяет значение во втором ряду во второй колонке с заменой колоды на эти значения. Аналогичным образом возраст других индивидов в домохозяйстве заменяет через пятое лицо значения в матрице интерполяции.

139. Отметим, что во время предыдущей процедуры интерполяции пола, пол 1 был приписан лицу 4. Ввиду того, что программа-редактор требует интерполяции

значения для пола, она не вносит в массив возраст этого лица в качестве обновления. Программа будет обновлять массив только теми значениями из записей, где пол и степень родства первоначально указаны правильно. Однако когда программа доходит до лица 6, она обнаруживает, что возраст не известен. Данное лицо мужского пола не является другим родственником главы домохозяйства. Исходя из этого, программа редактирования использует элемент матрицы интерполяции для лиц мужского пола, которые входят в группу «дру-

гой родственник» (четвертая колонка в первом ряду) и приписывает значение возраста для этой категории («другой родственник мужского пола» — в данном случае — 40).

140. Ни пол, ни возраст восьмого лица не указаны. Редактор исчисляет пол как женский, а затем распределяет возраст, исходя из этого выделенного пола и кода степени родства (5). В этом случае возраст составляет 37 лет.

141. Хотя программа редактирования исчислила возраст по известной степени родства, она пользовалась заранее выделенным значением пола для других переменных. В данном случае применение распределенных

величин для дальнейшей интерполяции является примером слабой процедуры редактирования (см. раздел 3 *d*, ниже). Здесь для интерполяции лучше поискать другие известные элементы данных, например, семейное положение.

142. После проверки десятого лица значения матрицы интерполяции даны на рисунке 13. В этом примере в обоих случаях исчисления использовалась исходная статическая матрица. Как правило, при интерполяции используется лишь небольшое число, если вообще используется, исходных величин. В большинстве же случаев используются величины, взятые из данных о зарегистрированном населении.

Рисунок 13. Пример матрицы динамической интерполяции после многократных изменений

	Степень родства				
	Глава домохозяйства (1)	Супруг/а (2)	Сын/дочь (3)	Другой родственник (4)	Не родственник (5)
Мужской (1)	39	35	13	40	44
Женский (2)	32	35	12	13	36

3. Вопросы динамической интерполяции (hot deck)

а) Вопросы территориальности

143. Если в программе редактирования для исчисления недостающих величин используется динамическая интерполяция, то следует попытаться использовать данные, отсортированные по наименьшему территориально определенному району. С такой процедурой должна увеличиться вероятность получения правильного ответа, поскольку те, кто проживает в одном и том же небольшом географическом районе, как правило, имеют более или менее однородные демографические и прочие характеристики. Если совокупность населения не однородна, то корреляции не будет, и поэтому группа редактирования должна рассматривать переменные в каждом конкретном случае.

б) Использование смежных элементов

144. Прежде чем прибегать к динамической интерполяции для получения недостающих величин следует попытаться использовать смежные элементы, чтобы присвоить значение, которое, вероятно, будет правильным. Например, если не указано семейное положение лица, то программа-редактор определит, есть ли у этого лица супруг(а) в домохозяйстве. Если так, то программа присвоит код для лица, состоящего в браке, не при-

бегая к матрице интерполяции. Однако если такого признака нет, то программа, возможно, будет полагаться на значение из матрицы интерполяции.

с) Как порядок переменных влияет на матрицы

145. Национальные бюро переписи/статистические управления, пользующиеся матрицами интерполяции, должны посмотреть, какие переменные им нужны, поскольку они устанавливают порядок их редактирования. Для элементов, касающихся населения, они захотят в первую очередь отредактировать данные о поле и возрасте с тем, чтобы их можно было использовать в других матрицах интерполяции. При общем редактировании не следует использовать в матрицах интерполяции неотредактированные переменные, хотя большинство пакетов компьютерных программ не отторгнет «неизвестные» ряды или колонки. В последующих матрицах интерполяции можно использовать эти элементы данных после редактирования. Однако по возможности следует исключать отредактированные данные из матрицы интерполяции.

146. Например, если программа-редактор исчисляет возраст по полу и степени родства, то элементы массива для этой матрицы интерполяции (пол по степени родства) не следует обновлять, если пол или степень родства были интерполированы. Как правило, только в

том случае, если возраст, пол и степень родства достоверны и согласуются между собой, пакет редактирования должен вводить возраст в элемент для соответствующего пола и степени родства. Однако иногда в силу других факторов использование отредактированных данных неизбежно.

d) *Сложность матриц интерполяции*

147. Делая матрицу интерполяции более детализированной, национальное бюро переписи/статистическое управление повышает вероятность получения согласующейся, «правильной» величины в матрице интерполяции. Например, программа может интерполировать семейное положение, используя только степень родства. Однако с возрастом увеличивается вероятность вдовства или развода. В связи с этим имеет смысл интерполировать семейное положение по возрасту и степени родства. Программа-редактор, которая использует возраст и степень родства, берет значение семейного положения от лица с теми же характеристиками в непосредственно предшествующей этому достоверной записи, хранящейся в матрице интерполяции.

148. Тем не менее, описанная выше процедура, может создать новые проблемы. Национальное бюро переписи/статистическое управление, как правило, редактирует элементы переписного листа в определенной последовательности, когда в случае применения нисходящего принципа возраст редактируется после семейного положения. В этом случае, если в записи нет данных о семейном положении и возрасте, то нельзя взять значение семейного положения непосредственно из предшествующей записи с такими же данными о возрасте и степени родства. В результате программа может не определить возрастную категорию для этой записи. Эту проблему можно решить по-другому, если в матрице интерполяции будет ряд или колонка для «не сообщенных» элементов. В этом случае программа сможет обозначать значение для семейного положения, используя категорию семейного положения непосредственно из предшествующей записи с теми же категориями родства и возраста с указанием «не сообщается». Однако против такого подхода можно привести два фактора. Во-первых, случай «не сообщается» в одной и той же комбинации настолько редкий, что будет трудно обновлять матрицу интерполяции для недостающего элемента. Во-вторых, по существу нельзя получить методом динамической интерполяции соответствующие величины для этих комбинаций значений «не известно», поскольку они не существуют в «реальном» мире.

149. Разрешение вышеописанной проблемы создает больше работы для обработчика данных, но дает в результате более чистый со статистической точки зрения

продукт. Программа редактирования сначала ведет проверку с тем, чтобы определить есть ли у этих элементов правильные коды. Если запись для данного лица не имеет правильного кода для этого элемента, этот элемент для данной записи в матрице интерполяции не используется. Обработчики данных могут облегчить этот процесс, создав более простую матрицу интерполяции. В продолжение ранее приведенного примера можно отметить, что если программа должна интерполировать семейное положение из-за того, что данная величина отсутствует, матрица интерполяции будет, как правило, иметь два измерения: возраст и степень родства. Если после проверки программа не найдет правильного кода для возраста, она будет интерполировать семейное положение только по степени родства. Поскольку данные о степени родства редактируются до информации о семейном положении, код родства будет достоверным. Те же принципы используются в программе для всех процедур динамической интерполяции.

e) *Разработка матрицы интерполяции*

150. Сотрудники, занимающиеся содержанием переписи, в сотрудничестве с обработчиками данных могут также подготовить соответствующие матрицы интерполяции (некоторые группы редактирования пользуются многократными матрицами интерполяции). Матрицы интерполяции обновляются только достоверными ответами; группы редактирования не пользуются распределенными или интерполированными величинами. Специалисты, занимающиеся содержанием переписи, и обработчики данных должны проверять спецификации редактирования и полученные методом динамической интерполяции данные на согласованность и полноту.

151. На разработку матрицы интерполяции нужно потратить немало времени и умственных усилий, включая научные изыскания для изучения вопроса об использовании административных данных и результатов предыдущих переписей или обследований. Но даже после научно-изыскательских работ редакторам не следует применять матрицы интерполяции произвольно. Когда в матрицах интерполяции присутствует внутренняя несогласованность, требуются значительные усилия для их согласования. Если в матрицах интерполяции не используются стандартные условные обозначения, то необходимо рассматривать каждую из них отдельно.

152. В примерах, приводимых в данном *Руководстве*, каждый элемент в матрицах интерполяции имеет одно значение, однако некоторые группы редактирования оставляют не одну возможность для каждого элемента. Эти элементы дают дополнительное измерение. Например, если возраст всех детей в семье не известен, скажем, в семье с четырьмя мальчиками, компьютер не

будет присваивать одно и то же значение четыре раза, создавая четверых близнецов. Вместо этого будут указаны четыре разных возраста. Однако даже здесь одно и то же значение может быть присвоено более одного раза в зависимости оттого, что хранится в матрицах.

f) *Стандартизированные матрицы интерполяции*

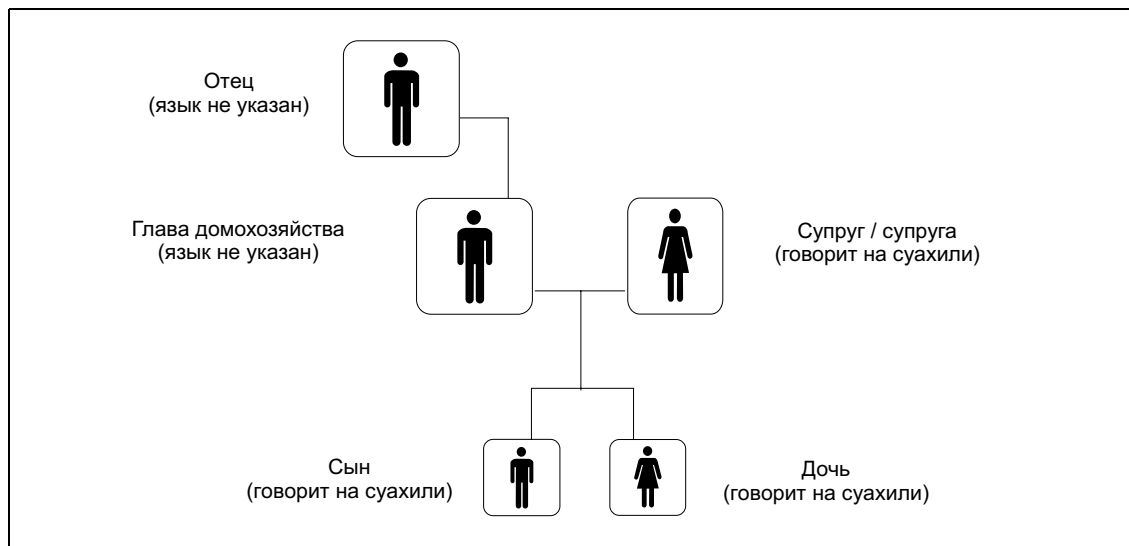
153. С помощью стандартизированных матриц интерполяции можно рационализировать процесс редактирования. Матрицы интерполяции со стандартными измерениями различных социальных и экономических переменных, таких как возрастная группа и пол, можно быстро проверить и затем использовать.

154. Например, национальное бюро переписи/статистическое управление может по своему усмотрению разработать матрицу интерполяции с тем, чтобы определить код языка при отсутствии данных. Почти навер-

няка программа-редактор проверит домашнее хозяйство, где проживает другое лицо, зарегистрированное как говорящее на данном языке. В случае неудачи программа может выбрать язык предыдущего индивида того же пола из той же возрастной группы (обновив матрицу интерполяции, если все три элемента достоверны). Такая процедура покажет вероятный язык, поскольку лица, говорящие на одном и том же или аналогичном языке, как правило, территориально находятся близко друг от друга.

155. На рисунке 14 переменная «язык» не содержит никакой информации в отношении главы домохозяйства. По той или иной причине оптический сканер или оператор могут не уловить информацию о языке или код или что-то другое может пойти неправильно. Однако поскольку супруг и дети все говорят на суахили, этот язык можно присвоить главе домохозяйства и отцу главы домохозяйства, чьи данные о языке тоже отсутствуют.

Рисунок 14. Пример главы домохозяйства и отца главы без указания языка



156. Если язык кого-либо из членов домашнего хозяйства не сообщается, программа-редактор должна сделать что-то еще. Во-первых, редактор ищет другие переменные для получения косвенной оценки используемого языка. Иногда на соответствующий исчисляемый язык указывает расовая или этническая принадлежность или место рождения. Если такой идентификатор имеется, то группа редактирования может его использовать для определения языка главы домохозяйства. Если нет, то редактор может использовать для интерполяции возраст и пол. Матрица интерполяции может выглядеть примерно как на рисунке 15.

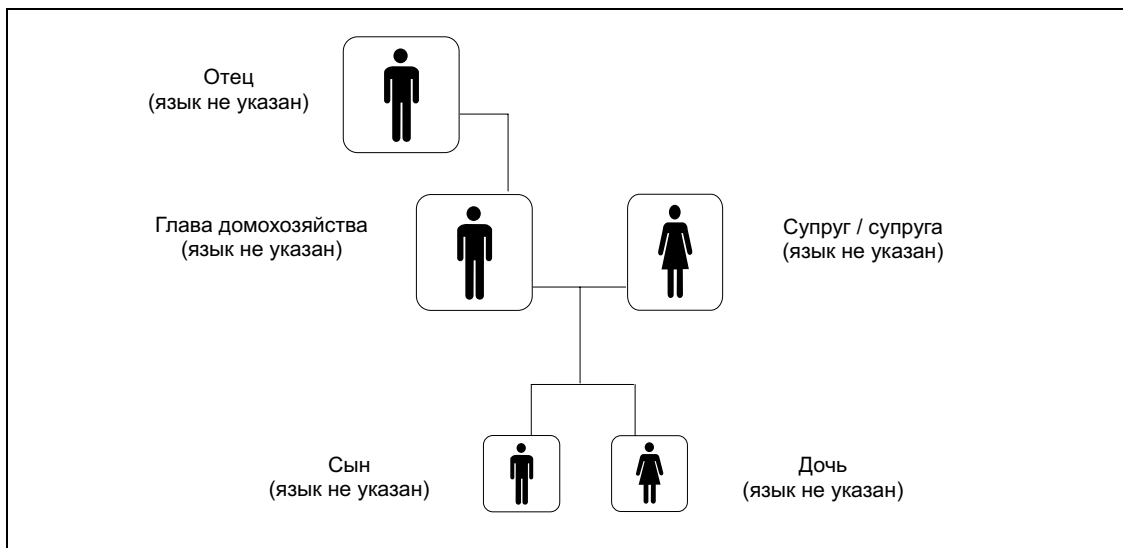
Рисунок 15. Исходные значения для матрицы, предназначенные для динамической интерполяции языка

Пол	Возраст					
	Младше 15 лет	15–29 лет	30–44 года	45–59 лет	60–74 года	75 лет и старше
Мужской	Язык 1	Язык 1	Язык 1	Язык 1	Язык 1	Язык 2
Женский	Язык 1	Язык 1	Язык 1	Язык 1	Язык 1	Язык 2

157. Если будет принято решение об интерполяции, то программа присвоит главе домохозяйства язык исходя из возрастной группы и пола. В этом случае позиции в матрице интерполяции будут касаться только предыдущих глав домохозяйства, поскольку все остальные лица в данном домохозяйстве получают тот же код языка, что и глава домохозяйства.

158. Здесь, если в домохозяйстве по-прежнему нет никого, кто указывает, что говорит на определенном языке, программа редактирования использует матрицу интерполяции для присвоения языка главе домохозяйства, исходя из возраста и пола главы домохозяйства. Присвоенный язык является в файле данных самым последним, на котором говорит другая глава домохозяйства того же возраста и пола. Поскольку матрица интерполяции «обновляется» постоянно, как только возникают приемлемые случаи, присвоенный язык, скорее всего, будет тем языком, на котором говорит данная общность.

Рисунок 16. Пример членов домохозяйства без указания языка



159. Исключения из правил редактирования будут делаться в самом начале редакторского прогона. Сотрудники должны внимательно следить за изменениями языка, которые могут происходить, когда они перемещаются из одного географического района в другой. Некоторые страны должны также следить за смешением языков с ограниченным распространением. Однако даже в этом случае, если нет выборочного недоучета

некоторых языков, доля распределенных и нераспределенных величин после интерполяции должна быть примерно одной и той же.

160. При редактировании еще можно обращать внимание на религию. Опять же вопросы о религии можно интерполировать по возрасту и полу. Программа редактирования будет продолжать обновление, когда имеется

вся информация, и будет вытаскивать ответы из матрицы интерполяции для информации из категории «не известно». Такая матрица интерполяции будет выглядеть как матрица для языка, но в элементах массива вместо языка будет указана религия.

161. Заметим, что такое разъяснение предполагает нисходящий, последовательный принцип. Группы редактирования, использующие сложные методы, такие как метод Фелледжи–Холта или новая методология интерполяции (НМИ) (см. приложение V) применяют все смежные приемы редактирования одновременно. Настоящая процедура также предполагает наличие определенного порядка редактирования.

162. Многие экономические характеристики, такие как показатель активности, время, отработанное в последнюю неделю, или недели и время, отработанное за последний год, можно интерполировать, используя аналогичные характеристики. Используя аналогичные матрицы интерполяции, программа редактирования может быстро проверить значение для характеристик переменных и процесс редактирования должен в целом пойти быстрее.

163. Заметим, что иногда трудно получить должным образом отредактированные характеристики для первых матриц интерполяции в серии. Как правило, статистическое управление не хочет использовать неотредактированные элементы в качестве измерений для матрицы интерполяции; программа редактирования не будет использовать ни пол, ни возраст в качестве измерений матрицы интерполяции до того, как эти показатели не будут отредактированы. Следовательно, в первых нескольких матрицах интерполяции будут использоваться разные переменные, которые не нуждаются в редактировании или переменные, значение которых изменить нельзя. Для самой первой матрицы интерполяции для элементов, касающихся населения, редактор может использовать количество лиц в жилищной единице, включая «0» для свободных единиц.

164. Для редакции данных о жилье в целом, в первой матрице интерполяции можно также использовать количество человек в жилищных единицах в качестве исходного измерения, но группа редактирования может изменить приемы для элементов, касающихся жилья, с тем, чтобы учесть свободные единицы. Например, в отношении жилья, если сначала редактируются такие элементы, как «строительный материал наружных стен» или «тип стен», исходные величины могут быть основаны на количестве лиц в жилищной единице, включая значение для свободной единицы.

165. Если в случае свободной единицы информация «тип стен» достоверна, редактор обновляет первый элемент, используя информацию о типе наружных стен. Если тип стен известен, то для занятой жилищной единицы редактор обновляет элемент, сообразуясь с количеством человек в данной единице. Однако если строительный материал для внешних стен не известен, матрица интерполяции выдаст значение для строительного материала наружных стен, исходя из количества человек в данной единице.

166. После первоначального использования этой матрицы интерполяции группа редактирования может пожелать переключиться на некоторые другие жилищные характеристики, такие как «тип крыши» или «владение». Независимо от выбранной характеристики должно проводиться четкое различие между единицами и обеспечиваться достаточное разнообразие, чтобы один и тот же критерий не выбирался повторно. Повторный отбор одного и того же критерия может дать значение квазистатической интерполяции, а не то, которое получается при динамической интерполяции (*hot deck*). Например, применение динамической интерполяции в отношении армейских казарм («помещений, предназначенных для коллективного проживания») может привести к тому, что одна и та же величина будет использоваться повторно, если единственными отобранными характеристиками будут возраст и пол. В этом случае все проживающие будут, вероятно, мужского пола и большинство будет попадать под ограниченный возрастной диапазон. Следовательно, эта конкретная матрица может и не дать наилучшие результаты. Если признак «владение» будет иметь достаточно разновидностей с достаточной долей владельцев и арендаторов, то такая переменная может сработать. В противном случае страна может использовать различные критерии для типов крыши.

167. В целом многие группы редактирования считают, что, используя сопоставимые измерения для матриц интерполяции, они меньше занимаются проверкой, получают результаты более быстро и, возможно, более точные.

g) *Когда динамическая интерполяция не используется*

168. Если даже группа редактирования решает совсем не использовать динамическую интерполяцию, свое значение все же сохраняет последовательность редактирования. Например, возраст связан со многими элементами, включая степень родства с главой домохозяйства, уровень образования, занятость и рождаемость (для женщин). Рассмотрим членов домохозяйства на рисунке 17.

Рисунок 17. Пример главы домохозяйства с детьми без указания возраста ребенка

Лицо	Степень родства	Возраст	Уровень образования	Работа	Род занятий	Дети, родившиеся живыми
1	1	40	12	1	33	ПРОПУСК
3	3	X	7	ПРОПУСК	ПРОПУСК	ПРОПУСК

ПРИМЕЧАНИЕ: X = неправильные данные.
ПРОПУСК = не относится.

169. Запись для лица три — степень родства 3 (ребенок), но без указания возраста. Чтобы определить возраст, программа редактирования может использовать разницу в возрасте между главой домохозяйства и ребенком (или величину, полученную методом статической интерполяции или величину, взятую от предыдущей единицы путем интерполяции). Если, например, разница составляет 25, то возраст ребенка получается 15 лет (возраст главы домохозяйства 40 лет минус разница в возрасте 25).

170. Количество лет обучения также известно и составляет в данном случае семь лет. Возраст 15 лет вполне может соответствовать этому уровню образования. Поскольку количество соответствующих лет обучения для того или иного возраста меньше чем возрастной диапазон для разницы в возрасте между главой

домохозяйства и ребенком, то лучше проверить, прежде всего, соответствует ли данным показателям уровень образования. И если данный уровень указан, то разница в возрасте, определенная путем статической (*cold deck*) или динамической (*hot deck*) интерполяции, может быть использована для получения соответствующего возраста. Если уровень не известен, то тогда разницу в возрасте между главой домохозяйства и ребенком можно использовать для присвоения возраста.

171. Однако даже информация о возрастной разнице может отсутствовать. Для большинства стран более вероятно отсутствие данных об уровне образования, чем о возрасте. Но в следующем примере показано, что может предпринять группа редактирования, если отсутствуют данные о возрасте и уровне образования.

Рисунок 18. Пример главы домохозяйства с ребенком при отсутствии данных о возрасте и уровне образования ребенка

Лицо	Степень родства	Возраст	Уровень образования	Работа	Род занятий	Дети, родившиеся живыми
1	1	40	12	1	33	ПРОПУСК
3	3	X	X	ПРОПУСК	ПРОПУСК	ПРОПУСК

172. На рисунке 18 нет ни данных о возрасте, ни об уровне образования, но другая информация имеется. Лицо 3 недостаточно взрослое, чтобы быть занятым и слишком молодое, чтобы иметь детей (или оно мужского рода). Используя информацию о занятости из набора статических величин можно получить возраст, но это значение будет ниже, чем наименьший возраст, приемлемый для работы. И в другом случае, если группы редактирования используют динамическую интерполяцию, значение в матрице интерполяции дает величину для возраста. Для отобранного возраста следует, вероятно, использовать возраст главы домохозяйства как одну из переменных для сохранения согласованности. Например, если главе домохозяйства не 40 лет, а 20, то, очевидно, было бы неправильным обозначать возраст 14 лет для лица 3. Когда возраст установлен, то

можно определить и уровень образования, и он тем самым должен согласовываться и с возрастом и с родом занятий.

173. Если группа редактирования решит интерполировать все элементы или большинство из них, она должна решить вопрос том, как логически построить редактирование. Для элементов, касающихся населения, редактирование должно начинаться со всех элементов, где потенциально могут быть неизвестные величины. Группа редактирования должна использовать любую другую имеющуюся информацию для того, чтобы легче было решить по каждому элементу, включать ли его в самую первую матрицу интерполяции. Детали разработки матриц интерполяции очень зависят от того, к какой стране они относятся, однако все нацио-

нальные бюро переписи/статистические управления, скорее всего, имеют для этого какую-то информацию.

174. Многие пакеты программ редактирования отслеживают число лиц в жилищной единице по мере их прохождения. Например, матрица интерполяции для неизвестного пола может позволить присвоить мужской или женский пол в зависимости от количества проживающих в жилищной единице. Следовательно, исходное значение, которое может быть выбрано для лица, пол которого неизвестен или недостоверен, в случае жилища, где проживает один человек может быть «мужской». Для жилища, где проживают два человека, исходное значение может быть «женский». Для жилища, где проживают три человека, значение может быть «мужской» и т. д. Матрица в этом случае может быть использована только в качестве последнего средства, после того, как все данные, такие как пол главы домохозяйства, супруг(а) и информация о рождаемости проверены, произведена проверка на согласованность и все несоответствия устранены.

h) *Какого размера должны быть матрицы интерполяции?*

175. Большинство пакетов вычислительного оборудования не отвергают многомерные матрицы интерполяции. Прежде чем создавать матрицы интерполяции, следует учитывать следующее:

i) *Проблемы, возникающие в случае слишком больших матриц интерполяции*

176. Одна из самых серьезных проблем, с которой сталкиваются национальные бюро переписи/статистические управления, где ведут совместную работу специалисты, занимающиеся содержанием переписи, и специалисты по обработке данных, — это слишком активные программы редактирования. Здесь очень легко увлечься разработкой пакетов редактирования, и в этом случае программирование занимает намного больше времени, чем это необходимо, и замедляет обработку данных переписи или обследования. Группа редактирования может, например, решить, что при определении возраста кроме таких показателей, как «пол», «уровень образования» и «активность» для женщин необходимо включать такой показатель, как «количество детей, родившихся живыми». Может быть, с добавлением позиции «количество детей, родившихся живыми», оценка возраста может получиться чуть точнее, но это может быть не оправдано из-за повышения сложности программирования. В этом случае группы редактирования должны решить, сколько измерений матрицы интерполяции дадут наилучшие результаты с точки зрения точности и эффективности. Слишком большие матрицы

интерполяции (где слишком много элементов) нельзя обновить до конца, и в этом случае могут извлекаться статические величины, что было бы неправильно.

ii) *Осмысление принципа работы матрицы интерполяции*

177. Ввиду того, что матрицы интерполяции могут быть слишком большими, дезориентировать могут еще и пути доступа. Важно, чтобы и сотрудники, занимающиеся содержанием переписи, и обработчики данных могли отслеживать все маршруты. Совместно они должны удостовериться, что матрица интерполяции выполняет поставленную задачу. Кроме того, те, кто занимается содержанием переписи, и обработчики данных должны вместе проверить, чтобы каждая переменная или измерение матрицы интерполяции применялись должным образом. При этом они должны убедиться, что все комбинации работают надлежащим образом.

iii) *Проблемы, возникающие в случае слишком маленького размера матрицы интерполяции*

178. Матрица интерполяции бывает слишком маленькой, если у нее очень мало измерений, или если из-за группирования (когда слишком мало возрастных групп или уровня образования), одно и то же значение матрицы интерполяции используется неоднократно до обновления. Например, без измерения пола в массиве возраста всем детям в семье, скорее всего, будет присвоен один и тот же возраст, если он неизвестен. Сотрудники, занимающиеся содержанием переписи, должны вместе с обработчиками данных проверить матрицы интерполяции на все возможные комбинации, с тем чтобы ни одна из них не повторялась слишком часто.

iv) *Элементы, представляющие трудности для матрицы интерполяции*

179. Некоторые элементы, такие как «род занятий» и «отрасль», оказались, как выяснилось, трудными для редактирования. Раздельные матрицы интерполяции в связи с родом занятий и отраслью могут дать несогласующиеся результаты, а попытки провести перекрестную проверку всех парных данных о роде занятий и отрасли могут оказаться дорогостоящими и сложными для исполнения. Например, если окажется, что парикмахер работает на рыбном заводе, то здесь потребуются какой-то иной вид редактирования. Кроме того, большое количество категорий занятости и отрасли может существенно затруднить проведение динамической интерполяции. В отношении некоторых элементов группа редактирования может решить, что в этом случае редактирование приведет к обратным результатам и предпочтительнее использовать «не указано» или «не сообщается».

В других случаях принцип статической интерполяции (*cold deck*) может оказаться достаточным.

4. Проверка матриц интерполяции

180. Базовая структура матрицы интерполяции в программном пакете редактирования должна выглядеть примерно так, как показано на рисунке 19. В спецификациях редактирования должны быть указаны массивы, используемые для интерполяции, и статические величины для исходного набора значений.

а) Создание исходной статической матрицы

181. При выполнении нижеописанной процедуры матрица интерполяции обновляется каждый раз, когда обнаруживается лицо с достоверными данными во всех трех элементах — в данном случае «степень родства», «пол» и «возраст». Однако если программа редактирования находит недостоверное (или неуказанное) значение пола, то матрица интерполяции выбирает величину на основе достоверных кодов степени родства и пола (переменных, которые уже отредактированы).

Рисунок 19. Выборочный набор величин для статической матрицы и выборочный код интерполяции

.							
.							
.							
22	A01-AGE-FM-SEXRL (2,6)						
23.	Глава домохозяйства	Супруг/а	Ребенок	Прочие родственники	Родители	Не сообщается	
24.	40	40	10	20	65	20	.Мужской
25.	40	40	10	20	65	20	.Женский
.							
.							
.							
40	если AGE = 0:98						
41	допустим A01-AGE-FM-SEXRL (SEX,RELATIONSHIP) = AGE						
42	еще						
43	сообщение 'Age is unknown, so imputed' AGE						
44	написать 'Age is unknown, so imputed, Age = ' AGE						
45	интерполировать AGE = A01 (SEX,RELATIONSHIP)						
46	сообщение 'AGE is now known' AGE						
47	закончить, если						
.							
.							
.							

б) Сообщения об ошибках

182. В пакетах редактирования должно быть несколько методов, чтобы удостовериться, что редактирование и интерполяция идут должным образом. Два таких свойства команды «message» (сообщение) и «write» (писать) рассматриваются ниже.

183. Одним из источников информации является отображение сообщения, как показано на рисунке 19, выше. Эта команда выдает конкретные сообщения и общий подсчет (общее число раз, когда появляется сообщение) для географических категорий (например, счетный участок, мелкое административно-территориальное деление, крупное административно-территориальное деление), а также для каждого переписного листа. Для всех переписных листов сводный отчет может примерно выглядеть так, как показано на рисунке 20.

Рисунок 20. Пример сводного отчета для ряда интерполяций на одну ошибку

Подсчет	Количество ошибок	Сообщение	Номер строки
–	14-1	Слишком много детей на одну женщину	2629
–	14-2	Слишком много детей на одну женщину	2645
2	14-3	Мальчики не указаны	2669
2	14-4	Девочки не указаны	2678
33	14-5	Месяц последнего рождения не указан	2723
7	14-6	Нет детей, родившихся живыми; разница в возрасте между матерью и ребенком, подтверждение	2892

ПРИМЕЧАНИЕ: Здесь «14» просто относится к элементу 14 в данном ряду; ошибки пронумерованы по порядку.

184. В отчете по переписному листу (рисунок 21) может быть указан номер переписного листа, в том числе все обозначенные географические коды. Затем в отчете могут быть перечислены ошибки, найденные в программе по элементам (в данном случае «возраст»), и по номеру строки в программе пакета программного

обеспечения, указанном ниже справа. В данном примере графа «возраст» была не заполнена, но матрица интерполяции по степени родства и полу данного лица выдала возраст 48 лет. В данном случае конкретный возраст был неизвестен, но команда «message» могла по желанию указать и эту информацию.

Рисунок 21. Выборочные сообщения об ошибках в переписном листе

Идентификационный номер переписного листа: 01 01 017		Номер строки
AGE (1) =	Возраст неизвестен, поэтому дается по исчислению	46
AGE (1) = 48	Возраст уже известен	

с) Заказная распечатка ошибок

185. В программном обеспечении может быть еще одна команда, позволяющая проводить более подробный анализ спецификаций редактирования и потока редактируемых данных. Данную команду можно использовать для того, чтобы показать информацию до того, как произведены изменения, а затем и все изменения. И, наконец, она показывает запись и записи уже после

произведенных изменений. Таким образом, аналитик может удостовериться, что программа редактирования надлежащим образом проходит по всем путям доступа. Результаты могут быть такими, как показано на рисунке 22. В первой строке выходных данных даются переменные (например, территория, степень родства, пол, возраст). Затем показаны поступающие данные, потом ошибки (в данном случае отсутствие категорий «возраст»), а затем данные после изменения.

Рисунок 22. Пример дополнительной распечатки ошибок по переписному листу, включая множественные переменные

	Территория	Участок	Глава домохозяйства	Степень родства	Пол	Возраст
Поступающие данные	01	01	17	1	1	
Ошибка						Возраст неизвестен, и по исчислению Age = BLANK
Отредактированные данные	01	01	17	1	1	48

186. Такая процедура помогает группе редактирования определить, правильно ли проходит по путям доступа программа редактирования.

187. Такой вид проверки является важной частью редактирования данных переписи и обследования. Описанный ниже метод представляет собой один из возможных способов проверки процедур редактирования. Этот процесс можно начать с проведения специалистами систематического анализа путем создания «безупречного» домохозяйства. Безупречное домохозяйство — это полное домохозяйство — глава домохозяйства, супруг(а), дети, прочие родственники и лица, не являющиеся родственниками — со всеми своими характеристиками. Безупречное домохозяйство может пройти все стадии редактирования без всяких ошибок. Процедура совершается следующим образом:

- a) обработчики данных вводят одну ошибку в каждое домохозяйство последовательно в соответствии с последовательностью спецификаций редактирования и программы редактирования;
- b) затем аналитик проверяет все пути доступа на ранней стадии редактирования;
- c) после того как программа редактирования пройдет надлежащим образом все пути доступа, обработчики данных прогоняют выборку из всего набора данных в поиске отклонений в фактическом наборе данных и с внесением изменений, если это необходимо;
- d) в заключение обработчики данных прогоняют весь набор данных.

188. Убедившись, что сообщения работают нормально и соответствующие изменения сделаны, обработчик данных может по своему усмотрению отключить их на уровне переписного листа. Если в больших странах прогонять все наборы данных целиком, оставляя функции оператора сообщений для каждого переписного листа, то получающееся в результате количество строк и перфоленты будет запретительным. Тем не менее сводный отчет по этим сообщениям следует продолжать, поскольку он дает полезную информацию по различным уровням территориального распределения. При этом выходные данные будут выглядеть примерно так, как показано на рисунке 22.

189. Компьютерное редактирование, как правило, предусматривает процедуру сохранения. В журнале редактирования показаны все изменения данных и единицы счета для случаев изменения и заменяемых величин. Обращение к журналу редактирования помогает определить, является ли количество изменений достаточно низким для группы записей, которые надо учитывать.

190. Если в конкретном элементе слишком много

ошибок, то этот элемент нельзя было предварительно проверить должным образом или сам по себе, или в связи с другими элементами, что указывает на то, что счетчики или респонденты не поняли данный вопрос. Иногда счетчики путаются и, например, собирают информацию о рождаемости только у взрослых мужчин, а не у женщин. Если данные будут собираться таким образом систематически, группа редактирования может попросить программистов переместить данные о рождаемости из рубрики «мужской» в рубрику «женский», если речь идет о супружеской паре. В противном случае группа редактирования вряд ли сможет на этом этапе исправить ошибку.

191. Как правило, программу редактирования необходимо прогнать через несколько разных файлов, с тем чтобы охватить все ситуации. Кроме того, обработчикам данных нужно будет внести изменения в связи с синтаксическими или логическими ошибками. Даже наиболее опытные специалисты по обработке данных иногда вводят с клавиатуры «более» вместо «менее», и эта ошибка обнаруживается только после нескольких прогонов, поскольку конкретная проблема может сразу и не проявиться. Также могут сразу не проявиться и незначительные дефекты логики. Здесь опять специалисты, занимающиеся содержанием переписи, и специалисты по обработке данных должны вместе разрешать эти проблемы по возможности в самом начале процесса редактирования.

d) *Сколько раз прогонять программу редактирования?*

192. Как разъясняется ниже, в целом программу редактирования неплохо было бы прогонять три раза.

193. При первом редакторском прогоне в матрицы интерполяции вводятся не те величины, которые созданы в первоначальной статической матрице, а реальные значения. Многие страны используют данные из других источников — из предыдущей переписи или обследования, или из административных записей, — чтобы ввести в массив статические величины. Обработчик данных прогоняет весь набор данных или значительную его часть, чтобы получить значения для матрицы интерполяции. Вероятно, статические величины из фактического набора данных будут более точными и более соответствующими действительности. При редактировании используется лишь около 2 процентов этой первоначальной статической матрицы: остальные значения выводятся путем динамической интерполяции.

194. При втором прогоне осуществляется собственно редактирование. Второй прогон программы редактирования состоит из нескольких повторных прогонов для того, чтобы охватить все ситуации. В это время обработчикам данных необходимо внести изменения для

того, чтобы исправить погрешности, появившиеся в результате ошибок в синтаксисе или логике. Кроме того, даже наиболее опытные специалисты по обработке данных могут совершать ошибки и, поскольку конкретная проблема может сразу не проявиться, ошибка может быть обнаружена только после нескольких прогонов. Точно так же могут сразу не проявиться и незначительные дефекты в логике.

195. Третий прогон программы редактирования делается для того, чтобы убедиться, что 1) в наборе данных не осталось ошибок и что 2) программа редактирования не внесла новых ошибок. После того как обработчики сделают последний редакционный прогон, в выходных результатах ошибок уже не должно быть. Если они остаются, то, вероятно, нарушена логика редактирования, поэтому обработчик данных должен ее изменить. Кроме того, этот прогон обычно подсказывает обработчику данных, не внесены ли случайно при редактировании новые ошибки логикой программы редактирования.

5. Признаки интерполяции

196. Признаки интерполяции используются, в частности, для того, чтобы сохранять информацию о неотредактированных данных. Как отмечалось ранее, многие группы редактирования беспокоит потеря потенциальной информации при замене неотредактированных ответов. В тех случаях, когда значение меняется из-за несогласованности, группы редактирования могут по своему усмотрению сохранить первоначальное значение или значения для того, чтобы после переписи привести демографический анализ или анализ ошибок. Как специалисты по вопросам переписи, так и программисты захотят проанализировать различные аспекты недостающих, недостоверных или несогласующихся данных. Члены группы редактирования должны удостовериться, что интерполированные и неинтерполированные распределения непротиворечивы и посмотреть, есть ли в плане редактирования и интерполяции какие-либо систематические ошибки. Например, иногда специалисты по обработке данных случайно используют только статические величины, потому что программа не проводит обновления матрицы интерполяции. Если в стране проводится предварительное тестирование переписи,

то группе редактирования, возможно, потребуется изучить взаимосвязи между некоторыми переменными после предварительного тестирования, для того чтобы завершить составление переписного листа. Во время предыдущих переписей, когда еще не было микрокомпьютеров с большими жесткими дисками, многие статистические управления не имели места на своих лентах или других носителях информации для сохранения дополнительных данных. Сейчас же для большинства стран хранение информации о неотредактированных данных уже не представляет никакой проблемы.

197. Некоторые страны в качестве признака для каждого элемента предпочитают использовать простую, двоичную переменную учета. Этот метод прост и занимает один байт для каждой переменной. Например, во время переписи 1990 года Бюро переписи США поставило признаки интерполяции для каждой переменной в конце каждой записи и в сообщениях о жилье, и о населении. Например, для каждой жилищной переменной величина для признака была сначала «0», а затем менялась на «1», если каким-то образом менялся исходный элемент. Программа не сохраняла первоначальные значения, хотя отделения переписи иногда их компилировали или для каждой записи, или в совокупности.

198. Для сохранения неотредактированных ответов есть и другие методы. На примере, приведенном на рисунке 23, национальное бюро переписи/статистическое управление изменило возраст супруга с 70 на 40 лет, используя матрицу интерполяции. Национальное бюро переписи/статистическое управление может легко поставить значение, которое было до интерполяции, в данном случае 70, там, где зарезервировано место для признаков интерполяции, и зарезервировать переменную, используемую в публикуемых таблицах для распределенного значения, в данном случае 40. Чтобы изучить изменения в наборе данных, статистическое управление может составить ряды распределения численностей или комбинационные таблицы распределенных и нераспределенных величин. Если после такого анализа последствия редактирования набора данных, таблицы, составленные с учетом редакционных изменений, окажутся подозрительными или неверными, группа редактирования может изменить ход редактирования или его часть.

Рисунок 23. Выборочные демографические записи с признаками для исчисляемых значений

Лицо	Пол	Возраст	Дети, родившиеся живыми (СЕВ)	Признак пола	Признак возраста	Признак СЕВ
1	1	40	ПРОПУСК	0	0	1
2	2	40	7	0	70	0

199. На рисунке 24 показан случай, когда женщина 13 лет зарегистрирована как родившая ребенка (в рубрике «дети, родившиеся живыми» стоит «1»). Однако группа редактирования решила, что минимальный возраст для первого рождения будет 14 лет и что рождение детей лицами женского пола младше 14 лет, вероятнее всего, ошибка, а не факт. Как всегда, при этом возникает вопрос, не представляет ли этот случай искажение в наборе данных по сравнению с реальным значением.

Рисунок 24. Пример признака для молодой женщины с пробелом в графе «рождаемость» и добавлением признака

Лицо	Пол	Возраст	Дети, родившиеся живыми (СЕВ)	Признак пола	Признак возраста	Признак СЕВ
С пробелом в графе «рождаемость»						
4	2	13	1	0	0	
С пробелом в графе «рождаемость» и добавлением признака						
4	2	13	ПРОБЕЛ	0	0	1

200. В соответствии с правилами редактирования информация о детях, родившихся живыми, при интерполяции «пропускается». Отметим, что признак СЕВ немного сложнее, поскольку он должен учитывать интерполированную позицию BLANK (ПРОПУСК), а также цифровые позиции. Предположим, специалисты, занимающиеся содержанием переписи, хотят изучить номера и характеристики лиц в возрасте 13 лет, о которых сообщается, что у них есть ребенок. Обработчики данных могут зарегистрировать исходную информацию на участке записи, оставленном для признаков, как правило, в конце записи. После этого в публикуемых таблицах в отношении этой женщины не будет информации о детях, родившихся живыми, но такая информация тем не менее будет доступна для последующих исследований. На более позднем этапе, в частности, при планировании последующего обследования или следующей переписи, группы редактирования могут использовать информацию о детях, родившихся у 13-летних женщин, чтобы решить, можно ли им снижать регистрационный возраст.

201. Одна из проблем, связанных с использованием признаков интерполяции, заключается в том, что вышеописанная процедура занимает достаточно много места в компьютере. Если признаки повторяют каждую переменную, то набор отредактированных данных будет приблизительно в два раза больше набора неотредактированной информации. Для многих стран это было бы неприемлемо с точки зрения длительного хранения. Однако исходные данные и отредактированная информация могут храниться для последующего воспроизведения.

202. Страны с очень большой численностью населения, возможно, предпочтут использовать признаки интерполяции на выборочной основе для целей научных

исследований. Например, какая-либо страна захочет создать набор данных с каждой сотой жилищной единицей. Затем пойдет редактирование этого уменьшенного набора с признаками интерполяции, с тем чтобы оценить, как она влияет на качество данных, и определить различия между неотредактированными и отредактированными данными.

Г. ДРУГИЕ СИСТЕМЫ РЕДАКТИРОВАНИЯ

203. Данное *Руководство* посвящено в основном нисходящему принципу машинного редактирования данных переписи и обследования. Некоторые страны используют другую, более сложную процедуру компьютерного редактирования, известную как редактирование переменных по групповому принципу (см. раздел С.2, выше). Впервые этот метод был разработан Феледжи и Холтом (Fellegi and Holt, 1976), и он применяется обычно к наиболее важным переменным в переписи или обследовании: возрасту, полу, степени родства и семейному положению. Однако его можно применять и к любой группе переменных или ко всем переменным, указанным в вопроснике переписи или обследования. При такой методике программа редактирования смотрит ответы на эти вопросы одновременно для одного лица или для всех членов домохозяйства для того, чтобы выявить недостающие или несогласующиеся ответы.

204. Принцип Феледжи–Холта получил дальнейшее развитие в статистическом управлении Канады, которое использовало его при проведении переписей у себя в стране с 1976 года по 1991 год. Для переписи 1996 года в Канаде этот метод был еще более усовершенствован и получил название «новая методология интерполяции» (НМИ). Он впервые позволил одновременно провести с минимальными изменениями интерполяцию

цифровых и качественных переменных, выполняя большие [редактирование и интерполяция] задачи (Bankier, Houle and Luc).

205. Если в процессе редактирования используется традиционный метод динамической интерполяции (*hot deck*), то интерполированная информация для ряда вопросов в переписном листе может касаться многих отдельных индивидов в зависимости от того, какая информация использовалась для обновления матрицы интерполяции. Например, если данные о поле, степени родства и семейном положении лица А правильные, то соответствующие матрицы интерполяции будут обновляться этими величинами. Если возраст лица А не указан или недостоверен, он, естественно, не будет использоваться для обновления матриц интерполяции. Фактически обновление этой величины произойдет за счет других элементов. Таким образом, если у следующего лица пол будет противоречивым и этот элемент придется интерполировать, то данные о поле будут братья у лица А. Если еще неизвестен и возраст, то программа редактирования будет использовать возраст какого-либо другого лица.

206. В новой методологии интерполяции для элементов используются доноры в надежде, что вся недостающая или несогласующаяся информация может быть взята у одного или у нескольких доноров. Для того чтобы получить всю или большую часть информации от одного донора, все записи данных должны храниться в памяти компьютера. И тогда, если неизвестны или недостоверны и возраст, и пол, одна и та же хранящаяся переменная даст значения для обоих элементов.

207. Цели, методологии автоматизированной динамической интерполяции должны быть следующими:

- a) подлежащие интерполяции домохозяйства должны иметь тесное сходство с домохозяйством, данные о котором не удалось отредактировать;
- b) подлежащие интерполяции данные о домохозяйстве должны поступать по возможности не от двух или более доноров, а от одного. Кроме того, подлежащие интерполяции домохозяйства должны иметь тесное сходство с этим одним донором;
- c) одинаково хорошие приемы интерполяции, основанные на данных об имеющихся донорах, должны иметь одинаковые шансы быть отобранными, для того чтобы избежать ложного увеличения численности небольших, но важных групп населения (Bankier, Houle and Luc).

208. По новой методике интерполяции эти цели достигаются путем выявления в первую очередь прошедших редактирование домохозяйств, максимально близких по своим характеристикам домохозяйству, данные о котором отредактировать не удалось. Это означает, что два домохозяйства должны быть равными по максимально возможному числу качественных переменных, имея лишь небольшие различия между цифровыми величинами. Домохозяйства с такими характеристиками называются «ближайшие соседи». Следующий шаг заключается в том, чтобы определить, для каждого ближайшего соседа, наименьшие подмножества несовпадающих переменных (и цифровых, и качественных), которые позволили бы в случае интерполяции отредактировать данные о домохозяйстве. Одна из этих операций по интерполяции, которая проходит через редактора и напоминает не прошедшее редакцию домохозяйство и прошедшие редакцию домохозяйства, затем произвольно отбирается (Bankier, Houle and Luc).

III. РЕДАКТИРОВАНИЕ СТРУКТУРЫ

209. При редактировании структуры проверяется охват и определяется, как согласуются различные записи. Такое структурное редактирование должно обеспечить, чтобы *a)* все записи о домохозяйствах и жилье для коллективного проживания на счетном участке были указаны, причем в надлежащем порядке; *b)* по всем занятым жилищным единицам были записи о проживающих там лицах, а по свободным единицам лишних записей быть не должно; *c)* домохозяйства не должны иметь двойных или недостающих записей о лицах; *d)* счетные участки не должны иметь ни двойных, ни недостающих записей о жилье. Таким образом, при редактировании структуры проверяется, чтобы переписные листы были в целом заполнены.

210. Конкретные приемы редактирования структуры, используемые для одной переписи или обследования, возможно, придется со временем менять, поскольку технология, применяемая для определения и исправления ошибок в структуре, меняется весьма быстрыми темпами. Исходя из этого, в данной главе рассматривается более общий вопрос достоверности и взаимосвязи данных в записях и между ними. В главах IV и V говорится о конкретных отдельных элементах переписи населения и жилого фонда.

A. РЕДАКТИРОВАНИЕ ГЕОГРАФИЧЕСКИХ ПАРАМЕТРОВ

1. Местоположение жилых помещений (населенный пункт)

211. Согласно *Принципам и рекомендациям в отношении переписи населения и жилого фонда, первое пересмотренное издание* (Организация Объединенных Наций, 1998 год, стр. 70) населенный пункт определяется как «четко выделяемая концентрация населения..., где жители проживают в граничащих друг с другом в совокупности жилых помещениях и у которой есть свое название или признанный статус». Дополнительную информацию, касающуюся месторасположения жилых помещений, можно найти в подразделах «населенный пункт» и «городское и сельское население» в пунктах 2.49–2.59 *Принципов и рекомендаций*. «Лицам, занимающимся проведением переписи жилого фонда, необходимо изучить эту информацию, поскольку географические понятия, используемые для описания ме-

Рамка 3. Руководящие принципы редактирования структуры

Редактирование структуры должно обеспечивать выполнение следующих задач:

- 1) обеспечивать, чтобы каждый пакет счетных участков (СУ) имел правильные географические коды (территория, район, СУ и т. д.), чтобы для наименования пакетов использовалась единая практика;
- 2) обеспечить, чтобы была включена каждая жилищная единица; и чтобы были внесены все домохозяйства на СУ;
- 3) объединять домохозяйства по соответствующим СУ, а СУ — сливать в соответствующий более высокий территориальный уровень;
- 4) способствовать правильному выбору между страницами, посвященными отдельным лицам, и страницами, посвященными домохозяйствам, в буклетах, переписных листах или вне их, исходя из численности населения и компоновки переписного листа;
- 5) причислять каждую индивидуальную запись к своему достоверному типу записи;
- 6) обрабатывать записи о жилье для коллективного проживания отдельно от жилищных единиц;
- 7) обеспечить соответствие между различными видами записей: например, по свободным единицам нет записей о лицах, по занятым единицам есть, по крайней мере, запись об одном лице. Обеспечить, чтобы количество записей о лицах для каждого домохозяйства соответствовало общему учету домохозяйств в записях о жилье. Обеспечить правильное количество переписных листов, когда для одного домохозяйства используется много документов и их надлежащую увязку;
- 8) устранить двойные записи внутри домохозяйств (двойные лица) и между домохозяйствами (двойные домохозяйства или части домохозяйств), чтобы избежать чрезмерного охвата;
- 9) обрабатывать записи с пробелами в рамках одного вида записи;
- 10) обрабатывать недостающие жилищные единицы.

стонахождения жилых помещений при проведении переписей жилого фонда, чрезвычайно важны как для проведения самой переписи, так и для последующего табулирования полученных результатов» (Организация Объединенных Наций, 1998 год, пункт 2.312).

212. При редактировании элементов, касающихся местонахождения, географические коды должны быть абсолютно точными. Получение полных, точных кодов для географической иерархии при обработке данных является одной из труднейших задач всей переписи. Если территориальная принадлежность будет закодирована неправильно, то операторы, вводящие данные, могут приписать жилищную единицу или единицы какой-то другой части страны. Зачастую исправить такого рода ошибку очень сложно.

2. Городское и сельское население

213. Традиционные различия между городскими и сельскими районами в рамках одной страны основываются на том утверждении, что городские районы, независимо от того, как они определяются, предполагают другой образ жизни и обычно более высокий уровень жизни по сравнению с сельскими районами. Во многих промышленно развитых странах это различие стирается, и основное различие в условиях жизни между городскими и сельскими районами все больше выражается в степени концентрации населения. Несмотря на то что разница в образе и уровне жизни между городскими и сельскими районами в развивающихся странах остается значительной, быстрая урбанизация в этих странах вызвала большую потребность в информации, связанной с различиями в размерах городских районов (Организация Объединенных Наций, 1998 год, пункт 2.53).

214. Большинство стран еще до проведения переписи определяют, какие географические районы считать «городскими», а какие — «сельскими», а после сбора данных вводят необходимые коррективы. Если страна присваивает коды городским и сельским районам (например, 1 для городских и 2 для сельских), то эти коды можно ввести при наборе данных с клавиатуры или можно определить во время редактирования, исходя из критериев, установленных группой редактирования. Если группа редактирования подготовит список территориальных единиц с разбивкой на городские и сельские, то обработчики данных могут легко ввести соответствующие коды в записи о жилом фонде.

215. Следует сделать так, чтобы характеристики населения в целом соответствовали счетному участку. Например, в некоторых странах некоторые категории квалифицированных специалистов, за исключением врачей, учителей и лиц с аналогичным родом занятий, следует искать в сельских районах, а некоторые категории

работников сельского хозяйства — в городских районах. Группа редактирования должна произвести проверку, с тем чтобы убедиться, что географический район классифицирован правильно.

В. ПРОВЕРКА ОХВАТА

1. Регистрация де-факто и де-юре

216. Национальные бюро переписи/статистические управления стремятся собирать данные переписи де-факто (где находятся люди в ночь переписи) или де-юре (где они обычно находятся). При проверке взаимосвязи между записями о жилье, в частности, в связи с учетом лиц в жилых помещениях, с записями об отдельных лицах необходимо учитывать вид переписи. Иногда страны собирают информацию и де-факто, и де-юре. Информация по каждому лицу может означать, проживает ли он/она 1) постоянно, 2) находится там временно, а обычно проживает в другом месте или 3) обычно проживает в данном домохозяйстве, но временно отсутствует. В таблицах с данными де-факто используется только (1) и (2), если присутствуют три вида; в таблицах с данными де-юре используется только (1) и (3), если присутствуют все три вида⁴. Программа редактирования должна быть составлена таким образом, чтобы при наличии всех трех видов записи данные должным образом согласовывались между собой. Если в записях де-факто есть несколько ответов, это может указывать на то, что на самом деле речь идет об отсутствующих на данный момент лицах или возникает другая связанная с регистрацией проблема, которая может потребовать специального решения.

2. Иерархия домохозяйств и жилищных единиц

217. В главе V рассматривается вопрос о взаимосвязи между домохозяйствами, жилищными единицами и жилыми помещениями. Применение этих понятий зависит от отдельной национальной организации по проведению переписи или статистического управления. Однако прежде чем перейти к редактированию отдельных жи-

⁴ Национальные организации по проведению переписи/статистические управления, пользующиеся этими категориями, должны использовать их весьма осторожно не только во время сбора и обработки данных, но и во время последующего анализа. Когда используются эти три категории, пользователи должны знать об отобранной совокупности, поскольку если анализировать весь набор данных, то некоторые лица будут включены дважды. Если нужна совокупность де-факто, из таблицы необходимо исключить категорию (3) — лица, временно отсутствующие; если необходима совокупность де-юре, из таблицы необходимо исключить категорию (2). Во время первоначального табулирования, составления таблиц для распечатанных отчетов и дополнительных носителей группа редактирования может выбрать для обработки подмножество из всего набора данных. Для последующих таблиц в документации файлов следует четко указывать, каким образом обрабатывать все возможные таблицы.

лицных единиц, группа редактирования должна разработать методы проверки соблюдения иерархии.

3. Фрагменты переписных листов

218. Прежде чем редактировать элемент за элементом, компьютерная программа должна проверить достоверные записи, недостающие записи и двойные номера строк в рамках редактирования структуры. Она должна также определить, относятся ли редактируемые записи к лицам, проживающим в жилых помещениях группами. Операторы могут сделать ошибку при вводе данных, а иногда они могут забыть исключить обрывочную информацию (части записей). Одна из функций предварительного редактирования должна состоять в том, чтобы посмотреть файл обрывочных записей, для того чтобы исключить их. Наиболее распространенной будет запись, содержащая географические коды, но без информации о населении или жилом фонде.

С. СТРУКТУРА ЗАПИСЕЙ О ЖИЛОМ ФОНДЕ

219. Один из вопросов, который можно включить в вопросник для сбора информации при проведении национальных переписей населения или жилого фонда, — это число жилищ в строении. В данном случае единицей учета является строение, и собираются сведения о числе обычных и основных жилищ в нем (Организация Объединенных Наций, 1998 год, пункт 2.418).

220. Термин «общее редактирование» относится к тем случаям, когда необходимо обеспечить, чтобы количество жилищных единиц как части строения соответствовало общему количеству жилищных единиц в записи о жилищах. В случае несоответствия количество жилищных единиц, указанных в качестве одной из характеристик строения, следует откорректировать, с тем чтобы оно соответствовало количеству записей о жилищных единицах. Если данное строение закодировано как имеющее пять жилищных единиц, а фактически в отдельных записях по этому строению указаны четыре жилищных единицы, группа редактирования должна решить, какую поправку внести: *a)* изменить первую цифру на основе учета отдельных записей (которые в большинстве случаев более приемлемы); или *b)* ввести еще одну запись, используя информацию об имеющихся записях (чего следует избегать).

Д. СООТВЕТСТВИЕ МЕЖДУ ЗАПИСЯМИ О НАСЕЛЕНИИ И ЗАПИСЯМИ О ЖИЛОМ ФОНДЕ

221. Если перепись или обследование включают в себя как записи о населении, так и записи о жилом фонде, то

при редактировании структуры необходимо удостовериться, что эти два вида записи согласуются между собой.

1. Свободные и занятые жилища

222. По свободной жилищной единице записей о населении быть не должно, а по занятой жилищной единице такие записи должны быть. Если есть демографические записи, а жилище зарегистрировано как свободное, то из категории свободного оно будет переведено в категорию занятого. Иногда запись составлена так, что статус свободного и владение входят в один и тот же элемент, поэтому такую информацию необходимо учитывать и при принятии решения. Кроме того, если есть ответ в графе «стоимость единицы» для единиц, занимаемых владельцем, или «арендная плата» для единиц, занимаемых арендатором, то программа редактирования использует в своем определении эту информацию; в противном случае может потребоваться матрица интерполяции.

223. Если там, где жилищная единица считается занятой, демографической записи нет, то группа редактирования должна решить, считать ли ее за свободную единицу или подставить данные о лицах из другой единицы. Если единица свободна, то методом интерполяции легко заменить эту переменную на свободную. Однако если жилищная единица занята, группа редактирования должна решить, стоит ли и как приписывать лиц из другой единицы, по возможности, с тем же количеством проживающих и аналогичными характеристиками. Если нельзя узнать характеристики неуказанных лиц, этот метод следует использовать, если вообще использовать, только когда группа редактирования решит, что у нее нет другого выхода. Три возможных варианта изложены ниже.

a) Оставить жилищную единицу в категории «свободная»

224. В этом случае группа редактирования принимает решение о том, что свободные жилищные единицы, данные о которых поступают с мест, должны оставаться как свободные, и поэтому они не интерполируются. О редактировании данных о свободных жилищных единицах говорится в главе V.

b) Повторное посещение жилищной единицы несколько раз для заполнения переписных листов

225. Национальное бюро переписи/статистическое управление может решить применить процедуры, требующие от счетчиков возвращения к данным о свободных жилищных единицах, пока они не убедятся в том, что они или свободны или заняты, и пока они не себе

рут, по крайней мере, минимальные характеристики. В этом случае группа редактирования должна разработать такие приемы, чтобы при редактировании проверить, свободна ли данная жилищная единица и достаточно ли характеристик, чтобы считать ее занятой. В зависимости от того, что группа редактирования считает «минимальной» информацией, применяется обычный метод редактирования, описанный в главе IV, или берутся данные из записей донора для «недостающих» лиц, о чем говорилось в пунктах 205–207, выше.

с) *Замена данных о недостающих лицах данными из другой жилищной единицы*

226. Процедуры замены данных о целых домохозяйствах или отдельных неуказанных лицах описываются в других подразделах настоящей главы. В соответствии с этими процедурами необходимо исходить из предположения о том, что неуказанные лица имеют те же характеристики, что и те, которые берутся для замены, что почти наверняка не всегда так, а сами процедуры очень сложны. Тем не менее без этих процедур подсчет количества человек и лиц по характеристикам может дать более низкие показатели.

2. *Двойной учет домохозяйств и жилищных единиц*

227. Двойной учет жилищных единиц происходит по целому ряду причин. Иногда оператор, вводящий данные, может дважды ввести одну и ту же жилищную единицу. Иногда разные операторы, вводящие данные, могут случайно набрать на клавиатуре одни и те же жилищные единицы или даже целые счетные участки из-за отсутствия гарантий качества в национальном бюро переписи/статистическом управлении. И, наконец, счетчик может неправильно записать территориальный код жилищной единицы, что ведет к двойному учету информации, когда две жилищные единицы имеют один и тот же территориальный идентификационный номер.

228. Если бюро отслеживает введенные пакеты данных, то дублирование не будет. Тем не менее следует разработать такую программу редактирования, которая исключит двойной учет домохозяйств, когда операторы по вводу данных дважды набирают одну и ту же информацию о домохозяйстве или домохозяйствах. Страны не должны сортировать свои данные до окончания структурной проверки и устранения проблем с двойными записями. До сортировки сотрудники могут исправить пакеты данных вручную; после сортировки сотрудники могут уже не обнаружить эту проблему. После сортировки данных программа редактирования может проверить двойные домохозяйства и путем интерполяции устранить последующие двойные данные.

3. *Недостающие домохозяйства и жилищные единицы*

229. Аналогичным образом после сортировки могут проявиться и недостающие домашние хозяйства. Например, программа редактирования предполагает получить последовательность домохозяйств в рамках самого низкого территориального уровня, например, 1, 2, 3, 4, а получает только 1, 2, 4. Тогда необходимо принять решение изменить нумерацию этих единиц или найти какой-либо «приемлемый» метод замены единицы 3 другой единицей.

4. *Соответствие между числом жильцов и суммарным количеством жильцов*

230. Число жильцов, зарегистрированных в записях о данной жилищной единице, должно точно совпадать с общим количеством членов домохозяйства. Программа редактирования суммирует количество человек, а затем сравнивает эту величину с числом жильцов в записи о жилищной единице. Если сумма отличается по значению от числа жильцов, необходимо или скорректировать количество жильцов так, чтобы оно равнялось количеству членов домохозяйства, или скорректировать отдельные записи. В главе V конкретно говорится о редактировании жилищных данных по числу жильцов.

а) *Когда число жильцов больше суммарного количества жильцов*

231. Если данные о числе жильцов в записи о жилищных единицах больше по значению суммы отдельных лиц, перед группой редактирования встает реальная проблема. Никто не может знать характеристики недостающих лиц. В связи с этим группа редактирования, решившая интерполировать данные о недостающих лицах по характеристикам или заменить их данными из аналогичных домохозяйств, может столкнуться с дилеммой. Данные о недостающих лицах заменять не следует. Однако если число жильцов приемлемо, можно сократить численность зарегистрированного населения. Группа редактирования должна проанализировать всю картину целиком и затем найти приемлемый путь.

232. Есть несколько способов обнаружения и замены недостающих записей, и ни один из них не является полностью удовлетворительным. Можно сохранять в памяти целые домохозяйства с разными, важными характеристиками. Если появляется домохозяйство с несколькими, но не со всеми его членами, то можно поискать файл с данными о домохозяйстве, где все или почти все известные характеристики совпадают, а затем скорректировать данные о недостающих лицах на основе данных о других лицах из домохозяйства-донора. Однако программировать такую операцию весьма

сложно, поэтому национальные бюро переписи/статистические управления, применяя такой подход, должны приступать к планированию задолго до проведения такой операции.

233. Разновидность этой процедуры состоит в том, чтобы пометить все домохозяйства с недостающими записями и продолжать редактирование остальных данных. В конце редактирования после того, как исправлены все отдельные записи, группа редактирования может предложить специалистам по обработке данных просмотреть файл и внести добавления и изменения, используя полностью отредактированный набор данных. Используя такой нисходящий принцип, группа редактирования может найти приемлемых доноров.

б) Проверка числа лиц по полу

234. Иногда количество жильцов в записи о жилищных единицах регистрируется по полу. В этом случае программа редактирования должна суммировать количество лиц по каждому полу отдельно. И если суммы отличаются от количества жильцов, одну из величин в каждом случае необходимо корректировать. Как правило, вместо добавления «недостающих» записей или исключения записей с полезной информацией записи о жилищных единицах корректируются.

с) Порядковая нумерация

235. Демографические записи должны быть расположены и пронумерованы по порядку. Такая нумерация должна указываться как переменная, например, номер строки или порядковый номер в переписном листе. Порядковые номера должны указываться в цифровой последовательности. Ошибки могут возникать из-за того, что иногда переписные листы или личные бланки оказываются испорченными, потому что счетчики собирают информацию в неправильном порядке или могут пропускать страницы, непреднамеренно оставляя пустые страницы в наборе данных. Хотя отсутствие последовательности, как правило, не влияет на процесс редактирования или составления таблиц, многие национальные бюро переписи/статистические управления предпочитают переставлять личные данные в правильном порядке. В связи с этим программа редактирования должна обладать способностью обнаруживать приводимые в неправильном порядке личные данные и вновь устанавливать их последовательность. Поскольку иногда повторное упорядочение может повлиять на степень родства с главой домохозяйства, это необходимо предусмотреть в спецификациях редактирования. Повторное упорядочение определенно повлияет на такие переменные, как номер строки матери и номер строки мужа.

5. Соответствие между жильцами и типом строения/домохозяйства

236. Вид отношений между членами домохозяйства должен согласовываться с типом жилищной единицы. Иногда данные о членах домохозяйства появляются в данных о жилище, зарегистрированном как жилые помещения для коллективного проживания или наоборот. В этих случаях вид отношений или тип жилищной единицы должен указываться с учетом численности домохозяйства и других переменных.

Е. Двойные записи

237. Двойные номера строк вряд ли появятся при оптическом считывании или ином сканировании переписных листов. В отношении бланков, вводимых с клавиатуры, национальное бюро переписи/статистическое управление может по своему усмотрению проверить вручную соответствие между списком домашних хозяйств и номерами строки для домашнего хозяйства, которые вводятся с клавиатуры. Такая ручная проверка может повысить качество вводимых данных, особенно при сопоставлении 1) имен на странице, где перечислены все члены домохозяйства с 2) данными по колонкам, рядам или страницам. Два человека, данные которых первоначально кажутся продублированными, могут на деле оказаться близнецами, если посмотреть их имена.

238. В вводимых бланках не должно быть дублирования номеров строки, если экраны данных или схемы пропусков установлены должным образом. Большинство современных пакетов программ создают порядковые номера автоматически при вводе данных. Ошибка может быть внесена, когда сотрудники вводят двойные данные для одного лица или когда ошибочный номер строки может создать двойную запись. По мере обработки каждой записи программа редактирования сопоставляет ее с предыдущими демографическими записями для конкретной жилищной единицы. При редактировании необходимо убедиться, что каждый номер строки установлен правильно. Двойные номера строки являются ошибкой и подлежат изменению.

239. Страны могут по своему усмотрению разработать собственные схемы ввода данных, не используя готовые пакеты. В этом случае группа редактирования должна выбрать приемлемый уровень ошибок. Для таких решений есть множество методов. Один из них заключается в следующем:

а) если две разные записи имеют одинаковый номер строки, а номер различающихся характеристик два или меньше, редактор устранил одну из записей, исходя из того, что она, вероятно, двойная;

- b) если различаются три или более характеристик, номер строки будет изменен.

Ф. СПЕЦИАЛЬНЫЕ СОВОКУПНОСТИ

1. Коллективное проживание

240. При редактировании структуры должны обрабатываться данные о лицах, проживающих в коллективах, таких, как учреждения, казармы, дома для престарелых или интернаты, в отличие от тех, кто живет в обычных жилищных единицах. Поскольку в коллективах, как правило, нет главы домохозяйства, страны должны определить, как лучше проводить различие между типами жилищных единиц. Один из методов состоит в том, чтобы использовать другой тип записи для коллективов. Еще можно присвоить конкретный код для отношений, которые будут означать помещение для «группового» или «коллективного» проживания.

- a) *Когда для коллективов используется другой тип записи*

241. Если национальное бюро переписи/статистическое управление принимает решение использовать отдельный тип записи, группа редактирования без труда сможет определить, какие записи относятся к коллективным. Таблицы с данными о коллективах легко составить, используя непосредственно только эти записи. Переменные, которые характерны исключительно для коллективных записей, например тип коллектива, можно редактировать и интерполировать раздельно. Переменные, исключенные из коллективных записей, можно легко проверить, чтобы удостовериться, что они действительно отсутствуют. Однако файл получается более объемным, поскольку эти записи должны быть короче, чем обычные демографические записи, но они занимают место как в прямоугольном файле. Кроме того, во время редактирования и интерполяции некоторые программы могут проверять определенные элементы в демографических и коллективных записях.

- b) *Когда переменная различает коллективы от других записей*

242. При использовании отдельной переменной вместо отдельного типа записи группе редактирования, возможно, будет труднее определить, какие записи относятся к коллективным или являются коллективными записями. В этом случае можно легко составить таблицы с данными о коллективах, опираясь лишь на саму переменную, которая отмечает, какие записи относятся

к лицам, проживающим коллективно. При этом переменные, относящиеся исключительно к коллективам, например, тип коллектива, можно редактировать и интерполировать отдельно. Переменные, которые исключаются из коллективных записей, можно легко проверить, чтобы удостовериться, что они фактически пусты, используя код для коллективов. Файл при этом получается более компактным, поскольку дополнительных записей о лицах, проживающих коллективно, не требуется, а они просто включаются в качестве демографических записей с другим кодом для переменной домохозяйства/коллектива. При редактировании и интерполяции программа проверит только демографические записи по некоторым элементам, а не будет проверять записи, касающиеся населения и коллективов.

- c) *Когда отсутствует код «тип коллектива»*

243. Код, обозначающий коллективы, может отсутствовать или быть недостоверным или может быть несоответствие между кодом коллектива и кодами отношений. Если код для коллективов отсутствует, а коды отношений указывают на коллектив, то в этом случае предлагается соответствующим образом изменить код коллектива. При наличии кода коллектива в отсутствие кода отношений, последний можно определить по типу коллектива.

- d) *Когда есть код коллектива, но все лица находятся в родственных отношениях*

244. Если есть код для коллектива, но все лица в данной жилищной единице находятся в родственных отношениях, если судить по кодам отношений, то тогда код следует изменить, чтобы он обозначал жилищную единицу. Вместе с тем, если эта единица зарегистрирована как домохозяйство, но там нет двух лиц, состоящих в родстве, то, может быть, целесообразно в этом случае заменить ее на помещения группового или коллективного проживания. В домохозяйстве может быть 5 или 6 не связанных родством лиц и все равно это не будет коллективом. Как подчеркивалось выше, для разрешения особых, необычных случаев членам группы редактирования, возможно, необходимо проконсультироваться между собой.

- e) *Различия между разными типами коллективов*

245. В большинстве стран есть разные типы коллективов. Часто информация классифицируется еще подробнее по конкретным типам помещений для коллективного проживания. Эту информацию можно или кодировать отдельно, как тип «коллективных помещений», или

включать в самых разных вариантах в коды отношений в домохозяйстве.

2. Жильцы без жилья

а) Сезонная миграция

246. В некоторых странах с сезонной миграцией счетчику понадобится знать, занята та или иная жилищная единица или свободна в связи со временем отметки. Так, если даже о домохозяйстве есть полная информация, оно может быть также учтено (зарегистрировано) в другом месте. Естественно, может быть и обратный вариант. Домохозяйство, у которого есть два жилища в разных местах, может быть полностью пропущено, если не проследить за этим.

247. Иногда целые домохозяйства регулярно живут одну часть года в одном месте, а остальную часть года — в другом месте. Национальные бюро переписи/статистические управления и группы редактирования должны решить, как поступать в разных ситуациях. Например, часть лиц проводит какое-то время ежегодно в другом месте, как те, кто живет в более холодной части страны в теплое время года, и те, кто живет в теплой части страны в холодное время года. Еще один случай связан с кочевниками, которые часть года передвигаются, а часть года ведут оседлый образ жизни, — возможно, ту часть года, когда в стране намечена перепись.

б) Бездомные

248. По определению, регистрация бездомных не дает информации о жилищах. Однако «фиктивная» запись, первоначально новая запись, которая включает в себя пустые величины для той или иной переменной, облегчает структурную проверку и делает эту запись совместимой со структурой других жилищных единиц. Группа редактирования должна в этом случае решить, стоит ли создавать такую фиктивную запись о жилищах, с тем чтобы содействовать обработке данных и составлению таблиц.

с) Беженцы

249. Беженцы тоже могут оказаться во временных помещениях, что может потребовать введения конкретной переменной, то есть отдельного типа записи или фиктивной записи о жилищах, чтобы учесть их состояние. Группа редактирования должна будет для этого разработать и применить соответствующие процедуры.

Г. ОПРЕДЕЛЕНИЕ ГЛАВЫ ДОМОХОЗЯЙСТВА И СУПРУГА(И)

1. Редактирование переменной, касающейся главы домохозяйства

250. При определении членов домохозяйства сначала по традиции определяется глава домохозяйства или базовое лицо, а затем остальные члены домохозяйства в соответствии со степенью родства с главой домохозяйства или базовым лицом. Глава домохозяйства определяется как лицо, которое признается таковым другими лицами, входящими в это домохозяйство. Страны могут использовать тот термин, который они считают наиболее приемлемым для определения этого лица, в частности, главы домохозяйства, базового лица домохозяйства и других до тех пор, пока определяемое таким образом лицо используется для определения отношений между членами домохозяйства. Каждой стране рекомендуется представлять в публикуемых отчетах используемые понятия и определения (Организация Объединенных Наций, 1998 год, пункт 2.67).

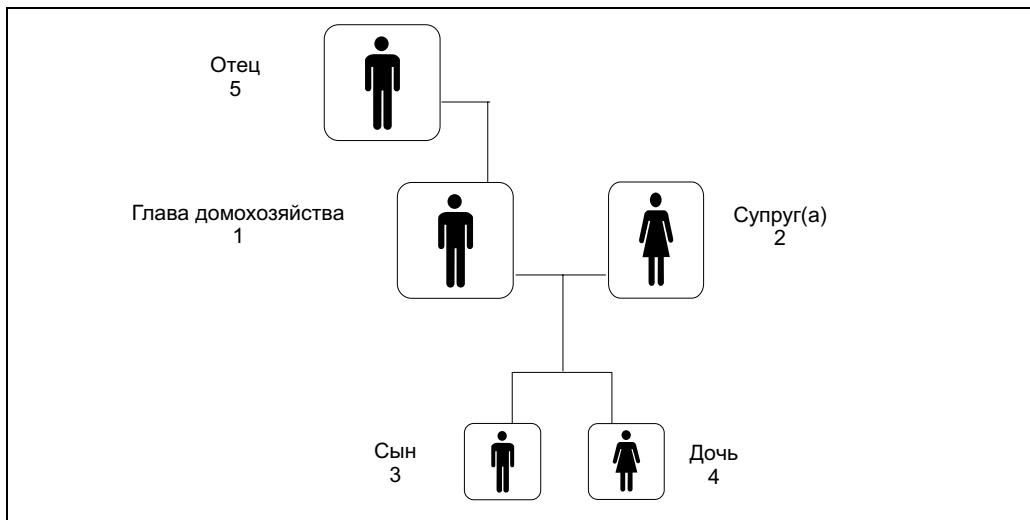
а) Порядок отношений

251. Порядок отношений в жилищной единице влияет на результаты редактирования, поскольку многие редактируемые данные построены на том, что глава домохозяйства — это первое лицо, и его/ее данные будут редактироваться в первую очередь. Например, при редактировании данных о главе домохозяйства в первую очередь проверяются такие переменные, как язык, этническая принадлежность и религия. Если о главе домохозяйства имеется достоверная информация, выраженная любой из этих переменных, то она интерполируется, чтобы получить данные о любом другом лице, домохозяйстве, которые отсутствуют, неправильно закодированы или неправильно введены с клавиатуры (см. главу IV). Данные о главе домохозяйства должны редактироваться в первую очередь, поскольку его/ее характеристики используются для указания или интерполяции данных о других членах домохозяйства.

б) Когда глава не является первым лицом

252. На процесс редактирования влияют действия счетчиков на местах, которые они предпринимают исходя из различных ситуаций, возникающих в связи с определением главы домохозяйства. Чтобы это лучше понять, рассмотрим сначала домохозяйство, показанное на рисунке 25.

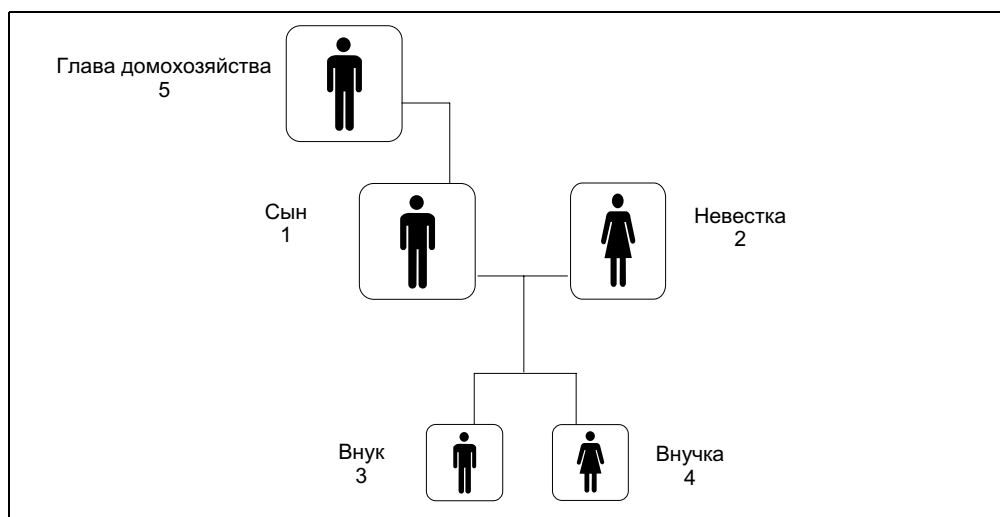
Рисунок 25. Пример домохозяйства, где глава домохозяйства зарегистрирован как первое лицо



253. На примере данного домохозяйства показана типичная ситуация, встречающаяся на местах: глава домохозяйства и супруг(а), их дети и отец главы домохозяйства. Если счетчик собирает информацию таким образом, то редактирование, основанное на том, что глава домохозяйства занимает первую позицию в домохозяйстве, пройдет гладко.

254. Однако если регистрация проводится таким образом, что дедушка определяется как глава домохозяйства, то конфигурация отношений меняется, как показано на рисунке 26. Такая ситуация может возникнуть, если счетчик, посетив дом, обнаруживает нуклеарную семью в составе мужа, жены и двух детей, и во время опроса отец главы домохозяйства входит в комнату и заявляет, что главой домохозяйства является он. Если взять за основу этого предполагаемого главу домохозяйства, то лицо 5 становится главой домохозяйства, лицо 1 — сыном, лицо 2 — невесткой и т. д.

Рисунок 26. Пример домохозяйства, где глава домохозяйства зарегистрирован как пятое лицо



255. Как становится ясно на примере этих двух домохозяйств, траектории редактирования, основанные на разных выделенных главах домохозяйства, будут разными. Чтобы определить фактического главу домохозяйства для последующего редактирования и составления таблиц, есть три возможности: *a)* можно использовать указатель, чтобы отметить, кто является главой, и этот указатель можно использовать для дальнейшего редактирования и составления таблиц; *b)* если глава не зарегистрирован как первое лицо, его/ее можно переместить на первую позицию, а тех, кто находится в списке выше, можно каждого переместить на одну позицию ниже; или *c)* можно изменить коды отношений и сделать главой первое лицо, невзирая на другие отношения.

- i) *Установление указателя для записи о главе домохозяйства*

256. При редактировании данных о главе домохозяйства используется указатель, чтобы определить номер строки главы домохозяйства в жилищной единице. Если глава остается на выбранной позиции, указатель можно поставить на эту позицию, и тогда всегда можно легко найти главу, когда это необходимо в конкретных целях редактирования или табулирования. Переменную «глава-указатель» можно установить на номер строки главы домохозяйства и использовать во время редактирования, чтобы вычислить или интерполировать недостающие или недостоверные характеристики других лиц в единице. Если глава является первым лицом в домохозяйстве, то значение переменной «глава-указатель» будет 1.

- ii) *Как сделать главой первое лицо*

257. Группа редактирования может решить переместить главу на первую позицию в домохозяйстве. В этом случае программирование несколько сложнее того, которое требуется для операции, описанной в подпункте (i), выше. Специалист по обработке данных должен разработать программу, которая переместит главу на первую позицию в списке, после него пойдет лицо, которое до этого было на позиции 1, затем лицо, которое было на позиции 2, и т. д. до тех пор, пока не дойдет до лица, указанного непосредственно перед тем, кто был главой. Таким образом, если глава находится на позиции 5, порядок следования изменится с 1, 2, 3, 4, 5 на 5, 1, 2, 3, 4. После такого изменения глава будет на позиции 1, и дальнейшее редактирование пойдет легче, поскольку глава всегда будет на этой позиции. Тем не менее при выполнении такой операции целостность набора данных несколько нарушится. Поскольку порядок лиц изменился, то аналитикам, возможно, будет трудно определить фактический порядок лиц, данные о которых

были собраны на местах, и возможное влияние такого порядка на толкование результатов.

- iii) *Изменение кодов родства, чтобы первое лицо сделать главой*

258. Если группа редактирования решит, что первое указанное лицо должно быть главой домохозяйства, то при редактировании необходимо выполнять процедуры *a* и *b*:

- a) первому лицу присваивается значение главы домохозяйства;
- b) затем для корректировки данных о домохозяйстве меняются значения для других членов домохозяйства. Например, на рисунке 26 родитель выступает в качестве главы домохозяйства. Когда лицо 1 становится главой домохозяйства, лицо 2 нужно указать как «супруга(у)», лиц 3 и 4 — как «ребенок», а лицо 5 — «родитель» (как показано на рисунке 25). При этом понадобится еще стандартная подпрограмма для хранения матрицы с исходными и измененными значениями.

259. При выполнении такой процедуры целостность набора данных нарушается еще больше. Порядок лиц не меняется как в предыдущем примере, и аналитикам будет трудно определить фактический порядок лиц, данные о которых собраны на местах. В данном случае все отношения изменятся, и аналитики не будут знать, кто был указан первоначально в качестве главы домохозяйства. Вместе с тем, если глава стоит на первой позиции, то таблицы номинально будет составлять легче. В отличие от предыдущего примера программистам при этой процедуре не надо физически перемещать записи.

- c) *Более одного главы*

260. Если обнаруживается более одного главы домохозяйства, то группа редактирования должна определить, кого назначить главой домохозяйства. Редактирование должно производиться на основе характеристик, установленных специалистами, занимающимися содержанием переписи, и ходом редактирования. Затем программа редактирования должна изменить степень родства другого лица (лиц), определенного как глава домохозяйства.

- d) *Если глава отсутствует*

261. Если же главы домохозяйства нет, то при редактировании необходимо определить, кого назначить главой домохозяйства. В этом случае при редактировании видимо придется откорректировать степень родства других членов домохозяйства.

2. Редактирование данных о супруге

а) Когда присутствует точно один супруг в моногамном обществе

262. Если есть точно один супруг, переменная «SPOUSE-POINTER» («супруг-указатель») отслеживает номер строки супруга для дальнейшего редактирования. Программа редактирования может включать в себя поиск противоположного пола главы домохозяйства и супруги(а), соответствующих возрастным различиям или другим смежным характеристикам.

б) Когда обнаруживается более одного супруга в моногамном обществе

263. В моногамном обществе, если в наборе данных обнаруживается более одного супруга, то программа редактирования должна определить, кто является супругом(ой), и изменить соответственно степень родства других лиц, которые были обозначены как супруги. Кроме того, специалисты, занимающиеся содержанием переписи, должны определить, какими должны быть характеристики и как пойдет редактирование.

в) Супруги в полигамном обществе

264. Если в полигамном обществе обнаруживается более одного супруга, группа редактирования, возможно, захочет оставить информацию как есть, или провести какую-то проверку на согласованность. Например, как минимум, у каждого из супругов в полигамном браке глава должен быть противоположного пола. И если супруги оказываются одного пола, то необходимо отредактировать ранее встречавшиеся данные о супругах одного и того же пола.

Н. ВОЗРАСТ И ДАТА РОЖДЕНИЯ

1. Когда есть дата рождения, но не указан возраст

265. Когда регистрируется дата рождения, а возраст нет, то эту информацию можно получить путем вычитания даты рождения из даты проведения переписи или обследования. Некоторые национальные бюро переписи/статистические управления предпочитают получать

данные о возрасте, основываясь только на годе проведения переписи и годе рождения, что дает значение с потенциальным отклонением. Если используются год и месяц, то возраст будет вычислен более точно, но самые точные результаты дадут день, месяц и год.

2. Когда возраст и дата рождения не согласуются между собой

266. Когда при переписи или обследовании собираются и данные о возрасте и дата рождения, «вычисленный» возраст получается путем вычитания даты рождения из базовой даты. Если это значение отличается более чем на один год от зарегистрированного возраста, то группа редактирования, возможно, пожелает произвести корректировку. Как правило, дата рождения предшествует зарегистрированному возрасту, и зарегистрированный возраст заменяется вычисленным возрастом.

I. УЧЕТ НЕДОСТОВЕРНЫХ ДАННЫХ

267. Некоторые группы редактирования, возможно, предпочтут выполнить процедуры подсчета количества недостоверных и несогласующихся данных для основных переменных (или всех переменных), таких как возраст и пол, прежде чем приступить к фактическому редактированию. Если группа редактирования готовится заранее или проводит периодические обследования с использованием этих самых элементов, то в их распоряжении могут быть несколько разных матриц динамической интерполяции. Если доля недостоверных или несогласующихся данных очень маленькая, группа редактирования может принять решение использовать для интерполяции только несколько переменных. Если доля ошибок больше, группе редактирования может понадобиться больше переменных для того, чтобы провести большое количество необходимых вычислений.

268. Как правило, небольшие матрицы интерполяции лучше, поскольку с ними легче обращаться при разработке программ редактирования и интерполяции, и их легче использовать во время фактического редактирования. Однако в случае неоднократного использования значений понадобится более крупная и более дифференцированная матрица интерполяции.

IV. РЕДАКТИРОВАНИЕ ДАННЫХ ПЕРЕПИСИ НАСЕЛЕНИЯ

269. Глава IV посвящена редактированию данных переписи населения, включая демографические данные, данные о миграции и социально-экономические характеристики. В спецификациях для этих элементов редактирования учитывается достоверность отдельных данных и согласованность данных переписи населения, а также согласованность данных переписи населения и жилого фонда. Знание некоторых взаимосвязей между этими элементами позволяет планировать процесс редактирования с точки зрения согласованности с тем, чтобы обеспечить более высокое качество данных для последующего составления таблиц. Например, в демографических записях не должно быть 15-летних женщин с десятью детьми или семилетними детьми, посещающими высшую школу.

270. При указании значений для элементов демографических данных группа редактирования должна решить, ставить ли «не указывается»; производить ли статическую интерполяцию (*cold deck*) для значения «не известно» или другого значения; или производить динамическую интерполяцию (*hot deck*) на основе характеристик других лиц или жилищных единиц. Во многих случаях предпочтительнее динамическая интерполяция, поскольку она исключает редактирование на этапе составления таблиц, когда информация есть только в самих таблицах, для принятия решений о неизвестных величинах. Матрицы интерполяции позволяют вводить данные для пробелов, недостоверных сведений или устранения несоответствий, когда нет других смежных элементов с достоверными ответами. В некоторых странах в общенациональном масштабе наблюдается некоторое многообразие демографических характеристик, хотя в большинстве населенных пунктов, если брать их отдельно, такого разброса нет. В других странах среди населенных пунктов могут быть значительные отклонения, особенно если речь идет о проживании в городских и сельских районах. Такое разнообразие необходимо учитывать при разработке матриц интерполяции и, в частности, при установлении исходных величин статической интерполяции. Группа редактирования должна уточнять обстоятельства, при которых следует вычислять данные для пробелов. Их следует брать из предыдущей жилищной единицы с аналогичными характеристиками.

271. В целях содействия обработке данных все демографические записи должны иметь регистрационные

номера. При редактировании структуры, о котором говорилось в главе III, производится проверка на соответствие между порядковым номером и последовательностью регистрационных номеров.

272. Группа редактирования должна редактировать каждую демографическую запись только в отношении подходящих элементов. Редактируемые элементы должны отличаться один от другого в зависимости от городских/сельских, климатических и иных условий. Редактирование желательно производить выборочно в зависимости от этих условий, но практически лишь немногие страны располагают временем и опытом в разработке и применении многомерных матриц для замены недостающих или несогласующихся данных. Еще меньше стран применяют эту дополнительную процедуру на практике.

273. Информация, собранная с помощью переписного листа, также зависит от отдельных демографических характеристик. Например, вопрос о рождаемости задается только женщинам, а об экономической активности — только взрослым.

274. Иногда группа редактирования в отношении определенных элементов должна оставлять запись «не сообщается». В группе редактирования может не быть хорошей основы для исчисления ответов по некоторым характеристикам. Решение оставить ответ «не сообщается» должно увязываться с требованием выдавать соответствующие характеристики в виде таблиц в целях планирования и выработки стратегии. Поскольку случаи «не сообщается» распределяются так же, как и регистрируемые случаи, распределение ответов «не сообщается», когда планирующим органам необходима выборочная информация, не должно вызывать трудности. Однако если случаи «не сообщается» в какой-то степени асимметричны, интерполяция после компилирования может быть проблематичной, особенно в отношении небольших районов или конкретных типов условий. Например, если респонденты из числа молодых девушек отказываются раскрывать информацию о рождаемости и данные о рождаемости не указываются, то редактирование не поможет получить эту информацию.

275. Редактировать демографические данные, как правило, сложнее, чем данные о жилом фонде, поскольку сводные таблицы, как правило, намного сложнее. Большинство стран составляют отдельные характеристики

жилого фонда только по разным территориальным уровням, но сводные таблицы с демографическими элементами могут быть многоуровневыми. Как разъяснялось выше, страны предпочитающие не использовать динамическую интерполяцию, должны определить идентификатор для значений «не известно», чтобы пользоваться им в случае недостоверных или несогласующихся ответов.

276. Если страны используют динамическую интерполяцию, группы редактирования должны разработать простые матрицы интерполяции с измерениями, которые дифференцируют демографические характеристики. Для большинства стран наилучшими первичными переменными для динамической интерполяции являются возрастная группа и пол, и их следует редактировать в первую очередь. Национальные бюро переписи/статистические управления, использующие метод редактирования переменных по групповому принципу, должны одновременно редактировать данные о возрасте, поле и другие переменные, такие как степень родства и семейное положение. К другим элементам, которые могут оказаться полезными при динамической интерполяции, относятся уровень образования и состояние занятости.

277. Группы редактирования должны работать очень тщательно, чтобы не исказить данные при интерполяции. Не следует исходить из того, что неинтерполированные и интерполированные данные будут обязательно иметь одинаковое распределение. Зачастую неизвестные данные искажаются сами по себе. Например, люди старшего возраста менее расположены сообщать свой возраст, чем молодежь.

А. ДЕМОГРАФИЧЕСКИЕ ХАРАКТЕРИСТИКИ

278. Данные о степени родства, поле, возрасте и семейном положении каждого человека являются основными для любой переписи и, вероятно, их следует редактировать вместе. Структура возраста и пола по совокупностям или подсовкупностям имеет основополагающее значение практически при любом планировании, которое ведется на основе данных переписи населения. Эти элементы также необходимы для составления наполненных информацией таблиц, поскольку, фактически, все другие виды анализа основаны на сводных таблицах с другими переменными по возрасту и полу.

279. Принцип использования переменных по групповому принципу (Фелледи–Холт) при редактировании данных переписи населения и жилого фонда был представлен в главе II настоящего *Руководства*. Поскольку демографические переменные являются неотъемлемой

частью планирования всех переписей, то этот подход следует использовать, если позволяют время и знание дела. Качество всего набора данных почти наверняка выиграет от того, если для определения ошибок или несоответствий будет применяться в приоритетном порядке редактирование по возрасту, полу и другим отдельным переменным. В первую очередь подвергаются редактированию элементы, где больше ошибок, а затем элементы, содержащие меньше ошибок или несоответствий.

1. Родство (P2A)⁵

280. Элемент родства используется для того, чтобы легче было определить структуру домохозяйства и семьи. Он указывается где-то в самом начале большинства переписных листов переписи и обследования и помогает удостовериться в том, что каждый, кто проживает в жилищной единице, зарегистрирован. Счетчик и респондент используют информацию о родстве между членами домохозяйства, чтобы убедиться в том, что никто не пропущен. Элемент родства помогает также произвести проверку на согласованность данных о возрасте и поле членов домохозяйства. Об определении единственного главы домохозяйства и не более одного супруга(и) (в неполнородном обществе) говорится в пунктах 250–264.

а) Редактирование данных о родстве

281. Поскольку статистические данные о степени родства приобретают все более важное значение, необходимо более тщательно подходить к разработке приемов редактирования, позволяющих составлять различные виды таблиц с указанием характеристик семьи как единицы учета. В этом, несомненно, поможет разработка в первую очередь соответствующих кодов родства (см. пункты 507–610 о записях типа семьи).

282. Когда данные о родстве вычислить нельзя, а динамическая интерполяция не применяется, то недостоверные или несогласующиеся ответы необходимо относить к категории «не известно». При динамической интерполяции родство можно вычислить из матрицы интерполяции по возрасту и полу или другим соответствующим характеристикам. Матрицы интерполяции не должны выдавать такие данные о родстве, которые

⁵ Номера в скобках в главах IV и V, указанные в *Принципах и рекомендациях в отношении переписи населения и жилого фонда, первое пересмотренное издание*, Статистические документы, серия M, № 67/Rev. 1 (издание Организации Объединенных Наций в продаже под номером R.98.XVII.8). Исходя из этого, «P2A» соответствует элементу переписи 2 а в «Перечне вопросов, задаваемых при переписи населения» в части II, главе V указанного издания, стр. 65–66. Жилищные единицы сопровождаются предложенными КОДАМИ, выведенными из основных и дополнительных вопросов, указанных в части II, главе VI *Принципов и рекомендаций*, пункты 2.293 и 2.416.

противоречили бы уже установленным отношениям в домохозяйстве. Например, не следует исчислять вторых и третьих супругов даже в полигамных домохозяйствах, если только группа редактирования не решит использовать такой прием.

b) *Когда главу необходимо указывать первым*

283. Если глава не указывается как первое лицо, то при редактировании структуры, о чем говорилось в главе III, выяснится, что для того, чтобы отследить положение главы, можно применить указатель. Если группа редактирования захочет указать главу как первое лицо, то его можно переместить на первую позицию или изменив порядок лиц, или оставив домохозяйство на месте, но изменив степень родства, как отмечается в главе о редактировании структуры. Первый метод требует значительного опыта программирования; второй метод может нарушить целостность набора данных, если не принять чрезвычайных мер.

c) *Когда степень родства кодируется без соблюдения определенного порядка*

284. Иногда счетчики собирают данные о родстве «наоборот»: вместо того, чтобы собирать данные о родстве каждого члена домохозяйства по отношению к его главе, они собирают эти данные о главе по отношению к каждому члену домохозяйства. Отсюда степень родства третьего лица указывается не как «ребенок», а как «родитель». И тогда домохозяйство может заканчиваться с указанием четырех родителей вместо четырех детей. Когда группа редактирования обнаружит, что такая проблема встречается систематически, она должна выработать решение, чтобы не вносить слишком больших нарушений в данные о домохозяйстве.

d) *Супруги в полигамном браке*

285. Если редактирование структуры будет проводиться, как указано в главе III, то оно обеспечит проверку моногамных домохозяйств на то, чтобы там был «один и только один» глава и «не более одного супруга». Что же касается полигамных домохозяйств, то здесь группа редактирования должна решить, когда полигамные отношения разрешены, а когда нет. Иногда домохозяйство, казалось бы, имеет полигамные отношения, а на самом деле это не так. Например, в домохозяйстве может быть указан глава и супруга, а другая пара зарегистрирована как «супруги» по отношению друг к другу, и в результате получается три супруга(и). При редактировании важно убедиться, что вторая пара — это действительно не отец и мать, не сын и невестка,

не сестра и зять или какая-либо иная комбинация. Иногда эти отношения можно определить с некоторой долей определенности, а иногда они не поддаются определению. Когда вышеуказанным отношениям присваивается код, группа редактирования должна исходить из того, чтобы произвести соответствующие исчисления. Когда в полигамных домохозяйствах дополнительные супруги являются фактическими супругами, то при редактировании следует произвести проверку по полу и, возможно, по возрасту.

e) *Появление нескольких родителей*

286. В данных о домохозяйстве не должно быть указано более двух «родителей», и они должны быть противоположного пола. Если появляется более двух родителей, то, вероятно, дополнительных родителей следует отнести к категории «прочие родственники». Иногда при переписи или обследовании присваивается код для «родителя» или «кого-либо из родителей» мужа или жены, что позволяет получить до четырех «родителей» вместо двух или более двух родителей каждого пола.

f) *Когда при переписи собираются данные о родстве по признаку пола*

287. Иногда при проведении переписи или обследования собираются данные о родстве по признаку пола: «муж» и «жена» отдельно, вместо характеристики «супруг(а)»; «сын» и «дочь» вместо категории «ребенок» и т. д. Если эти ответы не отредактировать, то в таблицах могут появиться данные, где дочери будут отнесены к мужскому полу, а мужа — к женскому. В этом случае группа редактирования должна установить приоритетность редактирования, то есть отдать предпочтение или родству или полу. В некоторых случаях, например, в отношении мужа и жены важнее отредактировать эти данные по отношению к другим, например, категории «маленький ребенок».

g) *Когда данные о родстве и семейном положении не согласуются между собой*

288. Данные о степени родства и семейном положении должны согласовываться между собой, если происходит наложение: лица, которые указывают в графе родство «супруг(а)», в графе «семейное положение» должны указывать «в браке». Группа редактирования выбирает, какую переменную изменить, если эти данные не согласуются между собой. Иногда ясности в отношении родства нет, и поэтому нужно тщательно подойти к разработке спецификаций редактирования. Например, во многих странах зять является одновременно

братом супруга(и) (и поэтому не должен быть женат), а также супругом(ой) сестры (брата) (должен быть женат).

289. В настоящее время появляются и другие, более новые проблемы в связи с сообщением данных о родстве. Когда два лица противоположного пола живут вместе, не состоя в брачных отношениях, то код отношений может быть или «не состоящий в браке партнер» или «супруг(а)». Если при переписи или обследовании указывается код для не состоящего в браке партнера, то тогда соответствующее семейное положение не должно указываться как «состоящее в браке», если данное лицо не состоит в браке с кем-либо еще помимо того, с кем они живут.

290. Аналогичная ситуация возникает, когда лица одного и того же пола проживают вместе, находясь или в любовных или не в любовных отношениях. Лица, не находящиеся в любовных отношениях, могут быть закодированы как «сожители» или «не родственники». В отношении тех, кто состоит в любовных отношениях, для некоторых стран может быть приемлема категория «не состоящий в браке партнер». В этом случае группа редактирования тоже должна решить, как указывать соответствующее семейное положение.

2. Пол (P3A)

291. Пол — это одна из самых простых характеристик с точки зрения сбора данных, но требует некоторого осмысления при редактировании. Она является одной из важнейших переменных, поскольку большинство демографических характеристик анализируются на основе пола. Исчисление данных о поле требует некоторого сопоставления с другими переменными. В некоторых случаях пол того или иного лица должен определяться на основе разницы между полами связанных с ним лиц, как правило, главы домашнего хозяйства или супруга(и), а также между родителями и родни со стороны мужа или жены. Вероятно, не следует отмечать пол как «недостовверный» или «не известен», поскольку он имеет столь важное значение в качестве переменной. В связи с этим следует подумать, как лучше всего получить результаты, сопоставимые с реальной ситуацией в стране.

292. Если то или иное лицо не является главой домохозяйства или супругом(ой) главы, то других лиц, на которых можно сослаться нет; поэтому следует проверять другие элементы в записи данного лица. И если есть данные о рождаемости, то женщине следует присвоить код. Однако, если, например, пол данного лица отсутствует или недостоверен, но есть супруг или супруга, пол которого или который указан, то при редактировании можно исчислить противоположный пол.

a) *Когда код пола достоверен, но глава и супруг(а) одного пола*

293. В тех случаях, когда встречаются большие противоречия, код пола необходимо менять, даже если есть достоверный код. Например, запись показывает, что есть состоящая в браке пара и при этом домохозяйство имеет главу и супруга или семейную пару в подсемействе. Если оба лица во второй паре регистрируют один и тот же пол, то можно использовать информацию о рождаемости и другие элементы данных, для того чтобы определить кто мужчина, а кто женщина. После этого можно исправить ошибочную запись.

b) *Когда в отношении мужчины есть информация о рождаемости, а в отношении взрослой женщины нет*

294. При редактировании можно обнаружить, что в отношении мужчины есть информация о рождаемости и/или детях, то есть ошибка могла возникнуть из-за личного номера матери или другой аналогичной переменной. Если супруг не указан, то пол можно поменять на женский, вместо того чтобы удалять информацию о рождаемости. И также в отношении взрослой женщины, если нет информации о рождаемости и не указаны дети, то можно поменять данные на категорию «мужской» при определенных обстоятельствах, установленных группой редактирования.

c) *Если код пола является недостоверным и указан супруг(а)*

295. Если графа «пол» не заполнена или содержит недостоверные данные, то программа редактирования должна использовать записи о степени родства с главой домохозяйства и пол супруга(и), если он достоверен, для того чтобы определить правильный код. Если степень родства с главой домохозяйства указана как «глава домохозяйства», то программа проверяет, указан ли супруг(а) [путем проверки данных о другом члене домохозяйства, в отношении которого указано «супруг(а)»]. При определении кода пола супруга код противоположного пола присваивается главе домохозяйства.

d) *Когда код пола для супруга(и) недостоверен*

296. Если степень родства с главой домохозяйства указана как «супруг(а)» и дается пол главы домохозяйства, программа присваивает этому лицу пол, противоположный по отношению к главе домохозяйства.

е) *Если код пола недостоверен и есть информация о женщине*

297. По многочисленным признакам в переписном листе можно узнать, является респондент женщиной. Если программа еще не определила пол лица, и есть какие-либо показатели, касающиеся женщины, то записи для этого лица следует присвоить женский пол. Например, если в редактируемых данных есть один или несколько элементов, касающихся рождаемости, то в этом случае пол можно указать как женский. К данным о рождаемости относятся такие, как дети, родившиеся живыми, дети, живущие в этом домохозяйстве, дети, живущие в других местах, умершие дети и дети, родившиеся живыми за последние 12 месяцев. Еще может быть так, что это лицо может быть матерью кого-то другого в данном домохозяйстве, с тем чтобы номер строки для этого лица равнялся номеру строки матери другого лица в данном домохозяйстве.

ф) *Когда код пола недостоверен и это лицо является супругом*

298. Если указанное лицо является мужем кого-либо еще в домохозяйстве, исходя из элемента, показывающего номер строки мужа, то в графе «пол» следует указать «мужской».

г) *Когда код пола недостоверен, а информации для определения пола недостаточно*

299. Если группа редактирования совсем не использует метод динамической интерполяции, то значение «пол не известен» необходимо присваивать. К сожалению, это означает, что во всех таблицах должна быть дополнительная колонка или дополнительный ряд или группа колонок или рядов для лиц, чей пол не известен. Поскольку пол — это двоичная переменная, то величины можно присваивать попеременно, начиная с одной из двух и используя при этом противоположный пол для второй недостоверной записи, и затем продолжать в том же порядке.

h) *Замечание об исчисленных соотношениях полов*

300. Если применяется метод статической интерполяции, то женский пол присваивается, как правило, чаще. Только у взрослых женщин есть данные о рождаемости, и их выбор искажается несколько случайно. В связи с этим, если имеющейся информации недостаточно, то лицо, о котором нет такой информации, чаще всего мужчина, а не женщина. Следовательно, необходимо предусматривать разработку матриц интерполя-

ции, где учитывалось бы общее соотношение между полами.

3. *Дата рождения и возраст (P3B)*

301. Возраст является одной из наиболее трудных характеристик с точки зрения сбора и редактирования. Однако вполне вероятно, что он является важнейшей переменной, поскольку все демографические характеристики анализируются, исходя из возраста. Редактирование данных о возрасте требует широкого сопоставления с другими переменными. В большинстве случаев в основе исчисления возраста должны лежать хранящиеся в памяти различия в возрасте родственников. Если, исходя из этого, возраст исчислить нельзя, то тогда следует использовать другие характеристики в записи об этом лице. Редактирование, возможно, потребует нескольких матриц интерполяции, включая возраст по полу, семейному положению, степени родства и посещению школы; возрастные различия между матерью и ребенком; возрастные различия между мужем и женой; и возрастные различия между главой домохозяйства и супругой(ом).

а) *Возраст и дата рождения*

302. При структурном редактировании возраст вычисляется по дате рождения. Однако при этом следовало бы рассмотреть разницу между возрастом и датой рождения. Как указывается в *Принципах и рекомендациях* (Организация Объединенных Наций, 1998 год, пункт 2.88), информацию о возрасте можно получить либо исходя из даты (года, месяца и дня) рождения, либо, задав прямой вопрос о возрасте на момент последнего дня рождения.

303. Дата рождения дает более точную информацию и должна использоваться по возможности при всех обстоятельствах. Если точный день или месяц рождения неизвестны, то их можно заменить указанием на время года. Вопрос о дате рождения является уместным, когда людям известна дата их рождения в соответствии как с солнечным, так и с лунным календарем, или когда годы обозначаются цифрами, или в традиционной народной культуре имеют названия в рамках регулярного цикла. Однако в отношении календарной системы, на которой основывается дата рождения, чрезвычайно важно, чтобы между счетчиком и респондентом было четкое взаимопонимание. Если возможны ответы некоторых респондентов со ссылкой на календарную систему, отличную от той, которую использовали другие респонденты, то в переписном листе необходимо предусмотреть отметку о применяемой календарной системе. Счетчикам не рекомендуется пытаться перево-

дать даты из одной системы в другую. Необходимый перевод наилучшим образом может быть осуществлен при редактировании данных с помощью компьютера (Организация Объединенных Наций, 1998 год, пункт 2.89).

304. По ряду причин на прямой вопрос о возрасте, возможно, будет дан неточный ответ. Даже если все ответы основываются на одном и том же методе определения возраста, респондент может не понять, о каком возрасте требуются сведения: на момент последнего дня рождения, следующего дня рождения или ближайшего дня рождения. Кроме того, могут возникать и другие проблемы: возраст может быть округлен до ближайшей цифры, оканчивающейся на 0 или 5; может указываться приблизительный возраст без должного указания на это или умышленно даваться неправильный ответ (Организация Объединенных Наций, 1998 год, пункт 2.90).

305. Многие национальные бюро переписи/статистические управления регистрируют или дату рождения, или возраст, а не то и другое вместе. Как отмечается в *Принципах и рекомендациях* (Организация Объединенных Наций, 1998 год), возраст в завершённые годы имеет очень важное значение: он используется при редактировании многих данных, а также в качестве измерения для многих матриц интерполяции. Не менее важно и то, что многие страны строят свою политику, исходя из данных о возрасте, поэтому необходимо прилагать все усилия, чтобы получать данные о возрасте наилучшего качества. Однако даже в идеальных ситуациях, некоторые данные о возрасте указываться не будут. Следовательно, необходимо обеспечить, чтобы возраст рассчитывался должным образом и согласовывался с другими ответами отдельных членов домохозяйства.

b) *Взаимосвязь между датой рождения и возрастом*

306. Если данные о возрасте не собирались отдельно от даты рождения, то возраст следует вычислить во время редактирования структуры. Редактирование данных о возрасте наравне с другими данными обеспечит всестороннюю проверку на согласованность данных в записях и между ними. Но сначала нужно вычислить возраст, исходя из даты рождения и даты проведения переписи. Необходимо проверить возраст после его вычисления, исходя из даты рождения, чтобы убедиться, что он вписывается в рамки переписи, если исходить из даты ее проведения. Возраст детей, родившихся в год проведения переписи, но после даты ее проведения, будет исчислен как 1, и его необходимо исправить. Детей, зарегистрированных после даты проведения переписи, вероятно, следует из нее исключать. Однако если после проверки обнаруживается, что дата рожде-

ния указана неправильно при регистрации счетчиками или при обработке, нужно использовать другие переменные для получения более точной оценки возраста.

c) *Когда рассчитанный возраст выходит за верхний предел*

307. Для переписей 2000 года и дальше большинство стран предпочитают записывать год рождения четырьмя цифрами. Приемлемый диапазон будет примерно с 1800-х годов, возможно, с 1880 года и позже, и до года переписи. Хотя для компьютера достаточно и трех цифр, использование этих трех цифр для указания года может ввести в заблуждение и счетчиков и работников статистических управлений. Иногда рассчитанный год будет выходить за верхнюю границу определенных переписью лет, и тогда потребуется вносить коррективы. Если перепись проводится в 2000 году, а человек сообщает, что родился в 1850 году, то расчетный возраст 150 лет, вероятно, выйдет за приемлемые пределы и его необходимо будет изменить.

d) *Редактирование данных о возрасте*

Рамка 4. Редактирование и исчисление возраста

При редактировании и исчислении возраста следует сделать следующее:

- a) приписать возраст на месте пробелов;
- b) проверить минимальный возраст тех, кто когда-либо состоял в браке;
- c) проверить минимальный возраст главы домохозяйства;
- d) проверить минимальный возраст родителей; и
- e) произвести любые другие проверки, характерные для каждой страны.

308. С помощью программы редактирования следует проверить согласованность указанного возраста данного лица с указанным возрастом его матери, отца или ребенка. Редактирование должно обеспечить минимальную разницу в годах между возрастом матери или отца и возрастом ребенка. В случае исчисления возраста проверки на согласованность должны производиться по таким данным, как годы, прожитые в данном районе, продолжительность проживания и самый старший законченный класс школы (уровень образования). Все эти проверки следует производить до изменения возраста или до указания исчисленного возраста.

309. Редактирование следует начинать с проверки на достоверность. Если возраст указан правильно, специалисты, возможно, пожелают проверить, согласуется ли возраст данного лица с возрастом его/ее матери (если в

домохозяйстве указана мать данного лица) и дети данного лица (если данное лицо является женщиной и имеет детей в домохозяйстве). Если данные о возрасте не согласуются между собой, то следует сделать заметку о возрасте этого лица и изменить его позднее.

е) *Редактирование данных о возрасте, если есть глава домохозяйства и супруг(а)*

310. На следующем этапе редактирования необходимо определить, есть ли супруг(а). Если есть, то его (ее) возраст следует проверить на достоверность (по меньшей мере, 10 лет в зависимости от установленного в стране минимального возраста вступления в брак). Если возраст не согласуется и если применяется динамическая интерполяция, то программа будет использовать специальное значение, полученное путем интерполяции, исходя из разницы в возрасте между мужем и женой. Диапазон возрастных различий меньше, чем сам возраст, поэтому в матрице интерполяции, используемой в программе, сохранится разница в возрасте (из предыдущих записей) мужа и жены. Это значение прибавляется или вычитается из возраста супруга, чтобы получить исчисленный возраст.

311. Для того чтобы этот исчисленный возраст согласовывался с другими характеристиками, в матрицу интерполяции следует также включить семейное положение, длительность проживания и самый старший законченный класс школы. Исключение этих переменных может привести к тому, что исчисленный возраст будет меньше количества лет, которые человек прожил в данном месте, или меньше уровня обучения. Например, матрица интерполяции может показать возраст 8 лет, но данное лицо уже может указать, что они прожили в этом месте 10 лет. Без других переменных, когда программа редактирования производит редактирование данных о годах проживания, другая матрица интерполяции изменит годы проживания с правильного значения на неправильное.

ф) *Редактирование данных о возрасте главы, когда нет супруга(и), но есть ребенок*

312. Когда сопоставить возраст супруга(и) при определении возраста главы домохозяйства нельзя, то программа может проверить степень родства. Если это отношение — «глава домохозяйства», программа редактирования может проверить другие записи о домашнем хозяйстве (если таковые имеются) в отношении сына или дочери, возраст которых, как известно, указан правильно. Программа проверяет возраст сына или дочери и исчисляет возраст этого лица с использованием динамической интерполяции по «возрастной разнице» аналогично тому, как это делается в отношении мужа и

жены. Как и ранее при исчислении возраста принимаются во внимание продолжительность проживания и высший уровень образования. Завершенный возраст, таким образом, будет согласовываться с этими переменными и не будет очевидных ошибок, возникающих при включении количества лет, прожитых в данном районе, и самого старшего законченного класса школы в матрицу интерполяции.

г) *Редактирование данных о возрасте главы домохозяйства при наличии родителя главы домохозяйства*

313. Если какое-либо лицо не подпадает ни под одну из описанных выше категорий, программа может включить поиск родителя этого лица в домохозяйстве. При обнаружении родителя такого лица можно подсчитать возраст с помощью матрицы интерполяции, используя разницу в возрасте. Разница в возрасте между ребенком и родителем обычно колеблется в более широких пределах, чем разница в возрасте между мужем и женой. По этой причине такое редактирование в программе применяется только тогда, когда методика, основанная на разнице в возрасте между мужем и женой, не дает результатов. В исчисляемых возрастных показателях следует учитывать показатели образования, последнюю законченную ступень учебного заведения, продолжительность проживания в данном районе, семейное положение, рождаемость и экономическую активность. В программе следует делать допущение о том, что какое-либо лицо находится, по меньшей мере, в минимально допустимом возрасте, если он/она когда-либо состоял(а) в браке, имеет детей или сообщает о любой экономической активности.

h) *Редактирование данных о возрасте главы домохозяйства при наличии внука*

314. Если какое-либо лицо не подпадает ни под одну из вышеупомянутых категорий, программа может включить поиск внука данного лица в домохозяйстве. При обнаружении внука данного лица можно исчислить возраст с помощью матрицы интерполяции, используя разницу в возрасте. Разница в возрасте между главой домохозяйства и его внуком колеблется в более широких пределах, чем разница в возрасте между мужем и женой или чем разница в возрасте между главой домохозяйства и его ребенком. По этой причине такое редактирование применяется в программе только тогда, когда методика, основанная на разнице в возрасте между мужем и женой и разнице в возрасте между главой домашнего хозяйства и его ребенком, не дает результатов. В исчисляемом возрасте следует учитывать показатели образования, включая последнюю законченную ступень учебного заведения, продолжительность проживания в данном районе, семейное положение, рож-

даемость и экономическую активность. В программе следует делать допущение, что какое-либо лицо находится, по меньшей мере, в минимально допустимом возрасте, если он/она когда-либо состоял(а) в браке, имеет детей или участвует в какой-либо экономической деятельности.

i) *Редактирование данных о возрасте главы домохозяйства, когда нет других возрастных показателей*

315. Если какое-либо лицо не подпадает ни под одну из вышеупомянутых категорий, программа может включать поиск другого родственника или лица, не состоявшего в родственных связях с главой домашнего хозяйства. При наличии такого лица, и если такое лицо находится в регистрируемом возрасте, группа редактирования должна решить, стоит ли ей включать любую имеющуюся информацию в матрицу интерполяции, используя различия в возрасте. Однако такие различия в возрасте между главой домашнего хозяйства и другими родственниками или лицами, не состоявшими с ним в родственной связи, колеблются в таких широких пределах, что группа редактирования может вообще отказаться от этих усилий и просто использовать другие переменные для динамических вычислений возраста главы домашнего хозяйства. В любом случае, в программе такое редактирование применяется только после того, как методика, основанная на разнице в возрасте между мужем и женой, между главой домохозяйства и ребенком, между главой домохозяйства и родителем и между главой домохозяйства и внуком, не дает каких-либо результатов. Однако если исчисляемый возраст определен, в нем должны учитываться показатели образования, последняя законченная ступень учебного заведения, продолжительность проживания в округе, семейное положение, рождаемость и экономическая активность. В программе следует делать допущение, что любое лицо находится, по меньшей мере, в минимально допустимом возрасте, если он/она состоял(а) когда-либо в браке, имеет детей или участвует в какой-либо экономической деятельности.

j) *Редактирование данных о возрасте супруга(и), когда уже определен возраст главы домохозяйства*

316. Редактирование возрастных показателей супруга(и) обычно приводится одновременно с редактированием возрастных показателей главы домохозяйства, поскольку для совместного редактирования требуется информация от обоих лиц. Если, однако, проводится раздельное редактирование, когда возраст супруга(и) указан неверно или не согласуется с другими переменными, для определения наиболее приближенной оценки возраста супруга(и) следует применять матрицу ди-

намической интерполяции, используя разницу в возрасте с главой домохозяйства и другие переменные. Как и в других случаях при определении возраста следует учитывать показатели образования, включая последнюю законченную ступень учебного заведения, продолжительность проживания в данном районе, семейное положение, рождаемость и экономическую активность. В программе должно делаться допущение, что лицо находится, по меньшей мере, в минимально допустимом возрасте, если он/она когда-либо состоял(а) в браке, имеет детей или участвует в какой-либо экономической деятельности.

k) *Редактирование данных о возрасте других супружеских пар в домохозяйстве, когда известен возраст одного из супругов.*

317. При редактировании следует, прежде всего, установить, принадлежит ли данная запись лицу, состоявшему в браке. При положительном результате программа может включать поиск супруга(и) данного лица среди других записей о домохозяйстве. Если поиск супруга(и) оказывается безрезультатным, программа приступает к следующей части редактирования. Если с помощью поиска обнаруживается один из супругов, возраст супруга(и) следует проверять на достоверность (минимальный возраст X лет, в зависимости от установленного в стране минимального брачного возраста). Если возраст не согласуется с другими данными и если применяется динамическая интерполяция, то тогда в программе используется специально вычисленное значение, полученное по разнице в возрасте между мужем и женой. Разница в возрасте колеблется не в столь широких пределах, как сами возрастные показатели, поэтому в матрицу интерполяции программы заносится разница в возрасте (из предыдущих записей) между мужем и женой. Это значение прибавляется или вычитается из возраста супруга(и) данного лица для определения исчисляемого возраста.

318. Чтобы исчисляемый возраст согласовался с другими характеристиками, в матрицу интерполяции следует также включать семейное положение, продолжительность проживания и наивысший полученный уровень образования. Исключение этих переменных может привести к тому, что исчисляемый возраст будет меньше количества лет, которые данное лицо прожило в данном месте или меньше возрастного показателя, который соотносится с уровнем образования.

l) *Редактирование данных о возрасте ребенка, когда уже определен возраст главы домохозяйства*

319. Если речь идет о сыне или дочери главы домохозяйства, исчисляемый возраст может быть получен с

помощью возрастного показателя главы домохозяйства, разницы в возрасте, продолжительности проживания и образовательной подготовки. Опять таки, при исчислении возраста следует учитывать показатели образования, включая последнюю законченную ступень учебного заведения, продолжительность проживания в данном районе, семейное положение, рождаемость и экономическую активность. В программе следует делать допущение, что какое-либо лицо находится, по меньшей мере, в минимально допустимом возрасте, если он/она когда-либо состоял(а) в браке, имеет детей или участвует в любой экономической деятельности.

m) *Редактирование данных о возрасте родителя, когда уже известен возраст главы домохозяйства*

320. Если речь идет о родителе главы домохозяйства, исчисляемый возраст может быть получен с помощью возрастного показателя главы домохозяйства, разницы в возрасте, продолжительности проживания и образовательной подготовки. При подсчете возраста следует учитывать показатели образования, включая последнюю законченную ступень школы, продолжительность проживания в округе, семейное положение, рождаемость и экономическую активность. В программе следует делать допущение, что любое лицо находится, по меньшей мере, в минимально допустимом возрасте, если он/она когда-либо состоял(а) в браке, имеет детей или участвует в какой-либо экономической деятельности.

n) *Редактирование данных о возрасте внука, когда уже определен возраст главы домохозяйства*

321. Если речь идет о внуке главы домохозяйства, исчисляемый возраст можно получить с помощью возраста главы домохозяйства, разнице в возрасте, продолжительности проживания и образовательной подготовки. Опять таки, при исчислении возраста следует принимать во внимание показатели образования, включая последнюю законченную ступень учебного заведения, продолжительность проживания в округе, семейное положение, рождаемость и экономическую активность. В программе следует делать допущение, что лицу исполнилось, по меньшей мере, двенадцать лет, если он/она когда-либо состоял(а) в браке, имеет детей или участвует в какой-либо экономической деятельности.

o) *Редактирование данных о возрасте всех других лиц*

322. Группе редактирования следует определить соответствующие матрицы интерполяции для других родственников и неродственников в домохозяйства. Руководящие принципы будут зависеть от конкретной пере-

писи или обследования и социально-экономических характеристик страны. Например, любое лицо, которое когда-либо состояло в браке, имело детей или участвовало в какой-либо экономической деятельности наверняка должно быть, по меньшей мере, в возрасте, равному минимальному возрасту, который устанавливают некоторые страны. На основе этой информации, если применяется метод динамической интерполяции, значение, полученное по матрице интерполяции, должно быть ниже минимального возраста. Аналогичным образом, если лицо достигает школьного возраста, имеет какую-либо школьную подготовку или умеет читать и писать, но не является главой домохозяйства, никогда не состояло в браке и не участвует в какой-либо экономической деятельности, то тогда данное лицо следует отнести к группе, возраст которой меньше минимального возраста взрослых, но больше или равен минимальному школьному возрасту. После этого можно определить матричное значение для лиц, младше минимального школьного возраста. При всей своей несовершенности данный метод ограничивает разброс значений, которые могут приниматься в матрице интерполяции.

4. Семейное положение (P3C)

323. В издании *Принципы и рекомендации в отношении переписи населения и жилого фонда* (Организация Объединенных Наций, 1998 год, пункты 2.96–2.98) семейное положение определяется как положение каждого человека по отношению к брачному законодательству или брачным обычаям данной страны. Категории семейного положения, подлежащие определению, включают, но не ограничиваются только этим, следующие категории: *a)* одинокие лица, никогда не состоявшие в браке; *b)* лица, состоящие в браке; *c)* вдовы и вдовцы, не вступившие в новый брак; *d)* разведенные, не вступившие в новый брак; *e)* супруги, живущие раздельно. В некоторых странах категория *b* может потребовать выделения подкатегории лиц, которые по договору состоят в браке, но еще не живут в качестве мужа и жены. Во всех странах категория *e* должна включать супругов, живущих раздельно как по решению суда, так и де-факто, что можно отразить в отдельных подкатегориях, если это необходимо. Независимо от того, что пары, живущие раздельно, по-прежнему считаются супругами (поскольку они не имеют права вновь вступить в брак), ни одна из подкатегорий категории *e* не должна включаться в категорию *b*. В некоторых странах необходимо будет учитывать брачные союзы, основанные на обычном праве (которые, согласно обычному праву, признаются законными и имеющими юридическую силу), а также незаконные брачные союзы, часто называемые фактическим браком.

а) Редактирование семейного положения

324. Группа редактирования должна принять решение относительно соответствующего минимального возраста вступления в первый брак для переписи или обследования. Минимальный возраст вступления в первый брак (примерно X лет) может отличаться в различных частях страны или в различных этнических группах. Если, например, сельское население вступает в брак раньше городских жителей, правила редактирования должны учитывать этот факт. Обычно национальное бюро переписи/статистическое бюро устанавливает этот возраст при самых ранних браках до регистрации для того, чтобы этот вопрос задавался только лицам старше установленного возраста. Лица младшего возраста попадают в категорию лиц, «никогда не состоявших в браке», автоматически. Если, однако, всем задается вопрос о семейном положении, группа редактирования должна разработать соответствующую программу редактирования для всего населения.

б) Отнесение к категориям «семейное положение» в случаях, когда не применяется метод динамической интерполяции

325. Хотя семейное положение следует табулировать только для лиц в возрасте X лет и старше, где X равно минимальному возрасту вступления в первый брак, группа редактирования должна определить целесообразность и объем редактирования. Если в стране для недостоверных или несогласующихся ответов используются только графы «не указано» или «не известно», то при обнаружении недостоверных или несогласующихся первичных данных неправильные ответы должны заменяться кодом для графы «не указано». Если для лиц младше X лет ответ «никогда не состоял в браке» пропущен, его следует интерполировать.

в) Отнесение к категориям «семейное положение», когда применяется метод динамической интерполяции

326. Если применяется метод динамической интерполяции, при редактировании семейного положения следует: а) интерполировать значения, когда вводимые данные находятся за пределами интервала; и б) проводить проверку на соответствие между зарегистрированным семейным положением, родством и возрастом.

д) Супруг(а) должен(на) состоять в браке

327. Всем лицам, которым присвоен код «супруг» в категории родства должен присваиваться код лиц, состоящих в браке.

е) Супруг(а) супружеской пары

328. Если номер строки супруга(и) лица А (лицо В) является переменной величиной, то тогда по отношению к лицу В лицо А должно квалифицироваться в качестве супруга(и); кроме того, оба лица А и В должны состоять в браке (как упоминается в пункте 327) и принадлежать к противоположному полу.

ф) Если супруг(а) является главой домохозяйства он/она должен(должна) состоять в браке

329. Если для семейного положения нет данных, но есть данные, в которых отношение к главе домохозяйства определяется как глава, в программу следует включить проверочный поиск присутствующего супруга (путем проверки отношения к другим членам домашнего хозяйства). При наличии супруга(и) программа определяет семейное положение главы домохозяйства как «состоящего в браке».

г) Глава домохозяйства без супруга(и) с детьми и без детей

330. При отсутствии супруга(и) и если данное лицо относится к мужскому полу и имеет детей, в программе проводится интерполяция семейного положения по возрасту при наличии детей. Если нет детей, программа может интерполировать семейное положение по возрасту без наличия детей. Мужчина, который является главой домохозяйства, но жена которого не живет в домохозяйстве, со всей определенностью относится к категории разведенных, живущих раздельно или вдовцов.

х) Если все остальные вычисления не дают результатов, интерполировать

331. Затем для лиц с кодами, выходящими за пределы диапазона, которым нельзя присвоить код на основе вышеупомянутых проверок, проводится проверка возраста. Если возраст зарегистрирован достоверно, т. е. младше X лет, то при этом следует присваивать категорию «никогда не состоял в браке». Во всех остальных случаях элемент данных следует классифицировать с помощью матрицы интерполяции. Следует составить матрицу интерполяции по полу и возрасту (двумерную), и полу, возрасту и родству (трехмерную), или по полу, возрасту, родству и количеству родившихся живыми детей (четырёхмерную). Опять таки, группам редактирования следует установить порядок редактирования, и таким образом при разработке матриц интерполяции важно помнить о том, какие элементы были отредактированы, а какие нет. Если до семейного по-

ложения были отредактированы только пол и родство, матрица интерполяции должна предусматривать графу «не сообщается» в других элементах.

i) *Зависимость между возрастом и семейным положением молодых людей*

332. Для всех лиц, сообщивших о достоверном семейном положении, кроме категории лиц, «никогда не состоявших в браке», в отношении возраста следует провести проверку на согласованность. Все лица, когда-либо состоявшие в браке, должны быть в возрасте X или старше, где X равно минимальному брачному возрасту, установленному в конкретной стране. Если возраст составляет менее X лет или пропущен, следует провести новую проверку на согласованность по другим основным переменным (таким как количество родившихся живыми детей или экономическая активность). Если данные по этим элементам недостоверны, то в графе семейное положение следует указывать «никогда не состоял в браке»; во всех остальных случаях семейное положение не должно меняться.

5. *Возраст вступления в первый брак (P4F)*

333. Согласно изданию *Принципы и рекомендации* (Организация Объединенных Наций, 1998 год, пункт 2.142) «датой первого брака» являются число, месяц и год, когда был заключен первый брак. Странам, где трудно установить дату первого брака, рекомендуется собирать информацию о возрасте вступления в брак или о том, сколько лет назад был заключен брак (продолжительность брака). Важно, чтобы сюда были включены не только первые браки, заключенные по договору, и фактические супружеские союзы, но и браки, заключенные на основе обычая, а также церковные браки. В отношении женщин, которые на момент переписи овдовели, жили с мужем раздельно или состояли в разводе, следует получить информацию о «дате/возрасте/количестве лет, прошедших с момента расторжения первого брака». Информация о дате расторжения первого брака (если такое имело место) позволяет получить необходимые данные для исчисления «продолжительности первого брака» на этапе обработки данных в качестве производного показателя. В странах, где сведения о продолжительности брака имеют более высокую надежность, чем сведения о возрасте, составление таблиц данных о родившихся живыми детях по признаку продолжительности брака позволяет получить более точные расчеты рождаемости, чем оценки на основе информации о детях, родившихся живыми, в их классификации по возрасту женщины. Данные о продолжительности брака можно получить путем вычитания возраста вступления в брак из возраста на момент переписи или не-

посредственно путем определения количества лет, прошедших с момента заключения брака.

334. Дату первого брака следует включать в перепись для всех лиц, когда-либо состоявших в браке (или только для женщин в соответствии с *Принципами и рекомендациями*). В программе должна проводиться проверка на соответствие. Для лиц, никогда не состоявших в браке, информацию сообщать не следует, но лица, состоявшие в браке, должны иметь достоверные число, месяц и год. Группам редактирования необходимо принять решение о том, должны ли быть достоверными число и месяц: страны, не использующие метод динамической интерполяции, могут квалифицировать число и месяц в графе «неизвестно»; страны, использующие динамическую интерполяцию, могут определять число и месяц, когда данные о них отсутствуют.

a) *Графа «семейное положение» для лиц, никогда не состоявших в браке, не должна заполняться*

335. Лица, которые никогда не состояли в браке, не должны сообщать о возрасте вступления в первый брак. Если для лиц, никогда не состоявших в браке, имеются достоверные данные, группа редактирования должна решить вопрос о том, изменить ли графу «семейное положение», или не заполнять графу «возраст» для этого лица. Если семейное положение не меняется, страны, использующие только графу «не указано», будут применять этот код. Страны, использующие метод динамической интерполяции должны, вероятно, использовать возраст и пол для получения соответствующего ответа о семейном положении.

b) *Для лиц, когда-либо состоявших в браке, следует отводить отдельную графу*

336. Что касается года первого брака, то страны, не использующие метод динамической интерполяции, могут применять графы «не указано» или «не известно». Страны, использующие метод динамической интерполяции, могут использовать другие переменные, например, такие, как возраст супруга(и) или разницу в возрасте между супругами, количество детей и детей, родившихся за последний год, для определения соответствующего года первого брака.

6. *Рождаемость: дети, родившиеся живыми (P4A) и выжившие дети (P4B)*

337. Термин «дети, родившиеся живыми» означает общее количество детей, родившихся живыми, из которого тем самым исключаются мертворожденные дети, выкидыши и аборт. Иногда демографы используют

выражение «живорожденные дети», но в настоящей работе будут использоваться термины «дети, родившиеся живыми» или «родившиеся дети».

338. Совокупность, в отношении которой следует собирать данные по каждому из вопросов, содержащихся в этом разделе, включает женщин в возрасте 15 лет (или другом допустимом минимальном возрасте) и старше, независимо от того, состоят они в браке или нет, или женщин конкретных подкатегорий, таких, как женщины, когда-либо состоявшие в браке. В странах, в которых не используют данные о женщинах в возрасте 50 лет и старше, все усилия следует сосредоточить на сборе данных только относительно женщин в возрасте от 15 до 50 лет, хотя при изучении последних данных о рождаемости, возможно, целесообразно уменьшить нижний возрастной предел на несколько лет (Организация Объединенных Наций, 1998 год, пункт 2.121).

a) *Собираемые элементы данных о рождаемости*

339. В издании *Принципы и рекомендации* (Организация Объединенных Наций, 1998 год) рекомендуется получать информацию об элементах рождаемости: детях, родившихся живыми, дате рождения последнего ребенка, родившегося живым, и возрасте матери при родах первого, родившегося живым ребенка. Ответы на вопрос о возрасте, дате вступления в брак или продолжительности брака могут существенно улучшить оценки рождаемости, основанные на показателе родившихся живыми детей. Кроме того, во многих странах по-прежнему собирается информация о детях, живущих на момент переписи, что является подспорьем особенно в ретроспективном анализе рождаемости.

340. В ходе переписей и обзоров информация о рождаемости собирается по всем женщинам с использованием установленного в каждой стране минимального возраста и иногда максимального возраста.

b) *Общие правила редактирования данных о рождаемости*

341. Цель редактирования данных о рождаемости состоит в том, чтобы согласовать вводимые данные между собой и с возрастом:

- a) общее количество детей, родившихся живыми, не может превышать возраст этого лица плюс определенный минимальный возраст, установленный в каждой стране. В целях редактирования см. раздел «Возраст при первом рождении» ниже для определения минимальной разницы в возрасте между матерью и старшим ребенком, родившимся живым;
- b) общее количество детей, родившихся живыми, не может превышать сумму детей, живущих в домаш-

нем хозяйстве, проживающих в других местах и умерших. Если общее количество превышает сумму этих слагаемых, то группы редактирования должны решить, какое из них имеет приоритет, и на этой основе может внести соответствующие коррективы;

- c) если данные собираются как по живущим, так и умершим детям, общее количество этих детей не может превышать количество детей, родившихся живыми;
- d) количество детей, родившихся живыми, не может быть меньше, чем данные о «детях, родившихся за последние 12 месяцев»;
- e) в зависимости от страны и фактического количества детей, родившихся живыми, и детей, живущих на момент переписи, можно использовать матрицу интерполяции для элемента данных о детях, родившихся за последние 12 месяцев, для распределения ответа по возрасту и детям, родившимся живыми. Однако необходимо проявлять большую осторожность при присвоении значения категории детей, родившихся за последние 12 месяцев, когда появляется пробел. Для многих стран пропуск этого элемента означает, что не родился ни один ребенок. Распределенные значения могут исказить данные;
- f) иногда страны собирают данные о детях, родившихся живыми, детях, живущих на момент переписи, и другие элементы данных о рождаемости в распределении по полу. В этих случаях отредактированные данные, представленные здесь, определяются в совокупности, но такие страны могут желать включать дополнительные проверки для учета дополнительной имеющейся информации. Такие дополнительные проверки включают точное определение того факта, что количество мальчиков, родившихся живыми, составляет сумму мальчиков, живущих на момент переписи, и умерших мальчиков, а количество девочек, родившихся живыми, является суммой девочек, живущих на момент переписи, и умерших девочек. Что касается редактируемых данных для детей, в отношении которых не проводится дифференциации по полу, необходимо предпринять соответствующие действия, когда эти суммы не равняются слагаемым.
- c) *Связь между детьми, родившимися живыми, и детьми, живущими на момент переписи*

342. Данные о детях, родившихся живыми, и детях, живущих на момент переписи, используются для опосредованных оценок как рождаемости, так и смертности. Результаты переписи или обследования организуются в разбивке по единому году или группе женщин с

пятилетним возрастным интервалом. С помощью различных алгоритмов получают постоянные или изменяющиеся оценки смертности. Однако, чтобы получить наилучшие результаты, группам редактирования необходимо проявлять осторожность при определении соответствующего вида редактирования имеющихся данных.

343. Частью проблемы, связанной с разработкой общей программы редактирования, является то, что различным странам требуются различные виды информации. Например, в различных странах собираются следующие комплекты данных:

- a) только дети, родившиеся живыми;
 - b) дети, родившиеся живыми; и дети, живущие на момент переписи (оба пола без дифференциации или отдельно по полам);
 - c) дети, родившиеся живыми; дети, живущие на момент переписи; и дети, которые умерли (как без дифференциации полов, так и в разбивке по полу);
 - d) дети, родившиеся живыми; дети, проживающие дома; дети, живущие отдельно; и дети, которые умерли (как без дифференциации полов, так и в разбивке по полу).
- d) *Редактирование, когда регистрируются только дети, родившиеся живыми*

344. Если страна не применяет метод динамической интерполяции, недостоверное или недостающее значение для категории «дети, родившиеся живыми», должно относиться к категории «не известно». В странах, где применяется метод динамической интерполяции, специалисты должны решить, желают ли они использовать динамическую интерполяцию для всех элементов данных. Если специалисты применяют этот метод, значение для детей, родившихся живыми, можно получить на основе одного и того же годового интервала для женщины и по меньшей мере на основе одной другой характеристики. Можно также использовать одномерную матрицу для одного и того же возрастного интервала только для матери. Другими характеристиками могут быть такие элементы, как достигнутый уровень образования или религия, поскольку известно, что во

многих странах рождаемость зависит от уровня образования или вероисповедания.

- e) *Редактирование, когда регистрируются дети, родившиеся живыми, и дети, живущие на момент переписи*

345. Если приводятся ответы как на вопрос о «детях, родившихся живыми», так и о «детях, живущих в момент переписи», в программе необходимо определить следующее:

- a) являются ли эти элементы внутренне сопоставимыми, равняется ли количество детей, родившихся живыми, количеству детей, живущих на момент переписи, или оно превышает это количество;
- b) согласуется ли каждый элемент с возрастом женщины;
- c) согласуется ли показатель «дети, родившиеся живыми» с показателем «дети, родившиеся в последний год» (или последнее деторождение), если такие данные собираются.

346. Демографы используют элементы данных о детях, родившихся живыми, и детях, живущих в момент переписи, для получения косвенных оценок смертности. В связи с этим при редактировании данных необходимо сохранять зависимость между двумя этими элементами. Иногда регистрируется только один из этих двух элементов, а остальные остаются неизвестными. При простом редактировании можно было бы допустить, что у детей, родившихся живыми, отсутствует смертность и указывать оба элемента как одинаковые. Однако при указании этих двух элементов в качестве одинаковых косвенная оценка смертности не учитывает младенцев, тем самым занижая показатель смертности и завышая показатель продолжительности жизни. Если таких случаев регистрируется в переписи или обследовании немного, это не приносит большого вреда. Однако, если это происходит с определенной частотностью, как можно было бы ожидать в тех странах, в которых используется косвенный метод, последствия могли бы быть значительными. Соответствующий пример приводится на рисунке 27.

Рисунок 27. Иллюстрация домохозяйства с информацией о рождаемости

Лицо	Степень родства	Пол	Возраст	Дети, родившиеся живыми	Дети, живущие на момент переписи
1	Глава домашнего хозяйства	1	60		
2	Супруг(а)	2	60	5	99
3	Дочь	2	40	3	3
4	Внучка	2	20	1	1
5	Внучка	2	18	0	0
6	Внучка	2	1		

ПРИМЕЧАНИЕ: 99 = недостающие или недостоверные данные.

347. В данном случае супруга сообщает о 5 детях, родившихся живыми, но по какой-то причине количество детей, живущих на момент переписи, не было зарегистрировано. Респондент или счетчик не сообщили об этой величине или позже оператор, введивший эти данные, ошибся с набором на клавиатуре. Многие страны разрабатывают программу-редактор, в которой значение «5» присваивается детям, живущим в момент переписи, исходя из количества детей, родившихся живыми. Однако при этом данные искажаются.

348. По сути дела, нет необходимости вообще менять это значение. Страны, которые не применяют метод динамической интерполяции, возможно, предпочтут не трогать «неизвестное» значение. Естественно, такое решение также приводит к искажениям, поскольку при таком редактировании принимается решение, что «неизвестные» и «известные» ответы имеют одинаковое распределение для таблиц. Если стране требуются данные о детях, родившихся живыми, и о детях, живущих на момент переписи, для определения косвенных оценок смертности, то, вероятно, такая страна сталкивается также с проблемами регистрации данных. Сохранение неизвестных значений в массиве данных также должно исказить результаты окончательного анализа. Женщин, для которых неизвестны ни количество детей, родившихся живыми, ни детей, живущих на момент переписи, нельзя учитывать при определении оценок смертности, поскольку невозможно определить разницу между детьми, родившимися живыми, и детьми, живущими на момент переписи.

349. В странах, в которых применяется метод динамической интерполяции, необходимо рассмотреть вопрос об определении недостающих данных, основываясь как минимум на других признаках рождаемости и возрасте женщин. Матрицы интерполяции могут обновляться, когда имеется достоверная информация о возрасте женщины, количестве детей, родившихся живыми, и детей, живущих на момент переписи, и могут применяться, когда недостает такого элемента. Когда

недостает информации о детях, родившихся живыми, в матрице интерполяции будут иметься возраст женщины и количество детей, живущих на момент переписи. Когда недостает информации о детях, живущих на момент переписи, в матрице интерполяции будет приводиться возраст женщины и количество детей, родившихся живыми.

350. Кроме того, при разработке матриц интерполяции важно помнить, что количество детей, родившихся живыми, и количество детей, живущих на момент переписи, должны соответствовать разнице в возрасте между матерью и старшим ребенком (если имеется такая информация) и общим количеством детей, родившихся живыми, при конкретном возрасте матери.

351. Например, разница между интерполированным количеством детей, родившихся живыми, и возрастом матери может быть, как минимум, 12 лет. В этом случае в матрице интерполяции, использующей группы женщин с возрастным интервалом в 5 лет, в ряде случаев будут интерполироваться несовместимые данные.

352. На рисунке 28 в сопутствующей матрице интерполяции в верхней части по горизонтали приводятся значения возраста женщин, а в левой части рисунка (по вертикали) приводится количество детей, родившихся живыми. Введенные данные представляют собой интерполированные значения для детей, живущих на момент переписи. В одних случаях ответы будут надлежащими, а в других — нет. Если в программе встречается женщина 19 лет с 5 детьми, родившимися живыми, то значение в 5 детей, живущих на момент переписи, вероятно, должно превышать критерий разницы в возрасте (разница в возрасте составляет 15 лет, исходя из количества детей, живущих на момент переписи, и зарегистрированного возраста). Однако для 15-летней женщины ни 5 детей, родившихся живыми (разница в возрасте составляет 10 лет), ни 4 ребенка, доживших до момента переписи (разница в возрасте будет 11 лет), вряд ли будут приемлемыми.

Рисунок 28. Первичные значения для определения детей, живущих на момент переписи, когда возраст и количество детей, родившихся живыми, достоверны

Дети, родившиеся живыми	Возраст												
	15	16	17	18	19	20	21	22	23	24	25–29	30–34	35+
0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	0	0	0
2		2	2	2	2	2	2	2	2	2	1	1	1
3			3	3	3	3	3	3	3	3	2	2	2
4					4	4	4	4	4	4	3	3	3
5					5	5	5	5	5	5	4	4	4

353. Более точной матрицей интерполяции будет матрица, в которой к молодым женщинам применяется единый возрастной интервал. Тогда в матрицу интерполяции будут вноситься лишь достоверные ответы о разнице в возрасте для конкретной возрастной группы, и только достоверные ответы можно получать на основе матрицы интерполяции.

f) *Редактирование, когда есть данные о детях, родившихся живыми, детях, живущих на момент переписи, и умерших детях*

354. Категория «дети, родившиеся живыми» представляет собой сумму показателей «дети, живущие на момент переписи» и «умершие дети». Ниже приводятся методы, иллюстрирующие, как можно ликвидировать любое несоответствие.

i) *Когда ответы имеются по всем трем элементам данных*

355. Если есть информация обо всех трех элементах, то в программе необходимо определить:

- a) являются ли все эти три элемента внутренне сопоставимыми, и является ли количество детей, родившихся живыми, суммой показателя детей, живущих на момент переписи, и умерших детей;
- b) соответствуют ли все три элемента возрасту женщины;
- c) согласуется ли количество детей, родившихся живыми, с количеством детей, родившихся в последний год (или последним деторождением), если такие данные собираются.

356. Если все эти элементы согласуются, то редактирование закончено. Однако, если они не согласуются,

редактирование должно решить эту проблему. Три элемента могут быть внутренне несопоставимыми: например, у женщины может быть 5 детей, родившихся живыми, но в момент переписи могут жить только два ребенка, а два — умереть. Группе редактирования следует решить, какая из переменных имеет приоритет по сравнению с другими. Во многих случаях женщина наверняка помнит всех своих детей, которых она родила, хотя она может забыть количество своих умерших детей. Тогда группа редактирования, возможно, предпочтет принять количество детей, родившихся живыми, и количество детей, живущих на момент переписи, и провести вычитание, чтобы получить новое, согласующееся значение умерших детей.

ii) *Когда сообщаются два элемента*

357. Поскольку категория детей, родившихся живыми (CEB), представляет собой сумму количества детей, живущих на момент переписи (CS), и умерших детей (CD), если имеются два элемента из трех элементов информации, то компьютерная программа может определить третью переменную:

- если известны CEB и CS, то $CD = CEB - CS$.
- если известны CS и CD, то $CEB = CS + CD$.
- если известны CEB и CD, то $CS = CEB - CD$.

Такие тесты обычно проводятся первыми. Как только программа определит, что все три элемента информации являются достоверными и согласующимися, то редактирование закончено.

iii) *Когда сообщается только один элемент данных*

358. Когда известен только один из трех этих элементов и если страна не применяет метод динамической интерполяции, два остальных элемента данных следует относить к категории «не известно». Если же

страна применяет метод динамической интерполяции, группе редактирования необходимо определить метод получения по меньшей мере еще одного элемента, а третий элемент следует затем определять путем вычитания или сложения. Двухмерную матрицу можно использовать для определения второго значения рождаемости, основанного на первом элементе или едином возрасте для женщин. Если, например, известно значение детей, родившихся живыми, количество детей, живущих на момент переписи, можно получить по матрице интерполяции, как описывалось выше, а затем путем вычитания получить количество умерших детей. Аналогичным образом, если известно значение детей, живущих на момент переписи, то количество детей, родившихся живыми, получают по матрице интерполяции единого возраста женщины и детей, живущих на момент переписи, а количество умерших детей получают путем вычитания.

iv) *Когда не сообщается ни один из этих элементов*

359. Когда нет данных ни об одном из этих трех признаков, группа редактирования должна принять решение о том, как поступать дальше. Если в стране не применяется метод динамической интерполяции, все элементы следует относить к категории «не известно», и их не следует применять в косвенных методах определения смертности или рождаемости. В странах, в которых применяется метод динамической интерполяции, специалисты должны решить, хотят ли они использовать метод динамической интерполяции для всех элементов данных.

360. Если специалисты примут решение применять метод динамической интерполяции, то показатель детей, родившихся живыми, можно получить на основе единого возраста женщин и по меньшей мере на основе одной из других характеристик. Можно также применять одномерную матрицу для единого возраста только матери. К другим характеристикам могут относиться такие элементы, как достигнутый образовательный уровень или вероисповедание.

361. Как только определяется первый элемент, можно предпринять упомянутые выше шаги по редактированию, когда сообщается только один элемент для получения второй характеристики рождаемости. Затем по первым двум элементам можно определить третий элемент. Все три элемента должны быть совместимыми, потому что матрицы интерполяции должны обновляться только тогда, когда все три элемента сопоставимы. Полученная рождаемость должна также быть сопоставимой с другими женщинами в географическом районе, поскольку информация, полученная от этих женщин, используется для обновления матрицы интерполяции.

g) *Интерполирование, когда есть данные о детях, родившихся живыми, детях, живущих дома, детях, живущих вне дома, и умерших детях*

i) *Когда имеются все четыре элемента*

362. Если имеются все четыре элемента данных, в программе необходимо установить:

- a) являются ли все четыре элемента внутренне сопоставимыми для того, чтобы количество детей, родившихся живыми, представляло собой сумму детей, живущих дома, детей, живущих вне дома, и умерших детей;
- b) согласуется ли каждый из этих четырех элементов с возрастом женщины;
- c) согласуются ли категории «дети, родившиеся живыми», с категорией «дети, родившиеся в последний год» (или последним деторождением), если такая информация собирается.

363. Если все эти элементы согласуются, редактирование закончено. Однако, если они не согласуются, при редактировании необходимо устранить эти несоответствия. Как и в случае с тремя элементами, описанным выше, все четыре элемента данных могут быть внутренне несопоставимыми. Опять-таки, группе редактирования следует принять решение, какая переменная имеет приоритет перед всеми остальными. Во многих случаях женщина, отвечающая на вопрос, наверняка помнит всех своих детей, которых она родила, хотя она может и забыть о тех из них, кто покинул дом, или точном количестве умерших детей. Тогда группа редактировать может предпочесть остановиться на количестве детей, родившихся живыми, и детей, живущих на момент переписи (суммой детей, живущих вне дома, и детей, живущих дома), и произвести вычитание, чтобы получить новые, согласующиеся значения для других переменных. Группе редактирования, возможно, придется разработать алгоритмы для различных комбинаций явлений.

ii) *Когда сообщаются три из этих четырех элементов*

364. Категория «дети, родившиеся живыми» (СЕВ) представляет собой сумму детей, живущих дома (СЛН), детей, живущих вне дома (СЛА), и умерших детей (СД). Если имеются любые три элемента из этих четырех элементов, то компьютерная программа может определить эти четыре переменные:

если известны СЕВ, СЛН и СЛА,
то $СД = СЕВ - СЛН - СЛА$.

если известны СЛН, СЛА и СД,
то $СЕВ = СЛН + СЛА + СД$.

если известны СЕВ, СЛН и СД,
то $СЛА = СЕВ - СЛН - СД$.

если известны СЕВ, СЛА и СД,
то $СЛН = СЕВ - СЛА - СД$.

iii) *Когда сообщаются два элемента из этих четырех элементов*

365. Если известны только два из этих элементов, то группе редактирования необходимо решить, что делать дальше. Как уже отмечалось выше, во многих странах женщины не сообщают о точном количестве умерших детей. Другой элемент, который, как правило, опускается, представляет собой информацию о детях, живущих вне жилищной единицы, данные о которых также обычно невозможно получить прямым методом. Поэтому следует проявлять особую осторожность при разработке вопросника, при регистрации и при обработке данных с целью получить наиболее качественные данные по всем элементам рождаемости.

366. Данные о детях, живущих в домашнем хозяйстве (СЛН), можно получить путем суммирования количества детей в жилищной единице. Если только одна женщина в единице имеет соответствующий признак родства, простой подсчет должен дать количество детей, живущих в домашнем хозяйстве. Если в домашнем хозяйстве этими признаками родства будет обладать больше женщин, то можно продолжать пользоваться программой-редактором при допущении, что данные о детях будут следовать немедленно за данными о матери в процессе сбора данных. Когда ничто другое не дает результатов, страны, в которых применяется метод динамической интерполяции, могли бы определять количество детей, живущих в жилищной единице, по возрасту матери и одной из других известных переменных. (См. ниже общие правила для определения отдельных элементов рождаемости по другим признакам и возрасту матери). Важно использовать единый возраст женщины, когда это возможно, а также единое количество детей, родившихся живыми, живущих в жилищной единице, живущих вне дома или умерших.

367. Например, данные о детях, родившихся живыми, и умерших детях могут быть достоверными, а данные о детях, живущих в домашнем хозяйстве, и детях, живущих вне дома, могут быть недостоверными. В этом случае количество детей, живущих дома, можно определить, суммируя количество детей, имеющих соответствующее родство с матерью (при допущении, что мать является главой домашнего хозяйства). Тогда в наличии будет три из четырех этих элементов, а чет-

вертый элемент, дети, живущие вне дома, можно определить путем вычитания: $СЛА = СЕВ - СЛН - СД$.

368. Однако когда известны только два элемента, скорее всего придется регистрировать детей, родившихся живыми, и детей, живущих дома. Женщины обычно охотно сообщают о детях, родившихся живыми, а информацию о детях, живущих дома, можно обычно получить путем наблюдения или работы с респондентами при регистрации, но такое решение не годится для детей, живущих вне дома, или умерших детей. Тогда при редактировании можно применить матрицу интерполяции, содержащую возраст женщины, и показатель детей, родившихся живыми (СЕВ), или, еще лучше, возраст женщины, детей, родившихся живыми (СЕВ), и детей, живущих дома (СЛН). С помощью этих переменных будет получена информация от той же женщины с теми же характеристиками для детей, живущих вне дома (СЛА).

369. В странах, в которых применяется только двухмерная матрица для возраста женщины и детей, родившихся живыми (СЕВ), не включающая третий элемент, есть риск, что дети, живущие дома (СЛН), получат значение для детей, живущих вне дома (СЛА), которое несовместимо с двумя другими. Например, если возраст женщины составляет 25 лет, а число СЕВ составляет 5, по матрице интерполяции можно получить значение 3 для детей, живущих вне дома. Если значение для детей, живущих дома, составляет 2, то редактирование данных проводится легко. Значение для умерших детей должно составлять 0, а элементы рождаемости должны быть $СЕВ = 5$, $СЛН = 2$, $СЛА = 3$, $СД = 0$.

370. Однако значение для детей, живущих дома, в действительности может составлять 4, причем для определения значения детей, живущих вне дома, используется только возраст женщины и количество детей, родившихся живыми. Цифра 3 для детей, живущих вне дома, приведет к несовместимости этих элементов. Значение для детей, родившихся живыми (5), будет меньше суммы живущих детей (4 дома и 3 вне дома), а общий результат будет равняться 7. Поэтому следует применять трехмерную матрицу: для 5 СЕВ и 4 СЛН, а значение в матрице интерполяции может составлять 1 для детей, живущих вне дома (нулевое значение следует определять путем вычитания количества умерших детей). В матрице интерполяции для детей, живущих вне дома, значение должно равняться 0 (а значение 1 должно определяться путем вычитания количества умерших детей). Для других пар известных элементов данных следует составлять аналогичные матрицы интерполирования, как это показано на рисунке 29.

Рисунок 29. Образец матрицы интерполяции, которую необходимо разработать для пар известных элементов данных

Если известны элементы		Применяйте метод динамической интерполяции для одного из этих элементов (и затем производите вычитание или сложение)	
Дети, родившиеся живыми	Дети, живущие дома	Дети, живущие вне дома	Умершие дети
Дети, родившиеся живыми	Дети, живущие вне дома	Дети, живущие дома	Умершие дети
Дети, родившиеся живыми	Умершие дети	Дети, живущие дома	Дети, живущие вне дома
Дети, живущие дома	Дети, живущие вне дома	Дети, родившиеся живыми	Умершие дети
Дети, живущие дома	Умершие дети	Дети, родившиеся живыми	Дети, живущие вне дома

371. В каждом случае имеются два из четырех этих элементов. Третий элемент определяется методом динамической интерполяции, а четвертый элемент — путем сложения или вычитания. Группы редактирования должны решить, какой оптимальный путь избрать, исходя из культурных традиций.

iv) *Когда сообщается только один элемент данных*

372. Когда из четырех элементов известен только один, ситуация осложняется. Странам необходимо решить, как им действовать, когда недостает такого значительного объема информации. Если применяется метод динамической интерполяции, в первой матрице интерполяции, как отмечалось выше, используется такой элемент, как один возраст женщины и один известный элемент для построения двумерной матрицы для интерполяции любого из остальных элементов. Как только определяются два элемента, два других остаются неизвестными по определению. Отсюда, дальнейшее применение динамической интерполяции для третьего элемента не должно приводить к несовместимости с остальными признаками, поскольку они неизвестны. Схема, обсужденная выше, для двух известных элементов и двух неизвестных элементов применяется для получения третьего элемента. Тогда четвертый элемент определяется путем вычитания. Все четыре элемента должны быть совместимыми. Для поиска возможного решения, когда известен только один элемент, см. пункт 376.

v) *Когда не сообщается ни один элемент*

373. Когда нет ни одного из этих четырех элементов, группа редактирования должна решить, что делать, если неизвестен ни один из элементов. Если в этой стране не применяется метод динамической интерполяции, все элементы должны относиться к категории «не известно» и не должны использоваться в косвенных методах определения смертности или рождаемости. В странах, в которых применяется метод динамической интерполяции, специалисты должны решать, хотят ли они применять интерполяцию для других элементов. Для поиска

другого решения, когда неизвестен ни один из элементов, см. пункт 376.

374. Если специалисты принимают решение применять метод динамической интерполяции, значение для детей, родившихся живыми, можно получить на основе одного и того же возраста женщины и по меньшей мере на основе одной из характеристик. Можно также применять одномерную матрицу для единого возраста только матери. К другим характеристикам могут относиться такие элементы, как образовательная подготовка или вероисповедание, поскольку известно, что во многих странах дифференциация по рождаемости существует в зависимости от образовательной подготовки или вероисповедания.

375. Как только определен первый элемент, использованный выше метод, когда известен только один элемент, можно использовать для получения второго элемента рождаемости. Тогда третий элемент можно получить по двум первым элементам, а четвертый элемент можно определить путем вычитания. Все эти четыре элемента должны быть совместимыми, потому что матрицы интерполяции необходимо обновлять только тогда, когда все элементы совместимы. Полученный показатель рождаемости также должен быть совместимым с другими категориями женщин в географическом районе, поскольку информация, полученная от этих женщин, используется для обновления матрицы интерполяции.

h) *Значение единого источника данных для всех признаков рождаемости*

376. Очень важно, и если вообще возможно, исчислять все признаки по одной женщине, когда неизвестен ни один показатель. Чтобы убедиться, что все данные берутся по одной и той же женщине-источнику, возможно, потребуется разработать матрицы интерполяции, в которых применяются все данные о рождаемости. В этом случае матрицы интерполяции можно было бы обновлять лишь тогда, когда в программе-редакторе установлено, что все элементы рождаемости согласуются.

7. Рождаемость:

дата рождения последнего ребенка (P4C)

377. Информация о дате рождения (число, месяц и год) последнего ребенка, родившего живым, и о поле ребенка используется для получения расчетных данных о текущей рождаемости. Впоследствии, на этапе обработки данных, можно получить число детей, родившихся живыми, в течение 12 месяцев, непосредственно предшествующих дате переписи, в виде оценки числа живорождений за последние 12 месяцев. Этот подход обеспечивает получение более точных данных, чем информация о числе рождений, приходившихся на женщину в течение 12 месяцев, непосредственно предшествующих переписи, для получения примерного значения текущих повозрастных коэффициентов рождаемости и других показателей рождаемости (*Принципы и рекомендации в отношении переписи населения и жилого фонда*, Организация Объединенных Наций, 1998 год, пункт 2.134).

378. Следует отметить, что информация о дате рождения последнего ребенка, родившегося живым, не позволяет получить данные об общем числе детей, родившихся живыми в течение 12-месячного периода. Даже если не было допущено ошибок при сообщении данных о последнем живорожденном ребенке, эти сведения позволяют установить число женщин, которые родили, по крайней мере, одного живого ребенка в течение 12 месяцев, а не число рождений, поскольку незначительная доля женщин родит более одного ребенка за один год (*Принципы и рекомендации в отношении переписи населения и жилого фонда*, Организация Объединенных Наций, 1998 год, пункт 2.135).

379. Информация должна собираться только в отношении женщин в возрасте от 15 до 50 лет, которые сообщили о рождении хотя бы одного живого ребенка в течение всей своей жизни. Кроме того, информация должна собираться в отношении женщин всех категорий семейного положения, от которых собираются сведения о рожденных ими детях с разбивкой по полу. Если данные о родившихся живыми детях собираются лишь по выборке женщин, то информация о текущей рождаемости также должна собираться по той же выборке. (*Принципы и рекомендации в отношении переписи населения и жилого фонда*, Организация Объединенных Наций, 1998 год, пункт 2.136).

380. В программу редактирования следует включать следующий вид редактирования. Дата рождения последнего ребенка по всем женщинам должна помещаться между установленным в стране минимальным возрастом и максимальным возрастом. Программа должна обеспечивать проверку на соответствие. Например, следует исключать все данные по мужчинам и женщинам, не включенным в отобранную возрастную группу.

Аналогичным образом, в отношении женщин в отобранной возрастной группе, имеющих порядок деторождений больше 0, следует указывать достоверное число, месяц и год последнего деторождения. Группе редактирования необходимо решить, должны ли быть достоверными число и месяц рождения: страны, не применяющие метод динамической интерполяции, относят число и месяц к категории «не известно». Группы редактирования, применяющие динамическую интерполяцию, могут исчислять число и месяц, когда они отсутствуют.

381. Если информация о годе рождения последнего ребенка отсутствует или является недостоверной, страны, применяющие метод динамической интерполяции, могут относить их к категории «не указано» или «не известно». В странах, применяющих метод динамической интерполяции, для получения даты рождения последнего ребенка можно применять другие переменные, например, такие, как возраст и количество детей, родившихся живыми.

8. Рождаемость: возраст при рождении первого ребенка (P4G)

382. Возраст матери при рождении первого живорожденного ребенка используется для косвенной оценки рождаемости, основанной на данных о первых родах, и в целях получения информации о наступлении детородного возраста. Если этот вопрос включается в перепись, то информацию следует получить по каждой женщине, которая родила живым хотя бы одного ребенка (Организация Объединенных Наций, 1998 год, пункт 2.143).

383. Возраст при рождении первого живорожденного ребенка определяется либо прямым методом с помощью точно определенного элемента «возраст при рождении первого живорожденного ребенка», либо по разнице в возрасте между текущим возрастом матери и возрастом старшего ребенка, если известен возраст старшего ребенка. Минимальный возраст детей, установленный в странах, не является минимальным биологическим возрастом. Если, например, минимальный допустимый возраст при рождении первого ребенка составляет 13 лет, то респондент может сообщить, а счетчик может зарегистрировать возраст какой-либо женщины при рождении, составляющий 11 или 12 лет. Тогда группам редактирования следует решить, менять ли минимальный допустимый возраст, исключить данные о деторождении или изменить либо возраст матери, либо возраст при первом рождении (применяя либо возраст ребенка, либо ее возраст в зависимости от переменных, используемых для определения разницы в возрасте). Аналогичным образом группы редактирования должны решить, является «пределный возраст» максимальным возрастом при рождении первого ребенка. Хотя жен-

щины могут рожать детей в возрасте свыше 50 лет, это событие случается не очень часто, и чтобы исправить ошибки, при редактировании необходимо установить, какие случаи являются реальными.

384. Важно помнить, что минимальный или максимальный возраст при рождении первого ребенка (разница в возрасте между матерью и ее старшим ребенком, проживающим в домашнем хозяйстве) должны соответствовать обычаям и традициям страны. Специалисты должны решить, когда какое-либо значение является «шумом», а не законным возрастом при рождении первого ребенка. Когда устанавливаются соответствующие правила, то специалисты должны определить, как решить эту проблему. Если не применяется метод динамической интерполяции, программа должна использовать категорию «не известно». Когда применяется метод динамической интерполяции, она может устанавливать возраст при рождении первого ребенка на основе данных о других женщинах того же возраста, аналогичного количества детей, родившихся живыми. Специалисты, определяющие матрицу интерполяции, возможно, пожелают учитывать такие факторы, как проживание в городских/сельских районах (если рождаемость отличается в этих двух районах), участие женщины в рабочей силе, хотя текущий статус в рабочей силе не всегда соответствует ее статусу при рождении первого ребенка, достигнутый уровень образования.

9. Смертность (P4D)

385. Информация о числе смертей за последние 12 месяцев используется для оценки уровня и структуры смертности с разбивкой по полу и возрасту в тех странах, где не представляется возможным непрерывно получать удовлетворительные статистические данные о смертности на основании записи актов гражданского положения. Для обеспечения надежности расчетных данных по этому вопросу необходимо получить как можно более полную достоверную информацию о числе смертей за последние 12 месяцев с разбивкой по полу и возрасту. Широкая практика последних десятилетий включения вопросов о смертности в переписные листы значительно повысила полезность применения косвенных методов для оценки в расчетах смертности взрослого населения (Организация Объединенных Наций, 1998 год, пункт 2.137).

386. Смертность по каждому домохозяйству лучше всего было бы определять в виде общего числа смертей за последние 12 месяцев, предшествующих дате переписи. В случае, когда невозможно получить информацию о числе смертей за последние 12 месяцев, рекомендуется по крайней мере собрать данные о смертности детей в возрасте до одного года. В отношении каждого на-

званного умершего лица должны также собираться данные об его имени, возрасте, поле, дате (число, месяц, год) его смерти. Следует принять меры к тому, чтобы четко объяснить опрашиваемому лицу, о каком именно учетном периоде времени идет речь, во избежание ошибок, которые могут возникнуть вследствие недопонимания этого момента. Например, учетный период можно точно определить, ориентируясь на какие-либо праздничные или исторические даты, отмечаемые в конкретной стране (*Принципы и рекомендации в отношении переписи населения и жилого фонда*, Организация Объединенных Наций, 1998 год, пункт 2.138).

387. В *Принципах и рекомендациях* (Организация Объединенных Наций, 1998 год) предлагается собирать данные об имени, возрасте и поле, а также числе, месяце и годе смерти лиц, которые умерли в год, предшествующий переписи. В странах, в которых не применяется метод динамической интерполяции, можно присваивать категорию «не известно» каждой из этих переменных, если они недостоверны. Для стран, где применяется метод динамической интерполяции, можно использовать возраст (в возрастных группах), пол и год смерти в качестве значений матриц интерполяции для других переменных. Фактические матрицы интерполяции, вероятно, относятся к какой-либо конкретной стране, и членам группы редактирования придется работать совместно для разработки соответствующих матриц интерполяции.

10. Сиротство в случае смерти матери/отца (P4E) и номер строки матери

388. Для сбора информации о сиротстве следует задавать два прямых вопроса: а) жива ли родная мать лица, зарегистрированного в домохозяйстве, на дату переписи; и б) жив ли родной отец лица, зарегистрированного в домохозяйстве, на дату переписи, независимо от того, зарегистрированы или нет его отец и мать в том же домохозяйстве. Это исследование должно дать информацию о биологических родителях. Следовательно, необходимо принять меры к тому, чтобы исключить приемных и патронатных родителей. Поскольку в живых, как правило, остается более одного ребенка, который может сообщить о своем сиротстве, необходимо составить вопросы так, чтобы избежать двойного учета родителей при ответах их детей. В этих целях следует задать два дополнительных вопроса, а именно: с) является ли респондент самым старшим из выживших детей его или ее матери; и d) является ли респондент самым старшим из выживших детей его или ее отца. Таблицы должны строиться с учетом только самого старшего выжившего ребенка (Организация Объединенных Наций, 1998 год, пункт 2.140).

389. Такие элементы, как «мать, живущая на момент переписи» и «номер строки матери» взаимосвязаны, и их следует редактировать вместе. В отношении тех, кто указывает в графе «мать, живущая на момент переписи» иные ответы кроме «да», необходимо проверить на достоверность номера строки матери; если запись сделана правильно, то матери, живущей на момент переписи, следует присвоить код для «да». В отношении тех, кто в графе «мать, живущая на момент переписи» указывает иное, чем «да», номер строки матери следует проверять, для того чтобы посмотреть, равен ли он 00 или номеру строки женщины в возрасте 12 лет или старше. И если один из этих случаев подходит, то программа исходит из того, что у данного лица есть мать или демографическому статусу матери присваивает «да». Если номер строки матери записан неправильно, и матери, живущей на момент переписи, присвоен код «нет» или «не знает», то запись номера строки матери следует устранить. Во всех других случаях матери, живущей на момент переписи, следует присваивать код «не знает», а любую запись в графе «номер строки» следует устранять.

390. Страна может предпочесть не редактировать номер строки матери в отношении тех лиц, которые указали «нет» или «не знает», если мать жива на момент переписи. Во всех других случаях номер строки можно проверить на согласованность или исчислить, используя такие показатели, как степень родства, номер строки, пол и степень родства и возраст того лица, кто указан как мать. Когда есть несоответствие или мать нельзя определить, то можно присвоить код «живет в другом месте».

В. ХАРАКТЕРИСТИКИ МИГРАЦИИ

391. В результате естественного прироста (рождаемость и смертность) и чистой миграции демографическая ситуация в стране со временем меняется. Миграция может быть долгосрочной (с рождения) или краткосрочной, которая определяется по предыдущему месту жительства и продолжительности проживания или по предыдущему конкретному моменту времени. Поскольку эти элементы часто взаимосвязаны, то некоторым странам, возможно, целесообразнее проводить совместное редактирование этих данных аналогично редактированию основных демографических переменных. Если используется нисходящий принцип, то порядок редактирования приобретает особое значение, поскольку некоторые элементы необходимо редактировать прежде других.

392. Элементы миграции зачастую требуют более подробных кодов по сравнению с другими данными, поскольку в целях планирования и осуществления по-

литики могут потребоваться более мелкие территориальные единицы. Подробная информация о небольших районах может потребоваться для планирования кадров для новой школы или больницы. Кроме того, могут потребоваться различные схемы кодирования и различные приемы редактирования для мест в стране и за ее пределами.

393. Данные о стране рождения и годах, прожитых в данном районе, следует проверять на согласованность между собой, поскольку между этими двумя элементами существуют очевидные взаимосвязи. Кроме того, существуют обоснованные взаимосвязи между ответами разных членов домохозяйства. Например, если нет ответа на вопрос о количестве лет, прожитых ребенком в данном районе, его можно вычислить по ответу матери, а программа редактирования проверит, чтобы исчисленное значение не превышало возраст ребенка.

1. Место рождения (P1C)

394. Под местом рождения в первую очередь имеется в виду страна, в которой лицо родилось. Следует отметить, что страна рождения не обязательно должна быть связана с гражданством (см. пункты 405-409). Для лиц, рожденных в стране проведения переписи (местных уроженцев), понятие о месте рождения включает также определенный вид территориальной единицы страны, в которой жила мать данного лица к моменту его рождения. Однако в некоторых странах место рождения уроженцев определяется как территориальная единица, в которой фактически родилось лицо. Каждая страна должна уточнить, каким определением она пользовалась при проведении переписи (Организация Объединенных Наций, 1998 год, пункт 2.29).

а) *Взаимосвязь между данными о стране рождения и годами, прожитыми в данном районе*

395. Вводимые данные о месте рождения (P1C) и продолжительности проживания (P1D) можно проверить на согласованность, поскольку между этими двумя элементами существует тесная взаимосвязь. Кроме того, взаимосвязи существуют между разными членами домохозяйства, и можно делать допущения, исходя из данных о других членах семьи относительно того, мигрировало ли данное лицо или нет.

б) *Присвоение признака «не известно» для недостоверных записей о месте рождения*

396. Если страна не собирается применять метод динамической интерполяции, любые недостоверные ответы о месте рождения должны обозначаться как «не извест-

но». Обычно страна не должна редактировать несогласующиеся ответы о членах семьи или географических районах, если кодирование произведено правильно.

c) *Применение метода статической интерполяции для места рождения*

397. Запись о стране рождения следует менять, только если она выходит за определенные пределы. Если для лет, прожитых в районе, присвоен код «всегда», то следует присваивать код для страны, и если указано иное, чем «всегда», можно использовать информацию о предыдущем лице. Например, если предыдущее лицо — это мать, то количество лет, в течение которых мать жила в данном районе, можно сопоставить с возрастом этого индивида. Если же значение, касающееся матери, больше или равно возрасту данного индивида, то следует присвоить код «эта страна»; в противном случае это следует относить к признаку «страна рождения матери». Если страну рождения нельзя присвоить, исходя из данных о матери, аналогичным образом можно использовать данные о других, имеющих к этому отношении лицах. Если после этих проверок запись классифицировать нельзя, то в графе «страна рождения» можно указать «не известна».

d) *Применение метода динамической интерполяции для данных о месте рождения*

398. Как и до этого, запись о стране рождения следует менять, только если она выходит за определенные пределы. Если в графе «годы, прожитые в районе» указано «всегда», то для признака «эта страна» следует присвоить код «страна рождения». Если запись иная, чем «всегда», то следует изучить информацию, полученную от других членов домохозяйства, чтобы найти признаки страны рождения данного индивида.

e) *Присвоение места рождения, если у данного лица есть мать*

399. Если в графе «страна рождения» пробел или запись недостоверна, а в графе «продолжительность проживания» указано не «всегда», а что-то иное, то можно произвести поиск данных о матери этого лица. Если в домохозяйстве есть мать, то изучается запись о продолжительности проживания матери. Если в графе «годы, прожитые в данном месте» для матери указано «всегда», то стране рождения данного лица можно присвоить признак «эта страна». Если мать этого лица не всегда жила в данном месте, а его возраст меньше или равен количеству лет, прожитых матерью в этом районе, программа может также присвоить стране рождения

признак «эта страна». Если возраст данного лица превышает количество лет, которые мать прожила в данном районе, а данные о стране рождения матери достоверны, то страна рождения этого лица классифицируется так же, как и страна рождения матери.

f) *Присвоение места рождения ребенка главы домохозяйства*

400. Если мать в домохозяйстве не указана, а это лицо является сыном или дочерью главы домохозяйства, то для того, чтобы получить данные о месте рождения можно произвести несколько проверок, используя информацию из записи о главе домохозяйства. Если в записи о главе домохозяйства в графе «годы, прожитые в данном районе» указано «всегда», то программа должна присвоить признак «эта страна» в качестве страны рождения в записи об этом индивиде. Если в графе «годы, прожитые в данном районе» для главы домохозяйства указано не «всегда», а возраст этого лица меньше или равен количеству лет, которые глава домохозяйства провел в этом районе, то программа также должна присвоить признак «эта страна» в качестве страны рождения данного лица. Однако если возраст данного лица превышает количество лет, которые глава домохозяйства провел в этом районе, программа должна присвоить признак страны рождения главы домохозяйства, если у него указан правильный код страны рождения.

g) *Присвоение места рождения ребенка, не являющегося ребенком главы домохозяйства*

401. В зависимости от того, старше или младше то или иное лицо определенного возраста (возраста X), установленного группой редактирования, могут производиться совершенно разные исчисления. Если возраст младше значения X, страну рождения необходимо исчислять, исходя из первой предшествующей записи для ребенка в возрасте до X по возрасту и полу.

h) *Присвоение места рождения взрослых женщин, имеющих мужа*

402. Если возраст этого лица составляет X или старше и это женщина, то программа должна проверить, есть ли у нее в домохозяйстве муж. Если у этой женщины муж есть и у него правильный код страны рождения, то программа должна указать код его страны рождения в ее записи. Если муж не имеет достоверного кода страны рождения, нужно проверить его запись о годах, прожитых в этом районе. Если код мужа для лет, прожитых в районе, «всегда», то код страны рождения женщины должен быть «эта страна». Если в графе «годы мужа в данном

районе» указано «всегда», то страну рождения женщины следует исчислять по возрасту и полу.

i) *Присвоение места рождения взрослых женщин без мужа*

403. Хотя женщина старше какого-то минимального возраста, установленного группой редактирования, не имеет мужа в домохозяйстве, она может быть матерью ребенка в этом домохозяйстве. В этом случае программа должна отыскать ее старшего ребенка. Если ребенка найти нельзя, программа может исчислить страну рождения по возрасту и полу. Если код ребенка для страны рождения указан правильно и количество указанных лет, прожитых матерью в районе больше возраста ребенка, то программа должна исчислить страну рождения по возрасту и полу. Если количество лет, прожитых матерью в данном месте меньше или равно возрасту ребенка, то программа должна присвоить ей страну рождения ребенка.

j) *Присвоение места рождения для лиц мужского пола*

404. Для того чтобы вычислить место рождения лиц мужского пола, программа редактирования должна попытаться найти данные о его жене или, если он является главой домохозяйства, программа должна попытаться найти данные о его детях. Сначала программа пытается найти данные о жене этого человека. Если она их находит, а количество его лет, прожитых в данном районе, меньше или равняется ее, то в записи об этом человеке указывается страна рождения жены. Если количество лет, прожитых этим человеком в данном районе больше, чем у его жены, то страну рождения следует вычислить по возрасту и полу, используя матрицу интерполяции. Если мужчина является главой семейного домохозяйства, имеет сына или дочь в этом домохозяйстве и количество лет, прожитых им в этом районе, меньше или равняется возрасту ребенка, то программа должна указать ту же страну рождения, что и его ребенка. Если количество времени, прожитого им в этом районе, больше чем возраст его ребенка, то программа должна произвести интерполяцию по возрасту.

2. *Гражданство (P3D)*

405. Сведения о гражданстве должны собираться таким образом, чтобы население можно было разбить на три категории: а) граждане по рождению; б) граждане, натурализованные путем подачи заявления о предоставлении гражданства, путем оптации, брака или иными способами; и с) иностранцы. Помимо этого следует собирать сведения о стране гражданства иностранцев. Необходимо записать страну гражданства как таковую,

а не использовать прилагательное, указывающее на гражданство, так как некоторые из этих прилагательных одновременно обозначают этнические группы. Кодирование информации о стране гражданства должно проводиться достаточно подробно, что позволило бы идентифицировать каждую из всех стран гражданства, представленных среди иностранного населения в стране. В целях кодирования странам рекомендуется использовать числовую систему кодирования, представленную в документе *Standard Country or Area Codes for Statistical Use* (United Nations, 1999). Применение стандартных кодов для классификации иностранного населения в соответствии со страной гражданства увеличит полезность таких данных и позволит организовать между странами обмен информацией об их иностранном населении. Если страна решит объединить страны гражданства в большие группы, им рекомендуется прибегнуть к использованию стандартных региональных и субрегиональных классификаций, описанных в вышеупомянутой публикации (Организация Объединенных Наций, 1998 год, пункт 2.105).

a) *Редактирование данных о гражданстве*

406. Гражданство зависит от принятых в каждой стране определений. В большинстве стран, но не во всех, те, кто родился в стране, автоматически становятся гражданами этой страны по рождению. Следовательно, при редактировании необходимо смотреть на взаимосвязь между местом рождения и гражданством и, возможно, приписывать «граждане по рождению» тем, кто родился в этой стране.

b) *Взаимосвязь между этнической/расовой принадлежностью и гражданством*

407. Некоторые страны собирают также информацию об этнической или расовой принадлежности, которая может дать дополнительные сведения, с тем чтобы использовать их при определении гражданства, особенно когда полученный ответ является недостоверным. Во многих странах данные об этническом происхождении и гражданстве мигрантов в первом и втором поколениях должны почти полностью согласовываться между собой. Для стран с давней историей международной миграции эта характеристика, возможно, не имеет такой ценности, но все же ее можно рассматривать наряду с другими переменными.

c) *Взаимосвязь между натурализацией и гражданством*

408. В странах, где есть натурализация, требования в отношении натурализации могут включаться, а могут и

не включаться в вопросы переписи. Если, например, требуется период постоянного проживания, то для проверки соблюдения срока, необходимого для натурализации, можно использовать элемент «продолжительность проживания». Затем, если то или иное лицо родилось за границей и указало неправильный или не согласующийся ответ в графе «гражданство», группа редактирования может в отношении гражданства присвоить признак «натурализованный». Другим лицам, которые не подпадают под требование о продолжительности проживания, необходимой для натурализации, может присваиваться «иностранец» методом статической интерполяции.

d) *Взаимосвязь между продолжительностью проживания и гражданством*

409. Элемент «продолжительность проживания» может в вопросник не включаться или может быть неоднозначным при определении гражданства или же группа редактирования может отказаться от его использования. Затем, если значение гражданства указано неправильно или не согласуется с местом рождения, то, если не применяется метод динамической интерполяции, указывается «не известно». Страны, применяющие метод динамической интерполяции для исчисления достоверных величин, должны пользоваться, по меньшей мере, двумя характеристиками (и одна из них, вероятно, должна быть «место рождения») для получения «известной» информации, исходя из аналогичных сведений о других лицах в данном географическом районе.

3. *Продолжительность проживания (PID)*

410. Под продолжительностью проживания имеется в виду промежуток времени вплоть до даты проведения переписи, выраженной в прошедших полных годах, в течение которых это лицо проживало: а) в населенном пункте, который является его или ее постоянным местом жительства в момент переписи; и б) в крупной или более мелкой административно-территориальной единице, в которой расположен этот населенный пункт (Организация Объединенных Наций, 1998 год, пункт 2.35).

a) *Редактирование данных о продолжительности проживания*

411. Как и страна рождения, продолжительность проживания имеет важное значение при составлении статистических данных о перемещении населения. В некоторых случаях какая-то подгруппа населения может быть намного более мобильной, чем вся нация в целом.

При редактировании этого элемента учитывается место рождения данного лица и ответы других членов домохозяйства. Данные о продолжительности проживания должны редактироваться вместе со сведениями о «прежнем месте жительства» или «местожительстве в определенный момент в прошлом».

b) *Проживание и продолжительность де-факто/де-юре*

412. Результаты редактирования могут зависеть от того, как проводится перепись: де-факто или де-юре. Поскольку во время переписи де-юре собираются сведения о постоянном месте жительства, данные о продолжительности проживания не могут выявить ту же информацию, что и при переписи де-факто, когда регистрация проводится по месту проживания в ночь переписи. Кроме того, при кодировании и редактировании необходимо учитывать тех, кто или «всегда» жил в данном месте, или «никогда не покидал» его. В отношении этих лиц программа редактирования должна пропустить согласованность и другие приемы редактирования.

c) *Отношение возраста к продолжительности проживания*

413. На начальном этапе процесса редактирования следует произвести проверку на согласованность между возрастом и местом рождения и на достоверность данных о годах, прожитых в данном населенном пункте или административно-территориальной единице. Количество лет, прожитых тем или иным лицом в одном населенном пункте или административно-территориальной единице, не может превышать возраст этого лица. Кроме того, тот, кто родился за пределами страны, не мог всегда проживать в данном населенном пункте или административно-территориальной единице. Программа должна приписывать «всегда» годам, прожитым в населенном пункте или административно-территориальной единице, если количество лет, прожитых в населенном пункте или административно-территориальной единице, превышает возраст, а страной рождения является эта страна. Если количество лет, прожитых в населенном пункте или административно-территориальной единице, превышает возраст, а страной рождения является не эта страна, то возраст данного лица следует соизмерять с годами, прожитыми в населенном пункте или административно-территориальной единице. В этом случае исходят из того, что данный человек, хотя и родился за пределами этой страны, переехал в данный населенный пункт или административно-территориальную единицу в возрасте менее одного года.

d) *Отношение места рождения к продолжительности проживания*

414. В том случае, если вводимые данные выходят за определенные пределы, следует произвести такую же проверку, как и в отношении места рождения. Нужно найти предыдущих родственников (мать, главу домохозяйства, мужа, ребенка). Интерполяция должна производиться на основе найденной информации. Однако прежде чем приписать то или иное значение, нужно убедиться, что оно согласуется с возрастом и местом рождения лица, чьи записи редактируются.

e) *Для тех, кто всегда жил в данном месте*

415. Если на вопрос о количестве лет, прожитых в данном населенном пункте или административно-территориальной единице, человек отвечает «всегда», но страна его рождения не «эта страна», то группа редактирования, возможно, соотнесет возраст данного лица с продолжительностью проживания в данном населенном пункте или административно-территориальной единице. Специалисты будут исходить из того, что, хотя человек родился за пределами этой страны, он переехал в данный населенный пункт или административно-территориальную единицу в возрасте менее одного года. На следующем этапе редактирования будет произведена проверка на достоверность данных о годах проживания в населенном пункте или административно-территориальной единице. Поскольку промежуток времени, в течение которого человек жил в данном населенном пункте или административно-территориальной единице, не может превышать возраст этого лица, его возраст будет соотнесен в этой ситуации с годами, проведенными в данном населенном пункте или административно-территориальной единице.

f) *Исчисление продолжительности проживания данного лица по продолжительности проживания матери*

416. Если данная категория не имеет достоверного кода, то программа может произвести проверку по записям в поисках данных о матери этого лица в данных о домохозяйстве. При положительном результате запись матери может дать полезную информацию для указания недостающих величин. Если мать данного лица всегда проживала в этом населенном пункте или административно-территориальной единице и ее страной рождения является «эта страна» (как и должно быть), то программа отнесет количество лет, прожитых этим лицом в населенном пункте или административно-территориальной единице, к категории «всегда». Если страной рождения матери является не «эта страна», да-

же если в категории, годы, прожитые ей в данном населенном пункте или административно-территориальной единице, указано «всегда», это свидетельствует о том, что в категориях матери что-то указано неправильно. Тогда программа не придаст значения стране рождения матери и припишет возраст по продолжительности проживания в населенном пункте или административно-территориальной единице. Если в отношении количества лет, прожитых матерью в данном населенном пункте или административно-территориальной единице, указано на «всегда», но код присвоен правильно, а возраст данного лица меньше количества лет, которые мать прожили в данном населенном пункте или административно-территориальной единице, то программа редактор вернется назад и проверит страну рождения матери. Если страна рождения матери указана как «эта страна», программа соотнесет возраст этого лица с годами, прожитыми в населенном пункте или административно-территориальной единице. Однако если возраст этого лица равен или превышает годы, прожитые матерью в населенном пункте или административно-территориальной единице, программа соотнесет годы, прожитые матерью в населенном пункте или административно-территориальной единице, с годами, прожитыми этим лицом в данном населенном пункте или административно-территориальной единице.

g) *Исчисление продолжительности проживания того или иного лица по продолжительности проживания ребенка*

417. Если данное лицо является ребенком (сыном или дочерью) программа редактирования должна проверить запись о главе домохозяйства для возможного получения информации, которая может помочь присвоению величин для недостающих сведений о продолжительности проживания. Если глава домохозяйства родился в «этой стране» и всегда проживал в данном населенном пункте или административно-территориальной единице, программа присвоит признак «всегда» годам, прожитым ребенком в данном населенном пункте или административно-гражданской единице. Если глава домохозяйства всегда проживал в данном населенном пункте или административно-гражданской единице, но родился не в «этой стране», возраст ребенка будет соотнесен с населенным пунктом или административно-гражданской единицей. Если в графе «годы, проведенные в населенном пункте или административно-гражданской единице», для главы домохозяйства указано не «всегда», но при этом код является правильным, эту информацию можно использовать, если она согласуется с редактируемыми данными о возрасте ребенка. Если возраст ребенка равен или превышает количество лет, прожитых главой домохозяйства в данном населенном

пункте или административно-гражданской единице, то программа использует годы, прожитые главой домохозяйства в населенном пункте или административно-гражданской единице, как годы, прожитые там сыном или дочерью. Если возраст ребенка меньше, чем глава домохозяйства проживал в населенном пункте или административно-территориальной единице, то программа присвоит значение в зависимости от страны рождения главы домохозяйства. Это значение будет «всегда», если глава домохозяйства родился в «этой стране»; в противном случае программа соотнесет возраст сына или дочери с годами, прожитыми в данном населенном пункте или административно-территориальной единице.

h) *Продолжительность проживания, когда нет никакой другой информации*

418. Когда все, о чем говорилось выше, не дает достоверного значения, программа может приписать в графе «годы в населенном пункте или административно-территориальной единице» в отношении этого лица «не сообщается» или «не известно». Если значение по-прежнему недостоверно, то «не известно» следует приписывать, если не используется метод динамической интерполяции. Страны, применяющие метод динамической интерполяции для исчисления недостоверных величин, должны пользоваться, по меньшей мере, двумя характеристиками для получения «известной» информации, исходя из аналогичных сведений о лицах в данном географическом районе.

4. *Прежнее место жительства (PIE)*

419. Под прежним местом жительства имеется в виду крупная или более мелкая административно-территориальная единица, либо другая страна, в которой данное лицо проживало непосредственно до приезда в нынешнюю территориально-административную единицу на постоянное жительство (Организация Объединенных Наций, 1998 год, пункт 2.38).

a) *Редактирование данных о прежнем месте жительства*

420. Элемент «прежнее место жительства» следует редактировать вместе с элементом «продолжительность проживания». Если человек родился в данном месте (страна, населенный пункт или административно-территориальная единица), в зависимости от элемента переписи (и никогда отсюда не переезжал), в этом вопросе необходимо ставить или пробел или поставить конкретный код «никогда не покидал». Однако пробелы могут вызвать проблемы при составлении

таблиц, поэтому группе редактирования необходимо решить, как лучше справиться с такой ситуацией.

b) *Прежнее место жительства при изменении границ*

421. Границы стран со временем меняются, поэтому нужно следить за тем, чтобы это соответствующим образом отражалось в схемах кодирования. Кроме того, коды следует устанавливать так, чтобы можно было составлять логические группы. Например, как говорилось выше, первая цифра в трехцифровом коде может представлять континент проживания, вторая цифра — регион на этом континенте, и третья цифра — страну в этом регионе.

c) *Если человек никуда не переезжал с рождения*

422. Обработчики данных составляют таблицы по отдельным элементам. Таким образом, специалисты должны быть уверены в том, что кроме других кодов, присваиваемых месту жительства, используется специальный код «родился здесь». В этом случае программа должна различать тех, кто родился в данном месте, и тех, кто родился в одном месте, а затем переехал в другое место в том же географическом районе.

d) *Использование данных о других лицах в единице*

423. Если прежнее место жительства указано неправильно или не согласуется с другими элементами, то, как правило, применяется прием, аналогичный тому, который используется при редактировании элемента «продолжительность проживания». Программа редактирования может просмотреть прежнее место жительства матери, если она указана в жилищной единице. Затем она может прочитать прежнее место жительства главы домашнего хозяйства, используя данные о детях и взрослых в тех странах, где взрослые перемещаются нечасто.

e) *Когда нет данных о других приемлемых лицах для редактирования прежнего места жительства*

424. Если вышеуказанные процедуры не дают достоверного значения, программа может для этого лица указать «не сообщается», или «не известно», когда речь идет о годах, прожитых на прежнем месте. Если значение остается неправильным, то следует приписать «не известно», когда не применяется метод динамической интерполяции. Страны, применяющие метод динамической интерполяции для исчисления недостоверных величин, должны пользоваться, по меньшей мере, двумя характеристиками для получения «известной» ин-

формации, исходя из аналогичных сведений о лицах в данном географическом районе.

5. Место жительства в определенный момент в прошлом (PIF)

425. Под местом жительства в определенный момент в прошлом имеется в виду крупная или более мелкая административно-территориальная единица либо другая страна, в которой данное лицо проживало в определенный момент, предшествовавший переписи. Выбранная базисная дата должна быть наиболее подходящей для национальных целей. В большинстве случаев — это срок в один год или пять лет, предшествовавших переписи. Базисная дата первого вида обеспечивает текущую статистику миграции в течение одного года, а базисная дата второго вида может быть более подходящей для сбора данных в целях проведения анализа международной миграции, хотя вероятно, может быть менее пригодной для анализа текущей внутренней миграции. При выборе базисной даты необходимо также принимать во внимание вероятность наличия у отдельных лиц способности точно вспомнить свое постоянное место жительства, где они проживали за один год или за пять лет до даты переписи. В странах, проводящих перепись каждые пять лет, пятилетний срок для большинства лиц может быть легко увязан со временем предыдущей переписи. В других случаях человек может более точно вспомнить, где он проживал за один год, чем за пять лет до проведения переписи. Однако некоторые страны могут использовать другую базисную дату, отличную от одного или пяти лет, предшествовавших переписи, поскольку оба промежутка времени могут вызвать трудности, связанные с восстановлением в памяти необходимых данных. При тех или иных национальных условиях может возникнуть необходимость устанавливать базисную дату, которая будет ассоциироваться с важным событием, запомнившимся большинству населения. Кроме того, для изучения международной миграции может быть полезной информация относительно года прибытия в страну (Организация Объединенных Наций, 1998 год, пункт 2.40).

426. Данные о месте жительства в определенный момент в прошлом редактируются аналогично данным о прежнем месте жительства. Как правило, страны спрашивают и «продолжительность проживания» и «прежнее место жительства», или просто «место жительства в определенный момент». Если человек родился в месте регистрации (страна, населенный пункт или административно-территориальная единица, в зависимости от элемента переписи) и никогда отсюда не переезжал, здесь нужно оставить или пробел, или поставить специальный код «никогда не покидал». Как указывалось ранее, пробелы могут вызвать проблемы при составле-

нии таблиц. Затем применяются те же процедуры для прежнего места жительства, которые описаны в трех предыдущих пунктах.

С. СОЦИАЛЬНЫЕ ХАРАКТЕРИСТИКИ

427. Социальные характеристики в каждой стране разные, но, как правило, они включают в себя элементы, характеризующие различные аспекты социально-культурных условий в стране. Элементы, касающиеся образования, включая грамотность, посещение учебных заведений и образовательную подготовку, а также область знаний и свидетельства об образовании можно классифицировать в соответствии с категориями, приведенными в пересмотренном издании Международной стандартной классификации образования (МСКО) 1997 года, выпущенном Организацией Объединенных Наций по вопросам образования, науки и культуры (ЮНЕСКО) (Организация Объединенных Наций, 1998 год, пункты 2.144-2.164). В данном разделе говорится также о нетрудоспособности, дефектах и инвалидности и причинах нетрудоспособности. Общие понятия и определения в связи с вопросами, касающимися нетрудоспособности, можно найти в Международной классификации дефектов, нетрудоспособности и инвалидности (МКДНИ) (выпущенной Всемирной организацией здравоохранения в 1980 году). Ниже рассматриваются также и другие социальные характеристики, такие как вероисповедание, язык и этническая принадлежность.

1. Способность читать и писать (грамотность) (P5A)

428. Желательно, чтобы сведения о грамотности собирались в отношении всех лиц в возрасте 10 лет и старше. Однако в ряде стран некоторые лица в возрасте от 10 до 14 лет, возможно, вскоре станут грамотными благодаря учебе в школе и процент грамотного населения для данной возрастной группы может не соответствовать действительности и вводить в заблуждение. Поэтому при проведении международных сравнительных исследований грамотности данные о ней следует собирать в отношении всех лиц в возрасте от 15 лет и старше. В тех случаях, когда страны собирают сведения в отношении лиц более молодого возраста, при составлении таблиц с данными о грамотности следует, по крайней мере, проводить различие между лицами младше 15 лет и лицами в возрасте 15 лет и старше (Организация Объединенных Наций, 1998 год, пункт 2.147).

429. Каждая страна должна установить минимальный возраст для составления таблиц с данными о грамотности; при этом группы редактирования должны решить, какой минимальный возраст нужно брать для редакци-

рования данных о грамотности, поскольку для внутреннего пользования могут понадобиться дополнительные таблицы. По мере составления переписного листа группы редактирования должны принять решение относительно минимального возраста с точки зрения сбора данных, и с какого образовательного уровня этот вопрос уже не стоит задавать. В связи с этим, если респондент уже достиг определенного уровня образования счетчику видимо уже нет необходимости задавать ему вопрос о грамотности.

430. При редактировании данных о грамотности, прежде всего, проверяется последняя законченная ступень; если в графе последняя ступень стоит «грамотен» исходя из спецификаций, то следует присвоить код «да». Лица на определенном уровне обучения должны считаться грамотными. В тех случаях, когда обнаруживается неправильный код для грамотности, следует написать значение. При этом должна стоять запись «не указано» или значение следует определить, используя матрицу интерполяции, исходя из конкретных переменных, таких как последняя ступень обучения и пол.

2. *Посещение учебных заведений (P5B)*

431. В принципе информация о посещении учебных заведений должна собираться для лиц всех возрастов. Это относится, в частности, к населению официально установленного возраста обучения, который составляет, как правило, от 5 до 29 лет, но может варьироваться в зависимости от страны и особенностей ее национальной системы образования. В случае более широкого диапазона собираемых данных, охватывающего сведения о посещении дошкольных учреждений и/или участия в других систематических образовательных и учебных программах, организуемых для взрослых на производственных предприятиях и предприятиях сферы услуг, в общинных организациях и других неучебных учреждениях, возрастной диапазон может устанавливаться в зависимости от конкретных обстоятельств (Организация Объединенных Наций, 1998 год, пункт 2.151).

а) *Редактирование данных о посещении учебных заведений*

432. В каждой стране группа редактирования должна решить, по каким возрастным группам собирать данные о посещении учебных заведений. Поскольку в большинстве стран обучение подразделяется еще и на несколько уровней, то для того, чтобы определить эти уровни по возрасту, специалисты должны также решить, какие возрастные группы соответствуют различным уровням обучения. При этом записи по всем ос-

тальным лицам необходимо изменить. Если программа редактирования дает несогласующиеся ответы для этой категории, необходимо изменить или возраст, или уровень обучения. Как правило, возраст устанавливается на момент редактирования, поэтому изменяется уровень обучения. Счетчикам следует дать указание опускать уровень обучения для лиц старше заранее установленного возраста, если это подходит для данной страны. В тех случаях, когда лица среднего возраста продолжают получать среднее или высшее образование, то может быть и нецелесообразным устанавливать верхний возрастной предел для обучения. Предполагается, что ответы и комбинации ответов проверяются еще до переписи на предварительном этапе, поэтому такие решения можно принимать до фактического проведения переписи.

б) *Прием на очное обучение или заочное обучение*

433. Некоторые страны могут по своему усмотрению собирать информацию об очном или заочном обучении. Если такой элемент включается в перепись, то его можно редактировать вместе с данными о посещении учебных заведений или в качестве отдельного пункта.

с) *Согласованность между сведениями о посещении учебных заведений и экономической активностью*

434. При редактировании в первую очередь следует проверять данные на согласованность с другими основными элементами, например, с основной экономической активностью. Если в данном об основной экономической активности указано посещение учебного заведения и респондент ответил, что его или ее основной активностью является посещение школы по элементу «посещение учебного заведения» нужно присвоить код для ответа «да», а основную экономическую активность указать как «студент» в целях согласования ответов между собой. Во всех других случаях должны приниматься любые достоверные ответы.

д) *Присвоение той или иной категории для недостоверных или несогласующихся ответов о посещении учебных заведений*

435. Если указанные данные выходят за установленные пределы, а запись, касающаяся последней законченной ступени, достоверна, то значение следует присваивать, используя матрицу интерполяции по возрасту, полу и последней ступени обучения. Если последняя ступень имеет неправильный код, то для присвоения значения о посещении учебных заведений следует использовать запись в графе «грамотность». Если грамот-

ность имеет неправильный код, то запись о посещении учебных заведений следует исчислять только на основе возраста и пола.

436. Матрицы интерполяции могут понадобиться для того, чтобы показать различные принципы, отражающие посещение учебных заведений по полу и возрасту (иногда по одному и тому же году) или по небольшим возрастным группам.

3. Образовательная подготовка (последняя завершенная ступень или уровень) (P5C)

a) Редактирование данных об образовательной подготовке

437. При редактировании данных об образовательной подготовке (последняя ступень или уровень) следует a) проверить на согласованность достоверную запись и возраст и b) интерполировать запись, если первоначальная запись выходит за установленные пределы. Как указывалось ранее, в странах, которые не используют метод динамической интерполяции значение должно быть «не указано». В странах, где используется метод динамической интерполяции, пол и один год, указывающий на возраст, потребуется для молодых людей, а пол и небольшие возрастные группы потребуются для детей чуть старше. Для стран, где указываются сведения о последней ступени или последнем уровне образования, могут понадобиться многомерные матрицы интерполяции.

b) Минимальный возраст для образовательной подготовки

438. В каждой стране группы редактирования должны определить минимальный возраст поступления в школу. При установлении минимального возраста высший завершённый уровень, как правило, не должен превышать возраст данного лица плюс какая-то постоянная величина (представляющая этот минимальный возраст поступления в школу). Кроме того, необходимо использовать один год, указывающий на возраст детей, поскольку обновление матриц интерполяции может привести к ошибкам в том случае, если возрастные группы весьма широкие по своему диапазону.

c) Соотношение возраста и образовательной подготовки

439. Группа редактирования должна также определиться с тем, сколько искажений допустимо в наборе данных. Как правило, лучше внести изменения в нескольких исключительных случаях, когда не согласо-

уются возраст и образовательная подготовка, чем принимать большое количество ответов, которые действительно не совместимы. Исходя из этого, в тех случаях, когда исходная запись выходит за установленные пределы или не согласуется с возрастом, то значение можно приписать. Для стран, где не используется метод динамической интерполяции, можно ввести «не указано». Там, где этот метод применяется, запись можно получить исходя из возраста (включая единый год для лиц школьного возраста), пола и посещения учебных заведений. ЮНЕСКО признает грамотность отдельно от образовательной подготовки и поэтому «способность читать и писать», вероятно, не следует использовать в качестве величины в матрице интерполяции.

4. Область знаний и свидетельства об образовании (P5D)

440. Информация о лицах, собираемая по признаку уровня образования и области знаний, важна для анализа соответствия между предложением и спросом на квалифицированную рабочую силу с учетом конкретной специализации на рынке труда. Это одинаково важно для планирования и регулирования численности учащихся учебных заведений и программ профессиональной подготовки различных уровней, типов и профиля (Организация Объединенных Наций, 1998 год, пункт 2.158).

441. Лица моложе 15 лет (или другого предопределенного возраста) не должны указывать информацию об области знаний и/или свидетельствах об образовании. Для лиц в возрасте 15 лет и старше должно быть соотношение между уровнем образовательной подготовки и областью знаний и/или свидетельствами об образовании. Каждый раз, когда возникают недостоверные записи, страны, не использующие метод динамической интерполяции, могут ввести «не известно». Страны, применяющие метод динамической интерполяции, могут рассмотреть вопрос об использовании возраста, пола, образовательной подготовки и, по возможности, рода занятий для исчисления области знаний и/или свидетельств об образовании.

5. Вероисповедание (P3E)

442. В целях переписи понятие вероисповедания может быть определено как a) религиозная или духовная вера, которой отдается предпочтение, независимо от того, представлена или нет эта вера организованной группой или b) принадлежность к организованной группе, придерживающейся определенной религиозной или духовной доктрины. Каждая страна, которая при проведении переписи учитывает вероисповедание, должна

пользоваться тем определением, которое лучше всего соответствует ее собственным нуждам, и должна указывать в опубликованных результатах переписи применяемое определение (Организация Объединенных Наций, 1998 год, пункт 2.109).

a) *Редактирование данных о вероисповедании*

443. Вероисповедание — это одна из переменных, соответствующих примерам, приведенным в главе II. Достоверное значение получается для каждого лица или непосредственно от другого члена домохозяйства, если такое значение имеется, или от другого главы домохозяйства с аналогичными характеристиками. Группа редактирования должна составить план логического редактирования для использования в отношении других социальных переменных. Сначала нужно определить главу домохозяйства и отредактировать данные о нем независимо от того, является ли он или она первым лицом в единице. Если лицо с неправильно указанным или неизвестным вероисповеданием является главой домохозяйства, то нужно сделать следующее:

b) *Вероисповедание главы домохозяйства, отличающееся от вероисповедания кого-либо другого в жилищной единице*

444. Сначала нужно определить, есть ли достоверные данные о вероисповедании кого-либо другого в жилищной единице, и приписать первый достоверный ответ о вероисповедании.

c) *Вероисповедание главы домохозяйства или кого-либо другого в жилищной единице*

445. Если вероисповедание кого-либо в домохозяйстве не указывается, следует или приписать «не известно» (если данная страна не использует метод динамической интерполяции) или вычислить вероисповедание по данным ближайшего главы домохозяйства с аналогичными характеристиками, включая возраст и пол, а также язык, место рождения и другие переменные с учетом соответствующих обстоятельств.

d) *Для лица помимо главы домохозяйства без указания вероисповедания*

446. Если данное лицо не является главой домохозяйства и не указывает свое вероисповедание, приписать вероисповедание главы домохозяйства.

6. Язык (P3F)

447. Имеется три категории данных о языке, которые могут быть собраны при проведении переписи (Организация Объединенных Наций, 1998 год, пункт 2.112):

- a) родной язык, определяемый как язык, на котором обычно говорили в доме опрашиваемого лица в его раннем детстве;
- b) обычно употребляемый язык, определяемый как язык, на котором в настоящее время говорит или на котором наиболее часто говорит опрашиваемое лицо в его или ее нынешнем доме;
- c) умение говорить на одном или нескольких обозначенных языках.

a) *Редактирование данных о языке*

448. Из трех разных категорий языка, которые могут быть указаны в переписном листе (родной язык, обычно употребляемый язык и умение говорить на одном или нескольких обозначенных языках) первые два, родной язык и обычно употребляемый язык, взаимосвязаны. Если в переписном листе указаны обе категории, группы редактирования должны взвесить возможность редактировать их вместе. Если один из них указан неправильно, другой можно использовать для восполнения записи.

b) *Редактирование данных о языке: глава домохозяйства*

449. Язык — это еще одна переменная, для примеров, представленных в главе II. Группы редактирования должны составить план логического редактирования для использования в отношении других социальных переменных, редактируя в первую очередь данные о главе домохозяйства. Если лицо, чей язык указан неправильно или неизвестен (родной язык или обычно употребляемый язык), является главой домохозяйства, в первую очередь необходимо определить, есть ли кто-либо еще в жилищной единице, чей язык указан правильно, и приписать первому правильно указанный язык. Если таких нет, нужно приписать «не известен», если не применяется метод динамической интерполяции, или вычислить язык по данным ближайшего главы домохозяйства с аналогичными характеристиками, включая возраст и пол, а также в зависимости от этих обстоятельств, другие языковые переменные, место рождения и другие переменные.

c) *Редактирование данных о языке: лица помимо главы домохозяйства*

450. Если данное лицо не является главой домохозяйства, а язык указан неправильно, то приписать ему язык главы домохозяйства.

d) *Редактирование данных о языке: использование данных об этническом происхождении или месте рождения*

451. Язык и этническое происхождение, а иногда и место рождения, тесно взаимосвязаны, и в некоторых странах эти данные можно редактировать вместе. Кроме того, группы редактирования должны рассмотреть возможность присвоения кодов, чтобы отразить соотношение этих переменных. В зависимости от количества цифр в коде и распределения языков и этнических групп в стране можно установить соответствия, с тем чтобы легче было восполнить неизвестные или несогласующиеся ответы.

e) *Редактирование данных о языке: родной язык*

452. Если родной язык не известен, но данное лицо является филиппинцем или родилось на Филиппинах, то можно приписать соответствующий эквивалентный язык: тагальский, илоканский или какой-либо другой язык, распространенный на Филиппинах. Язык присваивается, таким образом, только главе домохозяйства, а код этого языка присваивается другим членам домохозяйства, но в каждой стране группа редактирования должна рассматривать конкретные обстоятельства, в том числе территориальность (например, проживание в городских или сельских районах), возраст или другие элементы.

f) *Редактирование данных о языке: умение говорить на обозначенном языке*

453. Умение говорить на обозначенном языке — это третья переменная для примеров, представленных в главе II. Кроме того, в первую очередь следует редактировать данные о главе домохозяйства. Если значение для языка главы домохозяйства недостоверно или не известно, сначала надо посмотреть, есть ли кто-либо еще в жилищной единице с достоверным ответом об умении говорить на том или ином языке, и приписать первый достоверный язык. После этого, если таких в жилищной единице нет, или приписать «не известно», если в данной стране не используется метод динамической интерполяции, или вычислить язык по данным ближайшего главы домохозяйства с аналогичными характеристиками (например, возраст и пол, а также место рождения и другие переменные в зависимости от обстоятельств). Если данное лицо не является главой домохозяйства, а ответ об умении говорить на обозначенном языке недостоверен, то приписать данные главы домохозяйства об умении говорить на том или ином языке.

7. *Этническая принадлежность (P3G)*

454. Потребность в информации о национальных и/или этнических группах населения зависит от существующих в той или иной стране условий. Некоторые основные признаки, по которым различаются этнические группы, включают в себя этническую национальность (то есть страну или район происхождения в отличие от гражданства или страны юридической гражданской принадлежности), расу, цвет кожи, язык, вероисповедание, обычаи, касающиеся одежды и питания, племенную принадлежность или различные комбинации этих признаков. Кроме того, некоторые из употребляемых здесь терминов, например, «раса», «происхождение» и «племя» имеют целый ряд дополнительных оттенков значения. Исходя из этого определения и критерии, которыми пользуется каждая страна при изучении этнических характеристик своего населения, должны устанавливаться теми группами населения, которые страна желает выявить. По самому характеру данного вопроса эти группы весьма различны в разных странах; поэтому здесь нельзя рекомендовать приемлемые в международном масштабе критерии (Организация Объединенных Наций, 1998 год, пункт 2.116).

a) *Редактирование данных об этнической принадлежности*

455. Некоторые другие переменные, если они указываются в ходе переписи, могут быть полезными при «определении» этнической принадлежности, если она указана неправильно или неизвестна. Во многих странах существует взаимосвязь между местом рождения как в данной стране, так и в других странах, и этнической принадлежностью. «Родной язык» также зачастую является хорошим показателем этнической принадлежности во многих странах, поскольку категории, а значит и коды, будут одними и теми же или аналогичными.

b) *Редактирование данных об этнической принадлежности: глава домохозяйства*

456. Этническое происхождение также сообразуется с примерами, приведенными в главе II. Группа редактирования должна следовать той схеме, которая уже приводилась для других социальных переменных. Данные о главе домохозяйства должны редактироваться в первую очередь. Если лицо, данные об этническом происхождении которого недостоверны или неизвестны, является главой домохозяйства, следует сначала найти достоверные данные об этническом происхождении кого-либо другого лица в жилищной единице и приписать первый достоверный признак этнической принадлежности. Если таких данных нет, то тогда приписать «не

известно» или если данная страна не применяет метод динамической интерполяции, вычислить этническую принадлежность по данным ближайшего главы домохозяйства с аналогичными характеристиками (возраст и пол, а также язык, место рождения и другие переменные, которые подходят в зависимости от обстоятельств).

c) *Редактирование данных об этнической принадлежности: лица помимо главы домохозяйства*

457. Если лицо не является главой домохозяйства, а этническое происхождение указано неверно, то следует приписать этническое происхождение главы домохозяйства.

d) *Редактирование данных об этнической принадлежности: использование сведений о языке и месте рождения*

458. Этническое происхождение и язык, а иногда и место рождения тесно взаимосвязаны, и в отношении некоторых стран эти сведения можно редактировать вместе. Кроме того, группы редактирования должны продумать вопрос о присвоении кодов, с тем чтобы отразить соотношение этих переменных. В зависимости от количества цифр в коде и распределения этнических групп и языков в стране можно построить соответствия, которые помогут восполнить неизвестные или несогласующиеся ответы.

459. Например, если этническое происхождение не известно, но лицо говорит на одном из языков, распространенных на Филиппинах, и родилось на Филиппинах, можно присвоить соответствующее эквивалентное этническое происхождение, то есть филиппинское. Как правило, таким образом, присваивается только этническое происхождение главы домохозяйства (а другим членам присваивается этот код), но в каждой стране группа редактирования должна рассмотреть конкретные обстоятельства, в том числе территориальный признак (например, проживание в городских и сельских районах), возраст или другие элементы.

8. Нетрудоспособность (P8A)

460. Для того чтобы установить рамки аспекта нетрудоспособности как нетрудоспособное, определяется лицо, которое ограничено в виде или объеме деятельности, которую оно может осуществлять из-за испытываемых трудностей, вызванных длительным физическим или психическим недугом или иными проблемами со здоровьем. Непродолжительная нетрудоспособность, наступившая в силу временных обстоятельств, таких как перелом ноги или заболевание, исключается. Следует учитывать только ту нетрудоспособность, которая

продолжается не более шести месяцев (Организация Объединенных Наций, 1998 год, пункт 2.262).

461. Вопрос, заданный с целью выявления лиц с нетрудоспособностью, должен включать в себя широкие категории нетрудоспособности, с тем чтобы каждое лицо могло определить для себя наличие и отсутствие каждого вида нетрудоспособности. Рекомендуется использовать следующий перечень основных категорий на основе международной классификации дефектов, нетрудоспособности и инвалидности (МКДНИ): затруднения в связи со зрением (даже с очками, если они используются); в связи со слухом (даже с помощью слухового аппарата, если он используется); дефекты речи (разговор); затрудненность движения/передвижения (ходьба, подъем по лестнице, способность стоять); затрудненность телодвижений (дотянуться, согнуться, встать на колени); хватание/держание (использование пальцев при хватании или передвижении предметов); трудности в обучении (проблемы интеллектуального характера, задержка в умственном развитии); поведенческие проблемы (психологического, эмоционального характера); трудности в уходе за собой (купание, одевание, принятие пищи); прочее (указать). Лицо определяется как нетрудоспособное в том случае, если укажет на наличие у него одного или более отклонений, имеющих в списке (Организация Объединенных Наций, 1998 год, пункт 2.264).

a) *Редактирование данных о нетрудоспособности*

462. Если кто-то не отвечает на данный вопрос, то в этом случае трудно определить обусловлено ли это неспособностью или нежеланием респондента отвечать на этот вопрос или это связано с какой-либо другой причиной. Группа редактирования в стране должна решить, следует ли редактировать данные сведения обычным порядком, приписывая «не известно», если не используется метод динамической интерполяции, или использовать ответы других лиц, если применяется динамическая интерполяция. В качестве альтернативы специалисты могут решить, что следует принимать только те ответы, где указывается конкретная нетрудоспособность и что любой недостоверный ответ следует отмечать как «нетрудоспособность отсутствует». В этом случае динамическая интерполяция не применяется.

b) *Разные виды нетрудоспособности*

463. Страны, собирающие информацию о разных видах нетрудоспособности, должны видоизменить приемы редактирования. Программа редактирования должна следить за тем, сколько всего может быть видов нетрудоспособности и за дублированием и распределением этих видов нетрудоспособности. Как и прежде, боль-

шинство стран сочтут нецелесообразным использовать данные, построенные на сведениях о нетрудоспособности других лиц, поэтому в случае недостоверности, возможно, придется использовать такие признаки, как «не известно» или даже «не известно, есть ли нетрудоспособность».

9. Дефекты и инвалидность (P8B)

464. Уровни инвалидности могут быть оценены путем сравнительного анализа данных о тех лицах, которые заявляют о своей нетрудоспособности, и тех, кто дает на этот счет отрицательные ответы, по другим характеристикам, таким, как образование и занятость. Страны могут быть также заинтересованы в получении сведений относительно конкретных условий, при которых нетрудоспособные люди ощущают свою инвалидность, например, при пользовании общественным транспортом, на работе или во время проведения общественных мероприятий. Информация такого рода может быть использована для сокращения конкретных факторов, которые изолируют нетрудоспособных людей (физические препятствия, отсутствие соответствующего обслуживания, негативное отношение со стороны окружающих, предупреждения по отношению к нетрудоспособным). Вопрос об инвалидности должен быть нацелен на выявление тех проблем, которые мешают нетрудоспособным лицам с физическими или психическими недостатками участвовать наравне с другими с жизни общества. Чтобы дать представление об установке, в которой живет нетрудоспособный человек, надо рассматривать как физические, так и социальные факторы (Организация Объединенных Наций, 1998 год, пункт 2.276).

а) Редактирование данных о дефектах и инвалидности

465. В этом случае группа редактирования также должна решить, стоит ли редактировать эти сведения обычным порядком, присваивая признак «не известно», если не используется динамическая интерполяция, или использовать ответы других лиц с применением метода динамической интерполяции. В качестве альтернативы специалисты могут решить, что будут приниматься только те ответы, где указывается на инвалидность, а любой недостоверный ответ будет отмечаться как «инвалидность отсутствует».

10. Причины нетрудоспособности (P8C)

466. Информация о причинах нетрудоспособности важна для планирования и оценки профилактических программ. В силу ограниченности места в переписных

листах информация о причинах может быть получена посредством постановки общего вопроса об условиях, при которых возникла нетрудоспособность, вместо прямых вопросов относительно конкретных заболеваний или повреждений. При сборе данных о причинах нетрудоспособности следует использовать пять основных категорий: а) врожденные/внутриутробные; б) болезнь/заболевания, например, полиомиелит, проказа, катаракта; в) повреждения/несчастные случаи/травмы, например, дорожные и транспортные аварии, повреждения в результате случайных падений, пожаров, военных действий или случайных отравлений; г) прочее; д) неизвестные причины (Организация Объединенных Наций, 1998 год, пункт 2.277).

а) Редактирование данных о причинах нетрудоспособности

467. Группа редактирования в той или иной стране должна решить, следует ли редактировать данные сведения обычным порядком, присваивая признак «не известно», если не применяется метод динамической интерполяции, или использовать ответы других лиц, если применяется динамическая интерполяция. В качестве альтернативы специалисты могут решить, что можно принимать только те ответы, где указывается причина нетрудоспособности, и в этом случае матрица интерполяции использоваться не будет.

Д. ЭКОНОМИЧЕСКИЕ ХАРАКТЕРИСТИКИ

468. Информацию о статусе экономической активности следует, в принципе, собирать в отношении всего населения, но на практике такая информация собирается по каждому лицу, достигшему минимального возраста, установленного в каждой стране в соответствии с ее условиями. Минимальный возраст окончания школы не должен автоматически приниматься за нижний возрастной предел для сбора информации о статусе экономической активности. В странах, где обычно многие дети участвуют в сельскохозяйственных работах или других видах экономической деятельности (например, разработка полезных ископаемых, ткачество и мелкая торговля), потребуется выбрать более низкий минимальный возраст, чем в странах, для которых подобная занятость детей нехарактерна. В таблицах данных об экономических характеристиках должно как минимум проводиться различие между лицами в возрасте до 15 лет и лицами в возрасте 15 лет и старше; страны, где минимальный возраст окончания школы выше 15 лет и где есть экономически активные дети младше этого возраста, должны стремиться получить данные об экономических характеристиках таких детей в целях дос-

тижения международной сопоставимости данных, по крайней мере, для лиц в возрасте 15 лет и старше. Участие в экономической деятельности пожилых мужчин и женщин после достижения ими обычного пенсионного возраста также зачастую недоучитывается. Это требует пристального внимания для определения экономически активного населения. Максимальный возрастной предел для измерения экономически активного населения, как правило, не используется, поскольку значительное число пожилых людей старше пенсионного возраста могут регулярно или время от времени принимать участие в экономической деятельности (Организация Объединенных Наций, 1998 год, пункт 2.172).

469. Каждая страна должна определить минимальный возраст для участия в экономической деятельности. Странам, которые заинтересованы в сборе данных о детском труде, возможно, понадобится установить низкий минимальный возраст, но при этом необходимо помнить о том, что могут возникнуть некоторые искажения, если дети, которые не участвуют в экономической деятельности, ошибочно будут зарегистрированы как принимающие участие в этой деятельности. После установления минимального возраста производится редактирование экономической активности, с тем чтобы включить в таблицы данные о лицах в возрасте X лет или старше; поэтому редактирование о детях в возрасте до X лет можно производить только для того, чтобы убедиться, что все соответствующие графы не заполнены. Для того чтобы упростить составление всех таблиц, следует исключить все ответы, которые могли быть внесены для детей в возрасте до X лет.

1. Статус экономической активности (РБА)

470. Статус экономической активности составляется из нескольких экономических переменных, часть которых описывается ниже. Этих переменных может быть достаточно для сбора данных, но в целях обработки и анализа данных, возможно, потребуется изменить их классификацию.

471. «Статус текущей экономической активности» — это отношение лица к экономической деятельности, основанное на коротком учетном периоде, например, одна неделя. Использование показателя текущей экономической активности считается наиболее подходящим для стран, где на экономическую активность людей не оказывают большого влияния сезонные или другие факторы, вызывающие колебания в течение года. Этим учетным периодом в одну неделю может быть какая-то установленная фиксированная неделя в недавнем прошлом, полная последняя календарная неделя или последние семь дней до переписи (Организация Объединенных Наций, 1998 год, пункт 2.180).

472. Согласно *Принципам и рекомендациям* (Организация Объединенных Наций, 1998 год, пункт 2.182) **занятое население** включает всех лиц старше определенного возраста, которые на протяжении краткого учетного периода — одной недели либо одного дня — входили в одну из следующих категорий:

- a) занятые в сфере оплачиваемого труда:
 - i) **работающие**: лица, которые в течение учетного периода выполняли некоторую работу за заработную плату или оклад наличными или натурой;
 - ii) **имеющие работу, но не работающие**: лица, уже работавшие на своей нынешней работе, которые в течение учетного периода временно не работали, но числились официально на своем месте работы, что подтверждается, например, продолжающимся получением заработной платы/оклада, гарантией возвращения на работу по окончании непредвиденных обстоятельств, договоренностью о дате возвращения после короткого отсутствия на работе.
- b) самостоятельно занятые:
 - i) **работающие**: лица, которые в течение учетного периода выполняли некоторую работу ради получения дохода или семейной выгоды наличными или натурой;
 - ii) **имеющие собственное предприятие, но не работающие**: лица, которые имеют собственное предприятие, будь-то коммерческое предприятие, ферма или предприятие сферы услуг, и которые в течение учетного периода временно не работали по той или иной конкретной причине.

473. **Обычно экономически неактивное** население включает всех лиц, которые не отнесены ни к категории занятых, ни к категории безработных. **Обычно экономически неактивное** население следует классифицировать по следующим четырем группам (Организация Объединенных Наций, 1998 год, пункт 2.202):

- a) **учащиеся**: лица обоих полов, не относимые к **обычно экономически активному населению**, которые посещали любое обычное учебное заведение, государственное или частное, в целях систематического обучения на любом уровне в течение учетного периода;
- b) **лица, ведущие домашнее хозяйство**: лица обоих полов, не относимые к **обычно экономически активному населению**, которые выполняли домашние обязанности у себя дома, например, домашние хозяйки и другие родственники, несущие ответственность за уход за домом и детьми (однако домашняя

прислуга, работающая за плату, классифицируется как **экономически активное население**);

- c) **получатели пенсии или доходов от капитала**: лица обоих полов, не относимые к **обычно экономически активному населению**, которые получают доходы от собственности или инвестиций, проценты, арендную плату, гонорары или пенсии, за предыдущую деятельность и не могут быть классифицированы как учащиеся или лица, ведущие домашнее хозяйство;
- d) **прочие**: лица обоих полов, не относимые к **обычно экономически активному населению**, которые получают государственную помощь или частную поддержку и все другие лица, не относящиеся ни к одной из вышеуказанных категорий.

a) *Категории, относящиеся к статусу активности*

- i) *Незанятое население (P6A1) (пункты 494, 495, 496)*

474. В категорию **безработных**, согласно *Принципам и рекомендациям* (Организация Объединенных Наций, 1998 год, пункт 2.194), включаются все лица старше определенного возраста, которые в течение учетного периода отвечали следующим критериям:

- a) не имели работы: не были заняты в сфере оплачиваемого труда или не были самостоятельно заняты;
- b) в текущий момент готовы к работе: были готовы к оплачиваемому наемному труду или к самостоятельной занятости в течение учетного периода;
- c) искали работу: предпринимали конкретные шаги в определенный недавний период в поисках оплачиваемой работы по найму или самостоятельной работы. Конкретные шаги могли включать регистрацию на государственной или частной бирже труда; заявления к работодателям; регистрацию на рабочих площадках, фермах, у заводских ворот, на рынках или других людных местах; помещение объявления в газету или ответы на объявления; просьбу о содействии друзей и родственников; поиски земельного участка, здания или оборудования для создания собственного предприятия; поиски финансовых средств; и обращение за разрешением или лицензией. При классификации безработного было бы полезно проводить различие между лицами, ищущими работу впервые, и прочими ищущими работу лицами.

475. В целом для классификации в качестве безработного лицо должно удовлетворять всем трем вышеуказанным критериям. Однако в ситуациях, когда традиционные методы поиска работы имеют ограниченное

значение, рынок рабочей силы в значительной мере не организован или ограничен, когда поглощение рабочей силы на данный момент недостаточно, или когда рабочую силу в большинстве своем составляют самостоятельно занятые лица, стандартное определение незанятости может применяться при менее строгом соблюдении критерия «поиск работы». Такое смягчение критерия поиска работы предназначено в первую очередь для тех развивающихся стран, где данный критерий не применим ко всем случаям незанятости. Несмотря на это смягчение критерия поиска работы, которое в крайних случаях позволяет полностью отказаться от него, остаются применимыми два основных критерия, а именно: «не имеющие работы» и «готовые к работе в данный момент» (Организация Объединенных Наций, 1998 год, пункт 2.195).

476. Редактирование данных о незанятости — увольнение из-за отсутствия работы, поиск работы, может ли данное лицо получить работу, и последний год работы (если есть) — должно производиться одновременно. Кроме того, они должны быть сопоставимы с ответом об экономической активности и в большинстве случаев не должны указываться, если указаны проработанное время, отрасль, род занятий, категория работника и место работы. И если специалисты, занимающиеся содержанием переписи, решат, что в графе «увольнение из-за отсутствия работы» необходимо сделать запись, когда там пробел или ответ недостоверен, то тогда можно использовать матрицу интерполяции по возрасту и полу и, возможно, по образовательной подготовке данного лица.

- ii) *Поиск работы (P6A2)*

477. Редактирование показателя «поиск работы» должно производиться одновременно с редактированием показателя «увольнение из-за отсутствия работы» и «почему не ведется поиск работы». Сотрудники, занимающиеся содержанием переписи, должны разработать такие приемы редактирования, где использовались бы записи по этим элементам, с тем чтобы можно было исчислить другие элементы. При редактировании следует учитывать местные и региональные условия, а также показатели переписи или обследования.

- iii) *Экономически неактивное на данный момент население (P6A3)*

478. В категорию экономически неактивного на данный момент населения или в категорию лиц, не входящих в состав рабочей силы, относятся все лица, которые не были заняты и не были безработными в течение короткого учетного периода, используемого для оценки текущей экономической активности (Организация Объединенных Наций, 1998 год, пункт 2.205). В зависимо-

сти от причины, по которой они не являются на данный момент экономически активными, их можно классифицировать по одной из следующих групп:

- a) посещающие учебное заведение;
- b) выполняющие домашние обязанности;
- c) живущие на пенсию или доход от капитала;
- d) не работающие по иным причинам, включая нетрудоспособность или инвалидность.

О редактировании показателя «экономически неактивное на данный момент население» говорится выше в разделе о редактировании данных об экономической активности.

iv) *Почему не ведется поиск работы (P6A4)*

479. Данный элемент следует редактировать только в отношении тех, кто зарегистрировался как «не ведется поиск работы»; у всех других в этом месте должен быть пробел. В качестве альтернативы, если в графах «род занятий», «отрасль» и «статус занятости» есть достоверные записи, то следует ввести код для «имеющие работу, но не работающие». Этот код обозначает экономически активных лиц, которые имели работу, но не работали в течение учетного периода. Во всех остальных случаях, если не используется динамическая интерполяция, можно поставить «не известно». Для стран, где используется метод динамической интерполяции, можно поставить запись, применяя данные о возрасте, поле и основной активности.

480. Экономическая активность, как правило, подразделяется на следующие категории:

- 1) занятое население, работающее;
- 2) занятое население, не работающее;
- 3) самостоятельно занятые, работающие;
- 4) самостоятельно занятые, не работающие;
- 5) ищущие работу;
- 6) учащиеся;
- 7) выполняющие домашние обязанности;
- 8) живущие на пенсию или доход от капитала;
- 9) не работающие по иным причинам.

481. Для этой переменной первые четыре возможности имеются у экономически активного населения, а вторые четыре — у экономически неактивного. Работающее население (категории 1 и 3) занято, те, кто не работает (категории 2 и 4) могут быть не заняты или не входить в категорию экономически активного населения, в зависимости от ответов на вопросы о незанятости

(«увольнение из-за отсутствия работы», «поиск работы», «последний отработанный год»).

i) *незанятое население*

482. Если выбирается одна из категорий экономически активного населения (категории 1–4), следует указать переменные «отработанное время», «род занятий», «отрасль», «статус экономической активности» и «место работы». Если они не указаны, их следует отредактировать и заполнить или как «не известно» или с помощью статической или динамической интерполяции. Если выбирается категория от 1 до 4, вместо переменных «увольнение из-за отсутствия работы», «поиск работы» и «последний отработанный год» следует оставить пробелы. Если они указаны, надо изменить на ПРОБЕЛ.

ii) *экономическая активность незанятого населения*

483. Если выбирается одна из категорий для лиц из числа экономически неактивного населения (5–9), следует указать переменные «увольнение из-за отсутствия работы», «поиск работы» и «последний отработанный год». Если они достоверно не указаны, их следует отредактировать и указать как «не известно» или с помощью статической или динамической интерполяции. Если выбираются категории 5–9, вместо переменных «увольнение из-за отсутствия работы», «поиск работы» и «последний отработанный год» должны быть пробелы. Если они заполнены, следует оставить ПРОБЕЛ.

iii) *Экономическая активность учащихся и пенсионеров*

484. Если выбирается категория 6, учащиеся, сотрудники, занимающиеся содержанием переписи, должны решить, нужно ли вставить в графе «посещение учебных заведений» переменную «да, в учебном заведении». Если выбирается категория 8, пенсионер, сотрудники, занимающиеся содержанием переписи, должны решить, нужно ли устанавливать определенный возраст для выхода на пенсию.

iv) *Когда запись об экономической активности недостоверна и указаны переменные, касающиеся занятого населения*

485. Если запись об экономической активности недостоверна и если указаны некоторые переменные для отработанного времени, рода занятий, отрасли и места работы, то экономической активности респондента следует присвоить код от 1 до 4. Для выбора приемлемого ответа, возможно, понадобится матрица интерполяции.

- v) *Когда запись об экономической активности недостоверна и указаны переменные для незанятых на работе*

486. Если указаны какие-либо переменные для категорий для «увольнение из-за отсутствия работы», «поиск работы» и «последний отработанный год», то записи для экономической активности следует присвоить код от 5 до 9. Если данное лицо посещает школу, то код вероятно должен быть 6. Для лиц пожилого возраста значение кода должно быть, вероятно, 8. В других случаях специалисты, занимающиеся содержанием переписи, могут принять решение использовать матрицу интерполяции для присвоения приемлемого ответа.

- vi) *Когда запись об экономической активности недостоверна и не указаны никакие экономические переменные*

487. Если по каким-либо элементам экономической активности ответа нет, то специалисты, занимающиеся содержанием переписи, возможно, захотят использовать матрицы интерполяции для определения наиболее приемлемого ответа, а затем вычислить и другие экономические элементы.

2. Отработанное время (P6B)

488. Отработанным временем является все время, фактически затраченное на производство товаров и услуг в регулярные рабочие часы и в сверхурочное время, в течение учетного периода, принятого в переписи для определения экономической активности. В случае короткого учетного периода (например, недели, предшествующей проведению переписи), отработанное время следует измерять в часах. Если учетный период продолжительный (например, 12 месяцев, предшествующих проведению переписи), отработанное время следует измерять в неделях или днях, где это осуществимо. Отработанное время должно также включить время, затраченное на деятельность, которая хотя и не ведет непосредственно к производству товаров и услуг, все же определяется как часть производственных и служебных задач и обязанностей, например, подготовка, ремонт или содержание в порядке рабочего места или рабочих инструментов. На практике оно будет также включать время простоя в ходе этих видов деятельности, например, время, потраченное на ожидание, положение готовности к ожидаемому заданию и другие короткие перерывы. Следует исключить более длительные перерывы для приема пищи и время, когда работа не выполнялась по причине отпуска, праздничных дней, болезни или конфликтов (например, забастовок и локаутов) (Организация Объединенных Наций, 1998 год, пункт 2.210).

489. Данный элемент следует редактировать только в отношении тех, чей ответ об экономической активности был «имеет работу, работающий» или «самостоятельно занятый, работающий». Для некоторых стран отработанное время должно также включаться для лиц, ведущих домашнее хозяйство. Следует принимать категории, которые заранее определены группой редактирования. Если не применяется метод динамической интерполяции, то 0 или нецифровые коды следует заменить на «не сообщается», и специалисты, занимающиеся содержанием переписи, могут по своему усмотрению заменить переменную экономической активности на «не работает», если указанные часы равны 0.

490. В случае использования динамической интерполяции минимальные переменные для матрицы интерполяции включают в себя возрастные группы и пол, но можно использовать и другие переменные, например, основные категории образовательной подготовки, рода занятий или отрасли.

3. Род занятий (P6C)

491. Род занятий означает вид работы, выполнявшейся в течение учетного периода занятым лицом (или вид работы, выполнявшейся ранее, если речь идет о безработном), независимо от отрасли хозяйства или статуса занятости, к которым должно быть отнесено данное лицо (Организация Объединенных Наций, 1998 год, пункт 2.214).

492. Данный элемент следует редактировать только в отношении тех, чья экономическая активность указана как «занятый, работающий» или «самостоятельно занятый, работающий». Если не применяется метод динамической интерполяции, то пробелы, нулевые или недостоверные ответы следует заменить на «не сообщается».

493. Коды для рода занятий, как правило, разрабатываются таким образом, чтобы различные цифры представляли коды для основного и неосновного рода занятий. Дополнительные записи, которые почти неизбежны для категории «род занятий», сделают кодирование еще более трудным.

494. При использовании метода динамической интерполяции минимальные переменные для матрицы интерполяции включают возрастные группы и пол, но могут использоваться и другие показатели, такие как основные категории образовательной подготовки или отрасли.

4. Отрасль (P6D)

495. В соответствии с *Принципами и рекомендациями* (Организация Объединенных Наций, 1998 год, пункт

2.221) термин «отрасль» относится к деятельности предприятия, на котором занятое лицо работало в течение учетного периода, установленного для сбора данных об экономических характеристиках (или работало в последнее время, если речь идет о безработном). Для ориентации относительно выбора работы/деятельности, подлежащей классификации, см. пункт 2.212 (там же).

496. Данный элемент следует редактировать только в отношении тех, чья экономическая активность указана как «занятый, работающий» или «самостоятельно занятый, работающий». Если динамическая интерполяция не используется, то пробелы, нулевые или недостоверные ответы следует заменить на «не сообщается».

497. Коды для отрасли, как правило, разрабатываются таким образом, чтобы различные цифры представляли коды основных и неосновных отраслей. Дополнительные записи, которые почти неизбежны для этого элемента, сделают кодирование еще более трудным.

498. В случае применения метода динамической интерполяции минимальные переменные для матрицы интерполяции включают возрастные группы и пол, но могут использоваться и другие показатели, такие как основные категории образовательной подготовки или отрасли.

5. Статус занятости (P6E)

499. Термин «статус занятости» относится к положению экономически активного лица в связи с его или ее занятостью, то есть к виду прямого или косвенного договора о найме, заключенного данным лицом с другими лицами или организациями в отношении своей работы. Основные критерии, применяемые для определения таких групп классификации, — это тип экономического риска, элементом которого является степень связи лица со своим местом работы, и степень его полномочий в отношении предприятия и других работников, которую лицо имеет или будет иметь на этой работе. Следует побеспокоиться об обеспечении того, чтобы **экономически активное** лицо было классифицировано по **статусу занятости** на основе той же работы (работ), которая (которые) использовалась (использовались) для классификации лица по **роду занятий, отрасли** или **сектору** (Организация Объединенных Наций, 1998 год, пункт 2.226).

500. Экономически активное население следует классифицировать по статусу занятости (Организация Объединенных Наций, 1998 год, пункт 2.227) следующим образом:

a) наемные работники, среди которых можно различать работников со стабильными контрактами (включая штатных работников) и других работников;

b) работодатели;

c) самостоятельные работники;

d) работающие безвозмездно на семейном предприятии члены семьи;

e) члены производственных кооперативов;

f) лица, не поддающиеся классификации по статусу занятости.

501. Владельцев-управляющих инкорпорированных предприятий, которые, как правило, должны быть отнесены к числу работающих по найму, но которых в определенных описательных и аналитических целях может оказаться предпочтительнее группировать вместе с работодателями, следует выделять отдельно.

502. Данный элемент следует редактировать только в отношении тех, кто указал экономическую активность как «занятый, работающий» или «самостоятельно занятый, работающий». Если не используется динамическая интерполяция, то пробелы, нулевые или недостоверные ответы можно заменить на «не сообщается». В случае применения метода динамической интерполяции переменные для матрицы интерполяции включают возрастные группы и пол, но могут использоваться и другие показатели, такие как основные категории образовательной подготовки или отрасли.

6. Доход (P6F)

503. Вопросы переписи, связанные с экономическими характеристиками населения, представленными в первом пересмотренном издании *Принципов и рекомендаций в отношении переписи населения и жилого фонда*, касаются экономически активного населения, как оно определено в рекомендациях Международной организации труда (МОТ), где понятие экономического производства соотнесено с системой национальных счетов (СНС) (Организация Объединенных Наций, 1998 год, пункт 2.165). Экономически активное население включает всех мужчин и женщин, которые заняты или могут быть заняты в сфере производства товаров и услуг в соответствии с определением СНС в течение учетного периода времени (Организация Объединенных Наций, 1998 год, пункт 2.166). С учетом концептуального значения новых международных стандартов для характеристики экономически активного населения доход может быть определен как a) ежемесячный доход в деньгах и/или натурой от работы, выполняемой каждым экономически активным лицом или b) совокупный годовой доход домохозяйств в деньгах и/или натурой независимо от источника. Сбор достоверных данных о доходах, особенно доходов от самостоятельной деятельности и собственности, чрезвычайно труден при проведении общих опросов на местах в особенно-

сти при переписи населения. Включение дохода в денежной форме еще более осложнит проблему. Сбор данных о доходах при переписи населения, даже если он ограничивается данными о денежных доходах, вызывает особые проблемы, связанные с увеличением объема работы и, в частности, с ошибками в ответах. Исходя из этого такой вопрос, включая более широкое определение дохода, обычно считается более подходящим для использования при выборочном обследовании. Тем не менее, в зависимости от национальных потребностей страны могут пожелать иметь ограниченную информацию о денежных доходах. Собранные исходя из такого определения информация, может оказаться полезной для статистики распределения доходов, потребления и накопления в домохозяйствах в дополнение к непосредственным целям переписи (Организация Объединенных Наций, 1998 год, пункт 2.236).

504. В *Принципах и рекомендациях* выделяются два вида дохода: индивидуальный доход и доход домохозяйства. Оба элемента требуют аналогичного редактирования. Для индивидуального дохода, если не применяется метод динамической интерполяции, недостоверные ответы о доходе следует отмечать как «не указано» или «не известно». Если используется метод динамической интерполяции, то для создания матрицы интерполяции данных о доходе можно использовать показатели возраста, пол, образовательной подготовки, отрасли, род занятий и другие показатели.

505. Для этой переменной доход домохозяйства означает сумму всех доходов домохозяйства, которая вносится в запись о жилищах. Однако редактирование с применением метода динамической интерполяции происходит примерно также, когда используются не данные о каждом индивиде, а показатели возраста, пола, уровня образования главы домохозяйства.

7. Институциональный сектор (P6G)

506. Институциональный сектор занятости — юридическое оформление организаций и основные функции, поведение и цели предприятия, с которым связана работа того или иного лица (Организации Объединенных Наций, 1998 год, пункт 2.239).

507. Существует взаимосвязь между некоторыми возможными отраслями и родом занятий и институциональным сектором занятости (корпорация, государственный сектор, некоммерческие учреждения, домохозяйства и другие). Некоторые страны могут по своему усмотрению проверить эту взаимосвязь между переменными, с тем чтобы убедиться, что в таблицах не будет несоответствий, если свести эти показатели воедино.

508. При редактировании страны, не использующие

метод динамической интерполяции, должны будут обозначить институциональный сектор, если он неизвестен, как «не известно». Страны, использующие метод динамической интерполяции, должны рассмотреть возможность использовать показатель возраста и пола и, возможно, основной отрасли или рода занятий лиц с аналогичными характеристиками в данном географическом районе.

8. Место работы (P6H)

509. «Место работы» — это место, где занятое лицо в текущий момент выполняет свою работу и где занятое обычно по найму лицо выполняет основную работу, используемую для определения его/ее других экономических характеристик, таких как род занятий, отрасль и статус занятости. Информацию о месте работы можно использовать для разработки территориальных профилей занятой рабочей силы (в отличие от демографических профилей по месту жительства), однако главная цель состоит в увязке информации о месте работы с данными о месте жительства (Организация Объединенных Наций, 1998 год, пункт 2.245).

510. Поскольку показатель «место работы» используется в статистике о миграции, необходимо чтобы любые изменения в сообщаемой информации отражали конкретные географические районы. В связи с этим группа редактирования в конкретной стране может по своему усмотрению присваивать признак «не известно» для недостоверных ответов и анализировать только «известные» случаи.

511. Кодирование этого элемента потребует больше времени и будет более сложным, если будут приниматься дополнительные записи, которые должны кодироваться. Если для цифр будет установлен определенный порядок, например, первая цифра будет обозначать провинцию или область, вторая — район и т. д., то в этом случае кодирование, возможно, будет более эффективным и более точным.

512. При составлении матриц интерполяции разработчики данных должны убедиться в том, что в матрицы будут вноситься только наиболее вероятные территориальные единицы. Возможно, было бы целесообразно производить каждый раз статическую интерполяцию по каждой административно-территориальной единице или другому географическому району, чтобы убедиться, что нельзя для этого использовать предыдущие значения. В саму матрицу интерполяции можно включить показатели возраста и пола и, возможно, измененные основные категории главного рода занятий или отрасли. Кроме того, для показателя работы внутри страны и за ее пределами, возможно, понадобятся разные матрицы интерполяции.

V. РЕДАКТИРОВАНИЕ ДАННЫХ О ЖИЛОМ ФОНДЕ

513. В спецификациях для редактирования данных о жилом фонде учитывается достоверность отдельных элементов, а также согласованность между ними. Знание конкретных соотношений между элементами для данной страны позволяет планировать редактирование с точки зрения проверки на согласованность таким образом, чтобы обеспечить более высокое качество данных для сведения их в таблицы. Например, крыша жилищной единицы не должна быть из цемента, если стены построены из бамбука. И точно так же для того, чтобы в строении был туалет со смывом, ванна или душ, в нем должен быть трубопровод.

514. Как и в случае с данными о населении группа редактирования должна решить, что делать: вписывать «не указано», вычислять значение для показателя «не известно» или другое значение методом статической интерполяции (*cold deck*) или производить динамическую интерполяцию (*hot deck*) по характеристикам другой жилищной единицы. Здесь, как и во многих других случаях, предпочтительнее производить динамическую интерполяцию, поскольку она устраняет те виды исчислений, которые необходимы на стадии табулирования, когда есть информация только в самих таблицах, позволяющая принять решение о неизвестных величинах. Созданные таким образом матрицы интерполяции дают данные вместо пробелов, недостоверных величин или устраненных несоответствий, когда есть другие связанные с этим элементы с достоверными ответами. В некоторых странах могут быть различия в жилищных характеристиках по стране, но их почти нет в большинстве населенных пунктов, если их брать отдельно. В других странах могут быть существенные различия по конкретным элементам в разных населенных пунктах, особенно если брать городские и сельские районы. Эти различия необходимо учитывать при разработке матриц интерполяции и, в частности, в отношении исходных статических величин. Группа редактирования может по своему усмотрению уточнить обстоятельства, при которых следует заполнить пробел, взяв за основу предыдущую жилищную единицу с другими аналогичными характеристиками.

515. За исключением случая, когда в стране нет информации о помещениях коллективного (группового) проживания, каждому порядковому номеру должна соответствовать одна (и только одна) запись о жилище (см. глава III, Редактирование структуры). В главе о редакти-

ровании структуры излагается ряд процедур по обеспечению качества. В зависимости от решений группы редактирования программа редактирования может создать запись о жилище, если она отсутствует. Эта программа может также устранить одну или более записей, если они дублируются или повторяются несколько раз.

516. Теоретически, при редактировании записи о жилищах следует брать выборочно только допустимые элементы. Редактируемые элементы могут отличаться в зависимости от городских/сельских, климатических и иных условий. Однако на практике лишь у нескольких стран есть время и опыт разработки и применения многомерных матриц для изменения недостающих или несогласующихся данных. Еще меньше стран фактически применяют выборочное редактирование.

517. Информация, собираемая с помощью переписного листа, будет также зависеть от типа жилых помещений (жилищная единица или помещения для проживания группами) и от того, свободна она или занята. Редактирование данных о жилых помещениях для коллективного или группового проживания может ограничиваться только теми элементами данных, которые собраны в помещениях для проживания группами или иных жилищных единицах.

518. По определению бездомные не могут иметь записей о жилище. Если, тем не менее, такие записи есть, поскольку страна предпочитает иметь для них идентификаторы, то их можно обрабатывать так же, как и записи для помещений коллективного проживания, или надо будет применять совсем другие приемы редактирования или не применять их вообще.

519. Иногда для конкретного элемента можно допустить запись «не сообщается». Это может потребоваться тогда, когда у группы редактирования в данной стране нет хорошей основы для исчисления ответов по той или иной характеристике. Решение оставить запись «не сообщается» должно соотноситься с требованием выдавать соответствующие сведенные в таблицы характеристики в целях планирования и осуществления политики. Если планирующим органам необходима выборочная информация, поскольку признаки «не сообщается» распределяются так же, как и регистрируемые данные, то распределение признака «не сообщается» не должно вызывать никаких проблем. Однако если запись «не сообщается» вносит какие-то искажения, то интерпо-

ляция после сбора данных может быть проблематичной, особенно для небольших районов или конкретных условий. Например, если респонденты, занимающие по принятому в стране определению «нестандартное» жилье, отказываются раскрыть некоторые из характеристик этого жилья и счетчик их не регистрирует, планирующие органы могут не иметь возможности ввести программы по исправлению положения, чтобы облегчить нестандартные условия.

520. Редактировать данные о жилье, как правило, проще, чем демографические данные, поскольку сводные таблицы обычно намного проще. В большинстве стран отдельные жилищные характеристики собираются только по различным территориальным уровням. Как говорилось выше, страны, предпочитающие не использовать метод динамической интерполяции, должны определить идентификатор для категорий «не известно», чтобы использовать его, когда возникают недостоверные или несогласующиеся ответы.

521. В странах, где используется метод динамической интерполяции, группа редактирования должна разработать простые матрицы интерполяции с измерениями, где проводятся различия между жилищными характеристиками. Для большинства стран показатель «тип жилых помещений», будь то жилищная единица или помещения для коллективного проживания, включая тип единицы в рамках этих категорий, является наилучшим первичным показателем для динамической интерполяции.

522. Для некоторых стран территориальные единицы можно использовать как одно измерение этих матриц интерполяции. Можно использовать и показатель вида владения. Например, если в данной стране около половины единиц арендуется, а другие находятся в собственности, вид владения является удобным показателем для включения в качестве одного из измерений матрицы интерполяции. Однако если арендуется только около 5 процентов единиц, то более приемлемы будут какие-то другие характеристики. Вид владения часто является полезным показателем, который можно использовать в матрицах интерполяции, особенно в странах с большой долей основных видов владения. К другим заслуживающим внимания характеристикам относится тип стен и наличие электричества.

523. Для каждой страны конкретные переменные, включенные в качестве измерений в матрицы интерполяции, должны соответствовать показателям в наборе данных, поэтому в связи с элементами данных о жилье необходимо следить за тем, чтобы в отдельных элементах, а также в комбинациях элементов проводилось различие в количестве характеристик. В *Принципах и рекомендациях* проводится различие между основными и дополнительными элементами; в разделе А, ниже,

посвященном редактированию, выделяются основные элементы для включения в матрицу.

А. ОСНОВНЫЕ ВОПРОСЫ

524. Единицами учета при проведении переписи жилого фонда являются *a)* строения; *b)* жилые помещения и *c)* проживающие в них члены домохозяйств. В Организации Объединенных Наций был разработан список из двадцати основных вопросов, представляющих общий интерес и ценность, а также имеющих значение с точки зрения проведения статистических сопоставлений. Для удобства пользователей предлагаемые коды редактирования для этих и ряда других дополнительных вопросов приводятся ниже. Вопросы сгруппированы по типу единиц учета.

1. Строение: описание строения (H01)

525. В *Принципах и рекомендациях* (Организация Объединенных Наций, 1998 год, пункт 2.299) для строений, в которых хотя бы часть площади используется как жилье, рекомендуется применять следующую классификацию:

1. строения, имеющие одну жилищную единицу;
 - 1.1. отдельные;
 - 1.2. примыкающие;
2. строения, имеющие более одной жилищной единицы;
 - 2.1. до двух этажей;
 - 2.2. от трех до десяти этажей;
 - 2.3. одиннадцать и более этажей;
3. строения для лиц, проживающих в специализированных учреждениях.

526. В случае недостоверности показателя, если не применяется динамическая интерполяция, следует присвоить значение «не известно». Страны, применяющие динамическую интерполяцию для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, например, строительный материал для наружных стен, период строительства и/или тип жилищных единиц в строении (Организация Объединенных Наций, 1998 год, пункт 2.300), для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

2. Строение: строительный материал наружных стен (H02)

527. Этот вопрос касается строительного материала, из которого построены наружные (внешние) стены строения, где расположены группы жилых помещений. Если

стены построены из нескольких материалов, следует указать преобладающий материал. Выбор материалов (например, кирпич, бетон, дерево, необожженный кирпич) будет зависеть от тех материалов, которые наиболее часто используются в строительстве в данной стране и от их значения с точки зрения долговечности постройки или оценки ее прочности (Организация Объединенных Наций, 1998 год, пункт 2.304).

528. В случае недостоверности показателя, если не применяется динамическая интерполяция, следует присвоить значение «не известно». Страны, использующие динамическую интерполяцию для исчисления недостоверных показателей должны использовать, по меньшей мере, две характеристики, например, период строительства и/или тип жилищных единиц в строении, для того чтобы получить «известную» информацию по аналогичным жилищным единицам в данном географическом районе.

3. *Строение: год или период постройки (H03)*

529. Этот вопрос относится к возрасту строения, в котором находятся группы жилых помещений. Рекомендуется точно указывать год постройки для строений, возведенных в период между предыдущей и проводимой переписями, если этот период не превышает десяти лет. Если период между переписями превышает десять лет или если в данном районе еще никогда не проводилась перепись, точный год постройки нужно указать для строений, возведенных в течение предшествующих десяти лет. Что касается ранее построенных домов, то сведения следует собирать за периоды, которые дадут возможность определить истекший срок службы имеющегося жилого фонда. При сборе данных по этому вопросу могут возникнуть трудности, поскольку в некоторых случаях жильцы могут не знать года постройки строения (Организация Объединенных Наций, 1998 год, пункт 2.307).

530. Некоторые страны, даже используя динамическую интерполяцию, принимают ответ «не известно» для элемента, указывающего на год или период постройки. В этом случае страна может и не использовать динамическую интерполяцию для этого элемента, даже если используются матрицы интерполяции для других переменных. В такой ситуации, если показатель недостоверен, вероятно, следует все же присваивать значение «не известно», когда динамическая интерполяция не применяется. Страны, применяющие метод динамической интерполяции для недостоверных величин должны использовать, по меньшей мере, две характеристики, в том числе тип строения, строительный материал наружных стен и/или тип жилищных единиц строений для получения «известной» информации, ис-

ходя из аналогичных жилищных единиц в данном географическом районе.

4. *Жилые помещения:*

местонахождение жилых помещений (H04)

531. Местонахождение жилых помещений — это территориальный показатель, о котором говорится в главе III, посвященной редактированию структуры.

5. *Жилые помещения:*

тип жилых помещений (H05)

532. В изложенной ниже классификации описывается система трехзначных кодов, разработанных Организацией Объединенных Наций (1998 год, пункты 2.328–2.365) для группирования в широких категориях жилищных единиц и жилых помещений коллективного проживания со сходными характеристиками их структуры. Распределение жильцов (населения) по различным группам дает ценную информацию относительно имеющихся во время проведения переписи жилых помещений. Эта классификация дает также полезную основу для стратификации при проведении выборочных обследований. Жилые помещения можно подразделить на следующие категории:

1. жилищные единицы;
 - 1.1. обычные жилища;
 - 1.2. основные жилища;
 - 1.3. временные жилищные единицы;
 - 1.4. передвижные жилищные единицы;
 - 1.5. маргинальные жилищные единицы;
 - 1.5.1. импровизированные жилищные единицы;
 - 1.5.2. жилищные единицы в капитальных строениях, не предназначенных для жилья;
 - 1.5.3. другие помещения, не предназначенные для жилья;
2. жилые помещения коллективного проживания;
 - 2.1. гостиницы, пансионаты и другие сдаваемые внаем меблированные помещения;
 - 2.2. специализированные учреждения;
 - 2.3. лагеря;
 - 2.4. прочее.

533. Группам редактирования необходимо разработать такие методики редактирования, которые позволяют обеспечить внутреннюю согласованность информации, касающейся всех жилых помещений коллективного

проживания и жилищных единиц. Если показатель типа жилья не известен или недостоверен, то группе редактирования следует применить такую методику, которая позволяет использовать другие показатели, чтобы определить тип жилых помещений. В противном случае, когда показатель недостоверен, если не применяется динамическая интерполяция, следует присвоить значение «не известно». Национальным статистическим органам/бюро переписей применяющим динамическую интерполяцию для исчисления недостоверных показателей, необходимо использовать, по меньшей мере, две характеристики, например, тип строения, форма владения, количество комнат, размер площади или сведения о занятости помещений, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

6. Жилые помещения: сведения о занятости (H06)

534. Решение учитывать жилые помещения, жильцы которых временно отсутствуют или временно присутствуют, в качестве «занятых» или «незанятых» придется рассматривать в зависимости от того, проводится ли перепись населения де-юре или де-факто. И в том и в другом случае было бы полезно отличать по мере возможности жилые помещения, которые используются в качестве основного места жительства от жилых помещений, использующихся в качестве второго места жительства. Это особенно важно, если второе жилище значительно отличается по своим характеристикам от основного жилья, что, например, относится к тем случаям, когда семьи, проживающие в сельских районах, переезжают в определенные периоды года из своих постоянных сельских жилых помещений в простейшие строения, размещенные на сельскохозяйственных угодьях (Организация Объединенных Наций, 1998 год, пункт 2.369). Ниже приводится рекомендуемая классификация обычного и основного жилья:

1. занятое;
2. незанятое;
 - 2.1. не занятое на сезонной основе;
 - 2.2. не занятое на несезонной основе;
 - 2.2.1. сдается внаем;
 - 2.2.2. продается;
 - 2.2.3. предназначено на снос;
 - 2.2.4. прочее.

535. Если жилищная единица занята, то количество жильцов (код H17) и подсчет демографических записей не должны равняться нулю. Если никаких проживающих не зарегистрировано, то данная единица или свободна, или проживающие в ней отсутствуют. Как отмечалось

ранее, в разделе о структурном редактировании специалисты должны разработать соответствующие процедуры для того, чтобы определить, свободна ли данная единица. Если она зарегистрирована как занятая, но фактически свободна, то необходимо разработать какой-то метод для определения типа незанятости, указав «не известно» или используя метод динамической интерполяции. Если данная единица зарегистрирована как свободная, но можно определить, что она фактически занята, по имеющейся информации о количестве жильцов или подсчете демографических записей, то показатель занятости необходимо поменять на «занято».

536. Если показатель недостоверен, количество жильцов равно нулю и не представлено никаких демографических записей, то, если не используется динамическая интерполяция, следует приписать «свободно». Если показатель недостоверен, но количество жильцов не равно нулю или имеются демографические записи, следует указать «занято». Страны, применяющие метод динамической интерполяции для недостоверных величин (для исчисления типа «незанятости»), должны использовать, по меньшей мере, две характеристики для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе или в качестве альтернативы указать «не занято по неизвестной причине».

7. Жилые помещения: форма владения (H07)

537. Данный вопрос относится к форме владения самими жилыми помещениями, а не к земельному участку, на котором они расположены (Организация Объединенных Наций, 1998 год, пункт 2.370). Форму владения не следует смешивать с формой пользования. По этому вопросу должна быть получена следующая информация:

- a) принадлежат ли данные жилые помещения государственному сектору (центральное правительство, местные органы власти, государственные корпорации);
- b) являются ли данные жилые помещения частной собственностью (домохозяйств, частных корпораций, кооперативов, жилищных объединений или иных субъектов). Вопрос иногда ставится более широко, чтобы указать, полностью ли оплачено жилое помещение, производится ли его покупка в рассрочку или не взят ли под него залог. Классификация жилых помещений выглядит следующим образом:
 1. занимаемые владельцем;
 2. занимаемые другими лицами;
 - 2.1. в собственности государства;
 - 2.2. в частной собственности;
 - 2.3. прочее.

538. Если владение связано с пользованием, то это следует учитывать при разработке программы редактирования; если связи между ними нет, то тогда, вероятно, форма владения не будет зависеть от других жилищных переменных. В случае недостоверности показателя «форма владения», если не применяется динамическая интерполяция, следует присвоить значение «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, например, строительный материал стен, пользование жильем, тип жилищной единицы и количество комнат, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

8. Жилые помещения: число комнат (H08)

539. Комната определяется как площадь в пределах жилищной единицы или других жилых помещениях, окруженная от пола до потолка или до крыши или высотой не менее двух метров, достаточная для размещения одной кровати для взрослого человека, то есть не менее четырех квадратных метров. Таким образом, комнатами можно считать спальни, столовые, гостиные, кабинеты, пригодные для жилья мансарды, помещения для прислуги, кухни, помещения, используемые для профессиональных или деловых целей, а также другого рода отдельные помещения, используемые или предназначенные для жилья и отвечающие критериям в отношении стен и метража. Коридоры, веранды, вестибюли, ванны и туалеты не должны считаться комнатами, даже если они и отвечают этим критериям. Для целей национальной статистики можно отдельно собирать сведения по помещениям менее четырех квадратных метров, которые в других отношениях подходят под определение комнаты, если считается, что их число оправдывает сбор таких данных (Организация Объединенных Наций, 1998 год, пункт 2.375).

540. Поскольку число комнат может не зависеть от других жилищных переменных, в случае недостоверности показателя, если не используется динамическая интерполяция, то следует присвоить значение «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных величин, должны использовать, по меньшей мере, две характеристики, например, вид жилищной единицы, строительный материал стен, форма пользования жильем и незанятость, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

9. Жилые помещения: площадь помещения (H09)

541. Жилая площадь — это полезная площадь в жилищных единицах, то есть площадь, измеряемая в пре-

делах наружных стен жилищных единиц, исключая нежилые подвалы и чердаки. В многоквартирных зданиях все площади общего пользования должны исключаться. Подход к жилищным единицам должен отличаться от подхода к жилым помещениям коллективного проживания (Организация Объединенных Наций, 1998 год, пункт 2.378).

542. Жилая площадь может быть связана с числом комнат и/или числом спальных комнат, поэтому группа редактирования в конкретной стране, возможно, захочет принять это во внимание при разработке программы редактирования. Или же жилая площадь может не зависеть от других жилищных переменных, подлежащих редактированию. Возможно, необходимо будет указать единицу измерения, например, квадратные метры. В случае недостоверности показателя, если не используется динамическая интерполяция, следует присвоить значение «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных величин, должны использовать, по меньшей мере, две характеристики, например, тип жилищной единицы, строительный материал стен, форма пользования и незанятость, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

10. Жилые помещения: система водоснабжения (H10)⁶

543. Согласно *Принципам и рекомендациям* (Организация Объединенных Наций, 1998 год, пункт 2.381) основные сведения о системе водоснабжения, которые следует собирать во время переписи, касаются того, обеспечиваются ли жилищные единицы водопроводной водой. Из этих сведений будет видно, обеспечиваются ли жилые помещения водой из коммунальной системы водоснабжения или из индивидуальной установки, например, водонапорного бака или насоса. Единицей учета при этом является жилищная единица. Необходимо также указать, имеется ли водонапорный кран внутри жилищной единицы, а если нет, то находится ли он не далее определенного расстояния от входа. Рекомендуемое расстояние составляет 200 метров, исходя из того что в этих пределах жильцы смогут брать воду для бытовых нужд без чрезмерных усилий. Помимо месторасположения водопроводного крана особого внимания заслуживает также источник получаемой воды. В связи с этим с точки зрения водоснабжения рекомендуется следующая классификация жилищных единиц:

⁶ Жилищная единица фактически берется за единицу учета для следующих переменных: система водоснабжения, туалет и канализационное оборудование, ванное оборудование, кухонное оборудование, освещение и удаление мусора.

1. водопровод внутри жилищной единицы;
 - 1.1. от коммунальной системы водоснабжения;
 - 1.2. из частного источника;
2. водопровод вне жилищной единицы, но не далее 200 метров от нее;
 - 2.1. от коммунальной системы водоснабжения;
 - 2.1.1. только для индивидуального пользования;
 - 2.1.2. для совместного пользования;
 - 2.2. из частного источника;
 - 2.2.1. только для индивидуального пользования;
 - 2.2.2. для совместного пользования;
3. Отсутствие водопроводной воды (даже за пределами 200 метров от жилых помещений).

544. Коммунальная система водоснабжения представляет собой техническую службу, которая управляется и инспектируется государственными органами. Эксплуатантом таких систем является, как правило, государственный или муниципальный орган, но в некоторых случаях они принадлежат кооперативным или частным предприятиям.

545. Элементы, касающиеся водохозяйственных установок: системы водоснабжения, туалетов и канализационного оборудования, ванного оборудования и горячего водоснабжения, вероятно, следует редактировать вместе. Поскольку эти элементы тесно связаны между собой, если одного из них недостает или он указан неправильно, то другие можно использовать для исчисления недостающего показателя. В отношении территориальных районов, где нет проточной воды, специалисты, возможно, должны использовать специальные приемы редактирования данных о жилищных единицах. Если же в данном районе есть другие единицы с аналогичными характеристиками, то эти элементы рекомендуется использовать для динамической интерполяции, если таковая производится.

546. В случае недостоверности показателя системы водоснабжения, если не применяется метод динамической интерполяции следует приписывать «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей должны использовать, по меньшей мере, две характеристики. Это, как правило, тип жилищной единицы, туалет и канализационное оборудование, а также ванное оборудование. Эти данные можно использовать для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

11. Жилые помещения: туалет и канализационное оборудование (H11)

547. Некоторые страны сочли целесообразным расширить данную классификацию, включив в нее туалеты без слива, чтобы выделить некоторые широко используемые типы туалетов и указать на определенный уровень санитарных условий. Организация Объединенных Наций (1998 год, пункт 2.386) рекомендует следующую классификацию:

1. с туалетом внутри жилищной единицы;
 - 1.1. смывной туалет;
 - 1.2. несмывной туалет;
2. с туалетом за пределами жилищной единицы;
 - 2.1. смывной туалет;
 - 2.1.1. только для жильцов данного жилого помещения;
 - 2.1.2. для совместного пользования;
 - 2.2. несмывной туалет;
 - 2.2.1. только для жильцов данного жилого помещения;
 - 2.2.2. для совместного пользования;
3. отсутствие какого-либо туалета.

548. Данные о типе сантехнического канализационного оборудования — это еще один элемент, касающийся водоснабжения в жилищных единицах, и их следует редактировать вместе с другими элементами, связанными с водоснабжением. Такие показатели, как «для личного пользования», «для совместного пользования», «только для жильцов данного жилого помещения» и т. д., можно использовать при определении согласованности данных. Если они не согласуются, то их можно использовать для того, чтобы установить, какой путь редактирования избрать для выявления проблемы. Когда есть одна или более других связанных с водоснабжением переменных, то оценку неизвестной или несогласующейся информации можно произвести, не прибегая к показателю «неизвестно» или к динамической интерполяции. Однако если это не дает достоверного значения и если не используется динамическая интерполяция, следует приписывать «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, в том числе, как правило, тип жилищной единицы, а также водоснабжение, строительный материал стен, форма пользования и незанятость, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

12. Жилые помещения: ванное оборудование (H12)

549. Согласно *Принципам и рекомендациям* (Организация Объединенных Наций, 1998 год, пункт 2.390) необходимо получить сведения о том, установлена ли в пределах каждого комплекса жилых помещений стационарная ванна или душевая установка. Единицей учета для этого является жилищная единица. Могут быть собраны дополнительные сведения о том, находятся ли эти виды удобств в исключительном пользовании жильцов данной жилищной единицы или нет, и имеется ли для ванны горячая или только холодная вода. Однако в некоторых районах мира предложенная выше классификация может не отвечать национальным потребностям. Вместо этого может быть важным, например, провести разграничение между жилыми помещениями с отдельной ванной комнатой, строениями с открытыми кабинами для мытья и использованием общественными банями. Ниже приводится следующая рекомендуемая классификация жилищных единиц по наличию и типу ванного оборудования:

1. жилищная единица со стационарной ванной или душевой установкой;
2. без стационарной ванны или душевой установки;
 - 2.1. наличие стационарной ванны или душевой установки в нежилищной единице;
 - 2.1.1. только для пользования данной жилищной единицы;
 - 2.1.2. для совместного пользования;
 - 2.2. отсутствие стационарной ванны или душевой установки.

550. Данные о типах данного оборудования следует редактировать вместе с другими элементами, касающимися водоснабжения. Такие показатели, как «для личного пользования», «для совместного пользования» или «исключительного пользования», можно использовать для определения согласованности показателей, и, если они не согласуются, для того чтобы установить, какой путь редактирования избрать для выявления проблемы. Если есть одна или более переменных, касающихся водоснабжения, оценку неизвестной или несогласующейся информации можно производить, не прибегая к значению «неизвестной» или к «динамической» интерполяции. Однако если никакие приемы не помогают и если не используется динамическая интерполяция, в случае недоверности показателя следует приписывать «неизвестно». Страны, применяющие метод динамической интерполяции для исчисления недоверных величин, должны использовать, по меньшей мере, две характеристики. Как правило, это жилищная единица, водоснабжение, строительный материал стен, форма пользования или занятость, которые использу-

ются для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

13. Жилые помещения: кухонное оборудование (H13)

551. Согласно *Принципам и рекомендациям* (Организация Объединенных Наций, 1998 год, пункт 2.395) сбор данных о наличии кухни может предоставить удобную возможность и для сбора сведений о типе оборудования, используемого для приготовления пищи, например, кухонной плиты, электрической плитки или очага, а также о наличии кухонной раковины и места хранения продуктов, обеспечивающего их сохранность. Ниже приводится рекомендуемая классификация жилищных единиц по наличию кухни и других мест, отведенных для приготовления пищи:

1. жилищная единица, включающая кухню;
2. жилищная единица, включающая другое место для приготовления пищи;
3. жилищная единица без кухни или другого места для приготовления пищи;
 - 3.1. для приготовления пищи вне жилищной единицы;
 - 3.1.1. только для пользования данной жилищной единицы;
 - 3.1.2. для совместного пользования;
 - 3.2. отсутствие кухни или другого места для приготовления пищи.

552. При редактировании данных о кухонном оборудовании такие показатели, как «для личного пользования», «для совместного пользования», «для исключительного пользования» и т. д., можно использовать при определении согласованности показателей, и, если они не согласуются, для установления того, какой путь редактирования избрать для выявления проблемы. Когда есть одна или две переменных, касающихся кухонного оборудования, оценку неизвестной или несогласующейся информации можно производить не прибегая к значению «не известно» или к динамической интерполяции. Однако в случае недоверности показателя, если динамическая интерполяция не используется, следует приписывать «не известно». Страны, применяющие метод динамической интерполяции для исчисления недоверных показателей, должны использовать, по меньшей мере, две характеристики, которые, как правило, включают в себя тип жилищной единицы и еще водоснабжение, строительный материал стен, форма пользования и занятость, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

14. Жилые помещения: освещение (H14)

553. Следует собирать сведения о типе освещения жилых помещений — электричество, газ, керосиновые лампы или какие-либо другие источники. Если имеется электрическое освещение, то некоторые страны могут пожелать собрать сведения о том, поступает ли электроэнергия из коммунальной сети, с электростанции или из какого-либо иного источника, например от промышленного предприятия или горнодобывающего предприятия. Помимо типа освещения странам следует оценивать информацию о наличии электричества для иных целей (таких, как приготовление пищи, нагрев воды и отопление помещений). Если жилищные условия в стране дают возможность получать информацию о типе освещения, то в дополнительных обследованиях необходимости не возникнет (Организация Объединенных Наций, 1998 год, пункт 2.398).

554. В случае недостоверности показателя, если не используется динамическая интерполяция, следует присписывать значения «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных величин, должны использовать, по меньшей мере, две характеристики, включающие, как правило, тип жилищной единицы и затем электричество, строительный материал стен, форма пользования и незанятость, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

15. Жилые помещения: удаление мусора (H15)

555. Согласно *Принципам и рекомендациям* (Организация Объединенных Наций, 1998 год, пункт 2.401) этот вопрос касается сбора и удаления твердых бытовых отходов. Единицей учета здесь является жилищная единица. Ниже приводятся основные принципы классификации жилищных единиц по типу твердых отходов:

1. твердые отходы, собираемые на регулярной основе специально назначенными организациями или лицами;
2. твердые отходы, собираемые на нерегулярной основе специально назначенными организациями или лицами;
3. твердые отходы, собираемые организациями или лицами, взявшими на себя эту задачу;
4. жильцы выбрасывают твердые отходы на местную свалку, за состоянием которой следят власти;
5. жильцы выбрасывают твердые отходы на местную свалку, за состоянием которой власти не следят;

6. другие способы (включая сжигание отходов жильцами).

556. Данные о твердых отходах не зависят от других жилищных переменных. В случае недостоверности показателя, если не используется динамическая интерполяция, следует присваивать значение «не известно». Страны, применяющие метод динамической интерполяции для исчисления достоверных показателей, должны использовать, по меньшей мере, две характеристики. Как правило, они включают в себя тип жилищной единицы и затем строительный материал стен, форма пользования, незанятость или кухонное оборудование, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

16. Жилые помещения: занятость более чем одним домохозяйством (H16)

557. Занятость более чем одним домохозяйством не зависит от других жилищных данных. Если значение недостоверно, страна должна считать число глав домохозяйства и использовать это число.

17. Жилые помещения: число жильцов (H17)

558. Каждое лицо, обычно проживающее в какой-либо жилищной единице или в комплексе жилых помещений коллективного проживания, следует регистрировать как жильца. Исходя из этого, счетной единицей здесь является жилое помещение. Однако, поскольку переписи жилого фонда обычно проводятся с переписями населения, применимость этого определения зависит от того, содержат ли собранные регистрационные данные переписи населения указания на то, где именно каждое лицо находилось в день переписи, или указание на обычное место проживания этого лица. В отношении тех, кто занимает передвижные жилищные единицы, такие как лодки, фургоны, автоприцепы, следует четко разграничивать лиц, занимающих их в качестве жилого помещения, и лиц, использующих подобные единицы как средство транспорта (Организация Объединенных Наций, 1998 год, пункт 2.407).

559. «Число жильцов» связано с количеством демографических записей, и оба показателя должны быть идентичны. Если это не так, то необходимо внести коррективы в «число жильцов» или в «число демографических записей». Как правило, корректируется «число жильцов», с тем чтобы оно равнялось числу лиц в данной жилищной единице. Этот элемент не следует ни отмечать как «не известно», ни исчислять методом интерполяции.

18. *Жильцы: характеристики главы домохозяйства (H18)*

560. Характеристики главы домохозяйства берутся, как правило, из демографических записей, с тем чтобы легче было сводить информацию в таблице в целях планирования и проведения анализа. Эти данные, включая этническое происхождение, вероисповедание или доход, помогают определить различие в социальном статусе или потребностях. Поскольку эти характеристики уже редактировались как элементы переписи населения, то никакого дополнительного редактирования здесь не требуется.

19. *Жильцы: пользование жильем (H19)*

561. Согласно *Принципам и рекомендациям* (Организация Объединенных Наций, 1998 год, пункт 2.410) пользование определяется условиями, на которых домохозяйство занимает всю или часть жилищной единицы. Единицей учета здесь является домохозяйство, занимающее жилищную единицу. По форме пользования домохозяйства классифицируются следующим образом:

1. член домохозяйства является владельцем жилищной единицы;
2. член домохозяйства арендует всю жилищную единицу или ее часть;
 - 2.1. член домохозяйства является главным арендатором жилищной единицы или ее части;
 - 2.2. член домохозяйства является субарендатором жилищной единицы;
3. прочие формы.

Такую единицу, занятую без оплаты в денежной форме, разрешения или без разрешения владельца, особенно там, где такая практика распространена, следует учитывать отдельно.

562. Форма пользования может соотноситься с типом владения (H12), поэтому группа редактирования при составлении программы редактирования, возможно, должна будет учесть соотношение между этими двумя элементами. В случае недостоверности показателя формы пользования, если не используется динамическая интерполяция, следует присваивать значение «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, такие, как тип жилищной единицы, арендная плата и незанятость, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

20. *Жильцы: арендная плата за жилье и жилищные расходы владельца (H20)*

563. Данные об арендной плате за жилье и жилищные расходы владельца не зависят от других жилищных переменных за исключением, пожалуй, того, что расходы на арендную плату должны указываться только для сдаваемых в аренду единиц, а расходы владельца — только для единиц, занимаемых самим владельцем. Группа редактирования должна в каждом конкретном случае определять наиболее приемлемые взаимосвязи между этими показателями. В случае недостоверности показателя, если не используется динамическая интерполяция, следует присваивать значение «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных величин, должны использовать, по меньшей мере, две характеристики, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

В. ДОПОЛНИТЕЛЬНЫЕ ВОПРОСЫ

564. В данном разделе указывается ряд дополнительных вопросов, которые могут оказаться полезными для многих стран при планировании ими своих национальных переписей и обследований. Кроме того, вопросы сгруппированы по типу единицы учета (строения, жилые помещения и жильцы) и приводятся предлагаемые коды редактирования (Организация Объединенных Наций, 1998 год, пункты 2.416–432).

1. *Строение: число жилищ (A01)*

565. Схема редактирования данных о количестве жилищных единиц в строении разъясняется в пунктах 219–220 главы III в рамках редактирования структуры.

2. *Строение: наличие лифта (A02)*

566. Этот вопрос касается наличия лифта (поднимающейся или опускающейся платформы для транспортировки людей и грузов) в многоэтажном строении. Сведения собираются о лифте, работающем в течение большей части времени: иными словами о лифте, который функционирует большую часть времени при условии его регулярного ремонта (Организация Объединенных Наций, 1998 год, пункт 2.419).

567. Если строение одноэтажное или это одиночная обособленная единица, то лифт указываться не должен. Если все же он указан, группа редактирования должна решить, чему отдать предпочтение: числу этажей или

тому факту, что там есть лифт. Если предпочтение отдается лифту, то необходимо изменить число этажей, указав «не известно», или используя метод динамической интерполяции для получения другого показателя. Если предпочтение отдается числу этажей и строение имеет всего один этаж, то ответ на вопрос о наличии лифта нужно изменить на «нет».

568. При наличии лифта, который требует электроэнергии, необходимо проверить, электрифицировано ли данное строение.

569. И, наконец, в случае недостоверности данных о лифте, если не используется динамическая интерполяция, следует присвоить значение «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей должны использовать, по меньшей мере, две характеристики, такие как тип строения и строительный материал наружных стен для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

3. *Строение: сельскохозяйственное назначение (A03)*

570. Ряд стран сочли необходимым уточнять при проведении своих национальных переписей, имеют ли регистрируемые строения сельскохозяйственное назначение или нет. Строение сельскохозяйственного назначения представляет собой часть сельскохозяйственного владения и используется для сельскохозяйственных целей и проживания (Организация Объединенных Наций, 1998 год, пункт 2.420).

571. Данные о строениях сельскохозяйственного назначения не зависят от других жилищных показателей. Страны могут по своему усмотрению проверить их на соответствие с демографическими данными о роде занятий и отрасли. В случае недостоверности показателя, если не используется динамическая интерполяция, следует присписать значение «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

4. *Строение: строительный материал крыши (A04a)*

572. В некоторых случаях материалы, применяемые для крыш и полов, могут представлять особый интерес и использоваться для дополнительной оценки качества жилищ в строении. Этот вопрос касается материалов, используемых для крыш и/или пола (хотя в зависимости от конкретных потребностей страны, он может также относиться и к другим частям строения, таким как

каркас или фундамент). Единицей учета является строение. Регистрируются только наиболее распространенные материалы, и что касается крыши, то это может быть черепица, бетон, кровельное железо, пальмовые листья, солома, бамбук или аналогичные растительные материалы, а также грунт, пластиковое покрытие или некоторые другие материалы (Организация Объединенных Наций, 1998 год, пункт 2.421).

573. Иногда ответ на вопрос о строительном материале наружных стен не согласуется с ответом на вопрос о строительном материале для крыши; это может произойти, например, если строительный материал, определенный для стен, недостаточно прочный для того, чтобы поддерживать крышу. Как отмечалось выше, в этом случае специалисты должны решить, какой из двух показателей изменить и использовать значение «не известно». В случае недостоверности показателя, если не используется динамическая интерполяция, следует указывать «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, например, тип строения, тип жилищной единицы, строительный материал стен, форма пользования и незанятость, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

5. *Строение: строительный материал для пола (A046b)*

574. Данные о строительном материале для пола могут согласовываться или не согласовываться с данными о строительных материалах для крыш и стен. Если группа редактирования в стране обнаруживает несогласованность или недостоверные комбинации, она должна решить, приписывать ли значение «не известно» или использовать матрицы интерполяции для изменения одного или нескольких ответов. В случае недостоверности показателя, если не используется динамическая интерполяция, следует указывать «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, например, тип строения, строительный материал наружных стен, тип жилищной единицы, форма пользования и незанятость для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

6. *Строение: эксплуатационное состояние (A05)*

575. Данный вопрос касается того, нуждается ли строение в ремонте, или характера необходимого ремонта. Единицей учета является строение. Классифи-

кация строений по их эксплуатационному состоянию может включать такие показатели, как «ремонт не требуется», «нужен незначительный ремонт», «нужен средний ремонт», «нужен серьезный ремонт» и «ремонту не подлежит». Незначительный ремонт относится, главным образом, к регулярному поддержанию эксплуатационного состояния строения и его компонентов, например, замена разбитых стекол. Средний ремонт относится к устранению более серьезных дефектов, таких как недостающие водосточные желоба на крыше, большие участки осыпавшейся штукатурки или отсутствие надежных перил у лестниц. Серьезный ремонт необходим в случае значительных структурных дефектов строения, таких как отсутствие дранки или черепицы на крыше, трещины и отверстия в наружных стенах или недостающие ступеньки. Термин «ремонту не подлежит» указывает на то, что строение отремонтировать невозможно: это означает, что у строения так много серьезных структурных дефектов, что представляется более целесообразным его снести, чем ремонтировать. Чаще всего этот термин относится к строениям, от которых остался только каркас с полуразрушенными наружными стенами и/или крышей (Организация Объединенных Наций, 1998 год, пункт 2.422).

576. Данные об эксплуатационном состоянии строения не зависят от других жилищных показателей. Следовательно, в случае недостоверности показаний, если не используется динамическая интерполяция, следует указывать «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, например, тип строения, строительный материал наружных стен и тип жилищной единицы, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

7. Жилые помещения: число спален (A06)

577. Помимо регистрации числа комнат, в ряде национальных переписей собирается информация о числе спален в жилищной единице, которая является единицей учета по данному вопросу. Спальня определяется как комната с кроватью, используемая для ночного отдыха (Организация Объединенных Наций, 1998 год, пункт 2.423).

578. Иногда счетчики указывают число спален, которое больше, чем число комнат⁷. В этом случае, если страна использует значение «не указано» только для

⁷ Если указаны и комнаты, и спальни, то эти данные следует редактировать вместе, и число спален не должно превышать число комнат. Поскольку число спален — это «дополнительный» вопрос, то редактирование производится, если есть оба показателя.

недостоверных или несогласующихся ответов, то число спален следует обозначать как «не указано». Если используется динамическая интерполяция, спальню следует оценивать с помощью матрицы интерполяции, где в качестве одного из элементов стоит число комнат. Таким образом, число спален будет не больше, чем число комнат, потому что показатель для спален будет обновляться лишь в том случае, если значения для комнат и спален согласуются между собой. Проще всего использовать линейный массив с числом комнат в качестве элементов, а в этих элементах будет указано число спален. Более сложная матрица интерполяции может включать в себя число лиц в данной жилищной единице и тип структуры.

579. В случае недостоверности показателя спален, если не используется динамическая интерполяция, следует указывать «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики (одна из них — число комнат) для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

8. Жилые помещения: топливо, используемое для приготовления пищи (A07)

580. В связи с необходимостью строгого контроля за использованием природных ресурсов, в ряд национальных переписей жилого фонда включается вопрос о топливе, используемом для приготовления пищи. Единицей учета является жилищная единица; под «топливом, используемым для приготовления пищи», имеется в виду топливо преимущественно для приготовления горячей пищи. Если используется два вида топлива (например, электричество и газ), при переписи указывается наиболее часто используемое. Классификация видов топлива, используемого для приготовления пищи, зависит от национальных условий, и может включать электричество, газ, мазут, уголь, дрова и навоз. Было бы также полезным собирать эту информацию в отношении жилых помещений коллективного проживания, особенно, если их число в стране является значительным (Организация Объединенных Наций, 1998 год, пункт 2.424).

581. Ответ на вопрос о типе топлива, используемого для приготовления пищи, следует редактировать вместе с ответами о кухонном оборудовании. Группа редактирования определяет соотношение обоих показателей и разрабатывает приемы редактирования для их проверки на согласованность. Такие показатели, как «для частного пользования», «для совместного пользования», «для исключительного пользования» и т. д.

возможно будут использоваться для определения согласованности величин, и если они не согласуются, для того, какой путь редактирования избрать для выявления проблемы. Если есть один или два показателя, касающихся приготовления пищи, то оценку неизвестной или несогласующейся информации можно произвести, не прибегая к значению «не известно» или к динамической интерполяции. Однако в случае недостоверности показателя, если не используется динамическая интерполяция, следует указывать «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, в том числе такие, как кухонное оборудование, тип строения, строительный материал стен, форма пользования и незанятость для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

9. Жилые помещения: вид отопления и используемая энергия (A08)

582. Это вопрос относится к виду отопления жилых помещений и энергии, используемой для этой цели. Единицей учета являются все жилые помещения. Данный вопрос не актуален в ряде стран, где в силу их географического положения и климата нет необходимости в отоплении жилых помещений. «Тип отопления» означает систему, используемую для обогрева большей части площади. Это может быть либо центральное отопление для обогрева всех исключительно жилых помещений или отдельного их комплекса, либо нецентральное с отоплением отдельных жилых помещений печкой, камином или каким-либо другим нагревателем. Показатель «энергия, используемая для отопления» тесно взаимосвязан с видом отопления и касается преобладающего источника энергии, например твердого топлива (уголь, лигнит и их продукты, трава), мазут, газ (природный или сжиженный) и электричество (Организация Объединенных Наций, 1998 год, пункт 2.425).

583. Вид отопления и энергия, используемая для отопления, взаимосвязаны между собой, а также с наличием горячей воды и другими видами теплоснабжения, используемыми в жилищной единице, например с электричеством и магистральным газом. Группы редактирования должны учитывать наличие этих элементов при разработке спецификаций для программ редактирования по виду отопления и энергии, используемой для обогрева. Вид отопления может не зависеть от других жилищных показателей, поэтому эти данные можно редактировать отдельно. Однако, если показатель «энергия, используемая для отопления» неизвестен или не согласуется с другими данными, то программа мо-

жет проверить вид энергии, используемой для освещения. И, наконец, в случае недостоверности показателя, если не используется динамическая интерполяция, следует указывать «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, например, тип жилищной единицы, строительный материал стен, форма пользования и незанятость, для получения «известной» информации, исходя из аналогичных жилищных единиц в географическом районе.

10. Жилые помещения: наличие горячей воды (A09)

584. Это вопрос относится к обеспеченности жилых помещений горячей водой. Горячая вода — это вода, нагретая до определенной температуры и подаваемая в помещение по трубам через кран. Сведения могут собираться в отношении наличия горячей воды в жилых помещениях либо вне их только для части жильцов или для совместного пользования, либо ее отсутствия вообще (Организация Объединенных Наций, 1998 год, пункт 2.426).

585. Наличие горячей воды может быть связано со средствами подогрева воды, хотя использование солнечной энергии для подогрева воды может быть и не связано с другими жилищными элементами. Группа редактирования должна определиться в отношении соответствующих приемов редактирования в зависимости от других жилищных элементов и географического положения. В конечном итоге в случае недостоверности показателя, если не используется динамическая интерполяция, следует указывать «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, такие как вода, подаваемая по трубопроводу, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

11. Жилые помещения: магистральный газ (A10)

586. Этот вопрос относится к наличию или отсутствию подвода магистрального газа в жилые помещения. Магистральный газ обычно определяется как природный или промышленный газ, который распределяется по трубам и потребление которого регистрируется. Данный вопрос может быть неактуален для ряда стран, где нет развитой системы газовых магистралей или отсутствуют источники природного газа (Организация Объединенных Наций, 1998 год, пункт 2.427).

587. Магистральный газ не связан с другими жилищными элементами за исключением типа освещения и

топлива, используемого для приготовления пищи. Группа редактирования должна определить приемлемый путь редактирования, а также произвести проверку на согласованность. Если показатель остается недостоверным или не согласуется с другими данными, то, когда не используется динамическая интерполяция, следует указывать «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, такие как энергия, используемая для отопления, тип строения, тип жилищной единицы, строительный материал стен, форма пользования и незанятость помещений, для получения «известной» информации, исходя из аналогичных единиц в данном географическом районе.

12. Жилые помещения: наличие телефона (A11)

588. Этот вопрос относится к обеспеченности жилищной единицы телефонной связью. Телефон — это прежде всего телефонная линия, а не сам телефонный аппарат, поскольку к одной телефонной линии может быть подключено несколько телефонных аппаратов (Организация Объединенных Наций, 1998 год, пункт 2.428).

589. В процессе редактирования данные о наличии телефонов не соотносятся с другими жилищными элементами. Однако, если в каком то географическом районе телефонная связь отсутствует, то группа редактирования должна учитывать это при разработке программы. В случае недостоверности данных о наличии телефона, если не используется динамическая интерполяция, следует указывать «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, такие как тип жилищной единицы, строительный материал стен или форма пользования, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

13. Жилые помещения: использование жилищной единицы (A12)

590. «Использование жилищной единицы» означает полное или неполное использование жилищной единицы для целей проживания (обитания). Жилищная единица может использоваться для проживания или для коммерческих, а также производственных или некоторых других целей (Организация Объединенных Наций, 1998 год, пункт 2.429).

591. Показатель «использование жилищной единицы» не зависит от других жилищных элементов. В случае недостоверности показателя, если не используется

динамическая интерполяция, следует указывать «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, такие как тип жилищной единицы, строительный материал стен, форма пользования и владения, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

14. Жильцы: число автомобилей (A13)

592. Показатель «число автомобилей» относится к числу автомобилей или легковых автофургонов, имеющих обычно в пользовании у жильцов жилищной единицы. Как правило термин «обычно имеющиеся в пользовании» означает наличие автомобилей и легковых автофургонов, которые находятся либо во владении жильцов, либо в пользовании, в соответствии с более или менее постоянным соглашением, например, договором аренды (Организация Объединенных Наций, 1998 год, пункт 2.430).

593. Число транспортных средств не зависит от других жилищных показателей. Если в стране есть районы, где нет никаких транспортных средств, специалисты могут по своему усмотрению разработать специальные приемы редактирования для конкретных географических районов. Если этого не сделано, то в случае недостоверности показателя, если не используется динамическая интерполяция, следует указать «неизвестно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, такие как тип жилищной единицы, строительный материал стен, форма пользования и владения, для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

15. Жильцы: предметы длительного пользования (A14)

594. Сведения собираются о наличии таких предметов длительного пользования, как стиральные машины (A14 a), посудомоечные агрегаты (A14 b), холодильники (A14 c), морозильные камеры (A14 d), телевизоры (A14 e) и персональные компьютеры (A14 f) в зависимости от условий в той или иной стране (Организация Объединенных Наций, 1998 год, пункт 2.421).

595. Для большинства приборов необходимо наличие питания в жилищной единице. При указании данных о таких приборах группа редактирования должна разработать такую программу, которая проверяла бы нали-

чие электропитания с возможными исключениями для холодильника, который может работать от ДВС или в виде «ящика со льдом». Кроме того, если в какой то стране для того, чтобы работала стиральная машина или посудомоечный агрегат, необходима проточная вода, то при редактировании это также необходимо учитывать. Программу редактирования можно использовать для того, чтобы определить, требуется ли наличие конкретного элемента в зависимости от наличия электричества и воды, и следует произвести соответствующие действия, если возникают несоответствия. Кроме того, в отдельных частях страны может не быть электричества или проточной воды, и специалистам это, видимо, необходимо учитывать при разработке программ редактирования. В случае недостоверности или несогласованности показателя, если не используется динамическая интерполяция, следует указать «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, такие как тип жилищной единицы, наличие электричества, строительный материал стен и форма пользования, для получения «известной» информации, исходя из аналогичных единиц в данном географическом районе.

16. *Жильцы: наличие пространства вне помещения в пользовании домохозяйства (A15)*

596. Этот вопрос относится к наличию пространства, предназначенного для отдыха членов домохозяйства, занимающих жилищную единицу. Речь здесь может идти о пространстве вне помещения, которым пользуются как частью жилищной единицы (например, задний двор в отдельных домах), пространства, прилегающие к строению (например, задний двор и площадка для игр, находящиеся вблизи многоквартирных строений), пространства, входящие в общую зону отдыха в пределах

10 минут ходьбы от жилищной единицы (например, парки, спортивные площадки и т. п.) или пространства вне помещений в пределах 10 минут ходьбы (Организация Объединенных Наций, 1998 год, пункт 2.432).

597. Величина внешнего пространства, которым пользуются члены домашнего хозяйства, не зависит от других жилищных элементов. Однако в некоторых географических районах или определенных типах строений внешнего пространства может не быть. В этом случае группа редактирования, возможно, должна предусмотреть конкретные обстоятельства при разработке своих программ. В случае недостоверности показателя, если не используется динамическая интерполяция, следует указать «не известно». Страны, применяющие метод динамической интерполяции для исчисления недостоверных показателей, должны использовать, по меньшей мере, две характеристики, например, тип строения и тип жилищной единицы для получения «известной» информации, исходя из аналогичных жилищных единиц в данном географическом районе.

С. ЗАНЯТЫЕ И НЕЗАНЯТЫЕ ЖИЛИЩНЫЕ ЕДИНИЦЫ

598. Описанные выше приемы редактирования касаются данных о занятых жилищных единицах. Однако у занятых жилищных единиц и свободных жилищных единиц зачастую разные характеристики и одни и те же приемы к ним не подходят. Группа редактирования в национальном бюро переписи/статистическом управлении должна будет разработать разные программы редактирования для каждого типа единицы, если, как это обычно и бывает, для незанятых жилищных единиц собраны не все жилищные элементы. Группе редактирования надо будет уделить особое внимание переменным в матрице интерполяции, поскольку, вероятнее всего, они будут разными.

ПРИЛОЖЕНИЯ

Приложение I

ПРОИЗВОДНЫЕ ПЕРЕМЕННЫЕ

599. Для того чтобы наилучшим образом использовать данные переписи или обследования, в странах зачастую нужны переменные, представляющие собой комбинации из других переменных. Например, элемент, касающийся статуса экономической активности (глава IV, раздел D.7) — это уже комбинация нескольких переменных, собранных во время переписи. Вместо того, чтобы разрабатывать специальную программу для повторного кодирования информации каждый раз, когда национальному бюро переписи/статистическому управлению понадобится специальная таблица, специалисты по обработке данных могут записать программу для повторного кодирования один раз, хранить повторно закодированную информацию в файле конкретного лица и затем использовать ее для последующих таблиц. Национальные бюро переписи/статистические управления должны решить, как часто использовать повторно закодированные данные в таблицах и насколько необходимы будут конкретные повторно закодированные данные для решения вопроса о том, нужно ли составлять и хранить эти сведения. Необходимо помнить о том, что повторно закодированные данные также занимают определенное место в записях, касающихся конкретного лица. Чем больше численность населения, тем больше места будет использоваться.

600. Таким образом, можно создать множество переменных. Например, если сообщается дата рождения, а не возраст, то возраст можно определить путем вычитания даты рождения из базисной даты переписи, и эта информация будет храниться в записях. Аналогичным образом можно вычислить совокупный доход домохозяйства, суммируя доход каждого индивида и указывая эту сумму в записи, касающейся жилья, для дальнейшего использования.

601. Иногда производные переменные получаются из сочетания одного или нескольких показателей в одной записи, а иногда из нескольких записей. Например, классификация «Экономически неактивное население — посещение учебных заведений» может потребовать просмотра ответов на четыре вопроса. При разработке формата таблиц или планировании дополнительных таблиц использование производных переменных облегчит программирование и сделает его более эффективным, а также будет способствовать тому, чтобы данные оставались сопоставимыми с течением време-

ни. Ниже приводятся некоторые примеры возможных производных записей.

А. ПРОИЗВОДНЫЕ ПЕРЕМЕННЫЕ ДЛЯ ЗАПИСЕЙ О ЖИЛОМ ФОНДЕ

1. *Доход домохозяйства*

602. Производные переменные для дохода домохозяйства складываются из суммы дохода, включающего в себя все категории дохода всех лиц в домохозяйстве. Категории информации о доходе могут включать в себя заработную плату, доход от собственного дела, проценты и дивиденды, социальное обеспечение и пенсии, денежные переводы, гонорары и рентные доходы. Если собираются также сведения об общем доходе, то при редактировании следует проверить общий доход каждого лица путем суммирования отдельных категорий. Затем эта общая сумма проверяется по зарегистрированному общему доходу. Если суммированный общий доход не совпадает с зарегистрированным общим доходом, то группа редактирования должна разработать план корректировки. При этом необходимо или изменить общий итог, чтобы отразить сумму отдельных частей, или одну или более отдельных категорий дохода. Когда установлена общая сумма дохода для всех членов домохозяйства, показатель «доход домохозяйства» складывается из суммы отдельных доходов.

603. Группа редактирования должна учитывать ситуацию, при которой один или более человек в домохозяйстве имеют отрицательный доход из-за неудачного бизнеса или по другим причинам. В этом случае общий доход домохозяйства не увеличится, а уменьшится на сумму дохода этого конкретного лица.

2. *Семейный доход*

604. Производная переменная для семейного дохода складывается из суммы дохода, получаемого по всем категориям дохода всеми членами семьи. Семьи, в отличие от домохозяйств, как правило, состоят только из родственников, но это определение будет зависеть от ситуации в конкретной стране. Для некоторых стран домохозяйство и семья — это одно и то же, поэтому

производная переменная для семейного дохода будет необязательна. Сведения о семейном доходе могут включать в себя такие категории, как заработная плата, доход от собственного дела, проценты и дивиденды, социальное обеспечение и пенсии, денежные переводы, гонорары и рентные доходы. Если собираются также сведения об общем доходе, то при редактировании следует проверить общий доход каждого лица путем суммирования отдельных категорий. Затем эта итоговая сумма сверяется с зарегистрированным общим доходом. Если суммированный общий доход не совпадает с зарегистрированным общим доходом, то группа редактирования должна разработать план корректировки. Здесь необходимо изменить или общий итог, чтобы отразить сумму отдельных категорий, или одну или более отдельных категорий дохода. Когда устанавливается общий доход для всех индивидов, то семейный доход складывается из суммы отдельных доходов в семье.

605. Группа редактирования должна учитывать ситуацию, когда один или более членов семьи имеют отрицательный доход из-за неудачного бизнеса или по другим причинам. В этом случае общий семейный доход не увеличится, а уменьшится на сумму дохода этого конкретного лица.

3. Тип семьи

606. Для некоторых таблиц иногда полезно выявлять «тип семьи». Например, производная переменная для типа семьи может быть от единицы до восьми, представляя тип и состав семьи. Производная переменная для типа семьи может быть использована для того, чтобы посмотреть, как различные характеристики влияют на состав семьи.

607. Как указывается в *Принципах и рекомендациях в отношении переписи населения и жилого фонда* (Организация Объединенных Наций, 1998 год, пункты 2.60–2.66), каждая страна может давать определение семьи по-своему. По одному из определений семья состоит из главы домохозяйства и одного или более других лиц, проживающих совместно в одном домохозяйстве и связанных отношениями с главой домохозяйства, вытекающими из родства, брака или усыновления. Все члены домохозяйства, связанные родственными отношениями с главой домохозяйства, являются членами его или ее семьи. Однако не во всех домохозяйствах есть семьи, поскольку домохозяйство может состоять из группы лиц, не связанных между собой родственными отношениями, или из одного отдельно живущего лица.

608. Специалисты, занимающиеся содержанием переписи, могут подразделять семьи на «семьи в составе супружеской пары» или «прочие семьи» в зависимости от пола главы домохозяйства и наличия родственников.

Например в Соединенных Штатах Америки используются следующие коды на основе данных о типе семьи, выведенных из ответов на вопросы о поле и степени родства:

- a) Коды 1 и 2: **семья в составе супружеской пары.** Семья, где глава домохозяйства и его супруг(а) были зарегистрированы как члены одного и того же домохозяйства. Код 1 используется, если глава домохозяйства мужчина, код 2 — если глава домохозяйства женщина.
- b) Код 3: **прочие семьи: глава домохозяйства — мужчина, без жены.** Семья с главой домохозяйства мужчиной и без супруги.
- c) Код 4: **прочие семьи: глава домохозяйства — женщина, без мужа.** Семья с главой домохозяйства женщиной и без супруга.
- d) Коды 5 и 6: **несемейное домохозяйство.** Домохозяйство, которое не является семьей, поэтому супруга(и) или другого родственника главы домохозяйства нет; код 5 используется, когда главой домохозяйства является мужчина, код 6 — когда главой домохозяйства является женщина.
- e) Коды 7 и 8: **домохозяйство из одного человека.** Один человек, проживающий отдельно, считается домохозяйством, но не семьей, поскольку никаких родственников нет. Код 7 используется, когда главой домохозяйства является мужчина, код 8 — когда главой домохозяйства является женщина.

609. Более простой метод определения части вышеперечисленных категорий заключается в том, чтобы получить производную переменную под названием «глава домохозяйства, в браке, супруг(а) есть». Семейному состоянию главы домохозяйства можно присвоить новый код в зависимости от того, если в домохозяйстве супруг(а) главы домохозяйства. По каждой жилищной единице демографические данные сканируются в поиске лица с супругой(м) в качестве родственника. В жилищной записи в соответствующем поле ставится код для «да» или «нет». Для помещения коллективного проживания здесь можно оставить пробел или присвоить другой код. Затем для демографических таблиц состоящие в браке лица и при наличии супруги(а) будут определяться во время табулирования.

4. Лица, связанные родственными отношениями

610. Связанные родственными отношениями лица — это те, кто связан какими-либо отношениями с главой домохозяйства. Производная переменная для связанных родственными отношениями лиц складывается из суммы всех лиц, связанных отношениями с главой до-

мохозяйства. Этот показатель особенно важен, когда большое число лиц, не связанных между собой родственными отношениями, проживают совместно в жилищных единицах. Когда много лиц, не связанных родственными отношениями, проживают совместно, их часто классифицируют как лица, живущие в помещениях коллективного проживания или помещения группового проживания.

611. При разработке наборов данных национальные статистические управления часто разрабатывают производные переменные для различных наборов, связанных родственными отношениями лиц по возрасту. Например, производные переменные можно разработать для связанных родственными отношениями детей от 0 до 5 лет, связанных родственными отношениями детей от 5 до 17 лет, связанных родственными отношениями детей от 6 до 17 лет, связанных родственными отношениями детей от 0 до 17 лет, связанных родственными отношениями лиц в возрасте 65 лет и старше и связанных родственными отношениями лиц в возрасте 75 лет и старше.

612. Понятие «связанные родственными отношениями дети» в семье может включать, например, собственных детей главы домохозяйства и других лиц моложе 18 лет в домохозяйстве, независимо от семейного состояния, которые связаны родственными отношениями с главой домохозяйства, за исключением супруги(а) главы домохозяйства. По понятиям «связанные родственными отношениями дети» могут подразумеваться или не подразумеваться приемные дети, поскольку они не связаны родственными отношениями с главой домохозяйства, но это решение зависит от конкретной ситуации в стране.

5. Работники в семье

613. Иногда страны по своему усмотрению могут сопоставлять показатели домохозяйств по числу работников, например показатели распределения дохода по составу домохозяйства и числу работников на иждивенца. Страна может получать производную переменную для числа работников в семье, суммируя число лиц, которые работали по меньшей мере один час в базисный период, например неделю или год (или календарный год или последние 12 месяцев). Страна может использовать показатель, касающийся числа лиц, выполняющих работу в течение «последней недели», если данные собираются только за этот период.

6. Комплексное сантехническое оборудование

614. В переписном листе есть несколько элементов, предназначенных для получения данных о сантехниче-

ском оборудовании. Эти элементы, как правило, связаны с наличием водопровода, туалета со смывом и ванной или душа и обычно собираются как по занятым, так и незанятым жилищным единицам. Производная переменная для комплексного сантехнического оборудования может помочь в сопоставлении социально-экономических условий между районами или группами населения в определенный момент времени или в течение целого периода. Производная переменная для комплексного сантехнического оборудования может потребоваться, например, когда есть три вида удобств: водопровод (с горячей или холодной водой), туалет со смывом и ванна или душевая установка (в самой единице или за пределами строения, где расположена эта единица). Группа редактирования должна будет определить наиболее приемлемый набор переменных для комплексного сантехнического оборудования.

615. В данном примере производная переменная может потребоваться, если все три элемента запрашиваются отдельно и при редактировании наличие всех трех элементов будет определяться в сумме. Если в жилищной единице есть водопровод, туалет со смывом и ванна или душ, то следует указывать «имеется комплексное сантехническое оборудование». Без всех трех элементов следует указывать «нет комплексного сантехнического оборудования».

7. Комплексное кухонное оборудование

616. Во время переписи собираются данные о кухонном оборудовании по элементам переписного листа, касающегося оборудования для приготовления пищи, холодильника и раковины; эти данные собираются как по занятым, так и не занятым жилищным единицам. Можно считать, что в данной жилищной единице есть «все кухонное оборудование», когда оборудование для приготовления пищи (электрическая, керосиновая или газовая плита, микроволновая печь и неподвижные горелки или кухонная плита), холодильник и раковина с водопроводной водой расположены в том же строении, где зарегистрированы жилые помещения. Все это не обязательно может быть в одной комнате.

617. Производная переменная выводится, если вышеперечисленные три элемента запрашиваются отдельно, и при редактировании наличие всех трех элементов определяется по сумме. «Отсутствует комплексное оборудование» включает в себя такие условия, когда есть все три указанных кухонные удобства, но оборудование расположено в другом строении; есть некоторые, но все удобства; или нет ни одного из трех указанных кухонных удобств в строении, где зарегистрированы жилые помещения.

8. *Валовая квартирная плата*

618. Как правило, страны собирают данные о квартирной плате за наличный расчет или договорной квартирной плате. Из квартирной платы за наличный расчет, как правило, исключаются расходы на коммунальные услуги. Иногда странам нужны также сведения о валовой квартирной плате. Валовая квартирная плата — это договорная квартирная плата или плата за наличный расчет плюс среднемесячная стоимость коммунальных услуг (электричество, газ и вода) и топливо (включая нефтепродукты, уголь, керосин и дрова), если оплату этих услуг должен производить съемщик. Сведения о валовой квартирной плате используются для того, чтобы ликвидировать разницу, которая возникает в связи с разностью подходов к включению коммунальных услуг и топлива в арендную плату. Единицу, занимаемую съемщиком без оплаты наличными, можно в таблицах показать отдельно в графе «без оплаты наличными».

619. Производная переменная для валовой квартирной платы складывается из суммы производимой оплаты и оплаты за коммунальные услуги, если эти данные собираются отдельно.

В. ПРОИЗВОДНЫЕ ПЕРЕМЕННЫЕ
ДЛЯ ДЕМОГРАФИЧЕСКИХ ЗАПИСЕЙ1. *Статус экономической активности*

620. Производная переменная для статуса экономической активности может оказаться весьма полезной при составлении таблиц, но она требует информации по нескольким показателям. В соответствии с первым пересмотренным изданием *Принципов и рекомендаций в отношении переписи населения и жилого фонда* необходима реконфигурация нескольких переменных. Производная переменная может состоять из девяти категорий:

Экономически активное население*Занятое население*

1. Работающее

2. Имеющее работу, но не работающее

3. Служащие в вооруженных силах

Незанятое население

4. Ищущее работу

5. Переставшее искать работу

Экономически неактивное население

6. Лица, ведущие домашнее хозяйство

7. Учащиеся

8. Неспособные работать

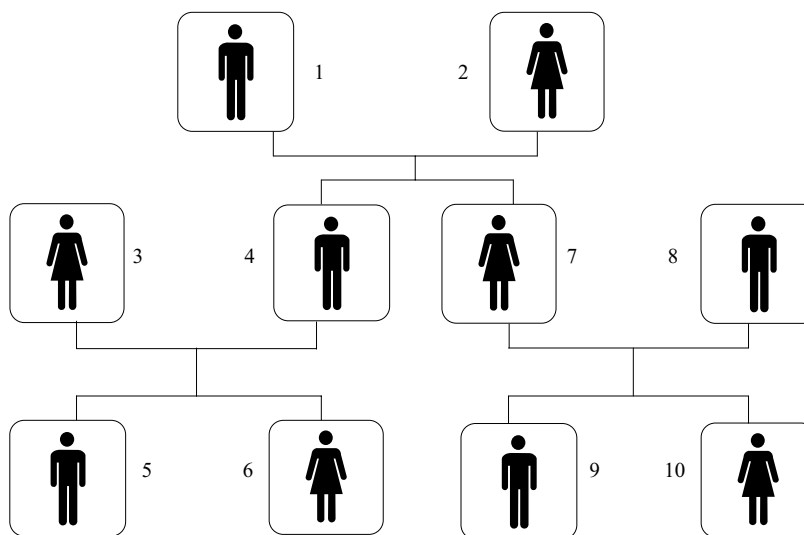
9. Прочие

621. Поскольку во многих смежных таблицах используются самые различные классификации экономической активности, группа редактирования должна подумать над тем, чтобы ввести в записи данных производную переменную, вместо того, чтобы обработчики данных реклассифицировали статус экономической активности при составлении таблиц. Реклассификация во время составления таблиц может привести к ошибкам, поскольку разные разработчики данных могут чуть по-разному произвести реклассификацию; даже одна программа может произвести реклассификацию по-разному, в зависимости от конкретных требований к редактированию или составлению таблиц. Специалисты по экономическим характеристикам должны подготовить спецификации для производной переменной.

2. *Число субсемей и родственники в субсемье*

622. Во всех странах есть как расширенные, так и нуклеарные семьи. Рассмотрим следующую расширенную семью (см. рисунок А.1.1, ниже). В домохозяйстве есть глава домохозяйства и супруг(а) (номера 1 и 2) с двумя детьми (номера 4 и 7). Их сын (лицо 4) женат и его жена приходится родителям невесткой (лицо 3), и у них есть два внука, являющиеся детьми их сына (лица 5 и 6). Их дочь (лицо 7) замужем, и ее муж приходится родителям зятем (лицо 8), и у них два внука, являющиеся детьми их дочери (лица 9 и 10).

Рисунок А.1.1. Расширенная семья



623. При проведении большинства переписей, если группа редактирования захочет изучить структуру семьи, которая показана, например, на рисунке А.1.1, то ей, возможно, трудно будет провести различие между внуками, поскольку лица 5, 6, 9 и 10 будут зарегистрированы как «внук/внучка» по отношению к главе домохозяйства. Замена кодов для субсемьи и члена субсемьи позволит произвести более подробный анализ состава семьи.

624. По одному из определений под субсемьей может подразумеваться «супружеская пара (муж и жена, зарегистрированные как члены одного и того же домохозяйства с неженатыми или незамужними детьми в возрасте до 18 лет или без таковых)». Группа редактирования может по своему усмотрению добавить к этому такое определение, как «один родитель с одним или более неженатых или незамужних детей в возрасте до 18 лет, живущих в домохозяйстве и связанных родством с главой домохозяйства или супругой(ом) главы домохозяйства, но не включая их». Число субсемей не учитывается в данных о семьях, поскольку члены субсемьи учитываются в составе семьи главы домохозяйства.

625. Производные переменные для субсемей, включая число и тип родственников, можно определить во время обработки данных. Можно разработать специальную программу редактирования для того, чтобы легче было определять субсемьи на основе родственных связей внутри домохозяйства. При определении каждой субсемьи — не главы домохозяйства/супруги(а) (с детьми или без детей) или не главы домохозяйства/родителей ребенка — каждой субсемье присваиваются номера по порядку. Затем различные степени родства могут быть обозначены кодами: «глава семейного

домохозяйства» — код 1, «супруга (супруг) главы семейного домохозяйства» — код 2 и ребенок «главы семейного домохозяйства» — код 3. Чтобы указать субсемью, должна быть, по крайней мере, одна пара родственников субсемьи: или «глава домохозяйства» и «супруг (супруга)» (коды 1 и 2) или «глава домохозяйства» и «ребенок» (коды 1 и 3). Если глава семейного домохозяйства, его супруга или супруг и ребенок живут все вместе, то субсемье будут присвоены коды 1, 2 и 3. Субсемьи подразделяются по типу: субсемьи в составе супружеской пары с собственными детьми или без детей; субсемьи в составе матери и ребенка; и субсемьи в составе отца и ребенка. К родителям-одиночкам относятся те, кто содержат семьи или субсемьи с одним родителем. К супружеским парам относятся мужья и жены в семьях и субсемьях в составе супружеской пары.

626. При разработке производных переменных степень родства с главой домохозяйства используется для определения родственных отношений в семье; поэтому, чем подробнее кодируются родственные отношения в данных переписи, тем легче производить сопоставление субсемей. Например, если такая степень родства как «зять» или «невестка» имеет собственный код, то программа сможет произвести согласование «сына/дочери» с «сыном/невесткой», то есть лицом противоположного пола для создания субсемьи. Без этой дополнительной информации согласование тоже можно произвести, но при этом понятие «прочие родственники» может быть двусмысленным или согласование может быть произведено неправильно. Аналогичным образом программа будет сопоставлять коды для «детей главы домохозяйства» с кодами для «зятя» или «невестки» и с кодами для «племянницы/племянника».

627. На рисунке А.1.2 показано домохозяйство с двумя субсемьями: субсемья 1 состоит из лица 3 (глава домохозяйства в субсемье 1), лица 4 (супруг) и лиц 5 и

6 (дети в субсемье 1). Субсемья 2 состоит из лица 7 (глава домохозяйства субсемьи 2), лица 8 (супруг) и лиц 9 и 10 (дети в субсемье 2).

Рисунок А.1.2. Выборочное домохозяйство с двумя субсемьями

Номер лица	Степень родства	Пол	Субсемья	
			Номер	Степень родства
1	Глава домохозяйства	М		
2	Супруга	Ж		
3	Сын	М	1	1
4	Невестка	Ж	1	2
5	Внук	М	1	3
6	Внучка	Ж	1	3
7	Дочь	Ж	2	1
8	Зять	М	2	2
9	Внук	М	2	3
10	Внучка	Ж	2	3

3. Собственные дети

628. Иногда страны хотят получить сведения о «собственных детях», которые являются биологическими детьми главы домохозяйства и/или супруги(а). «Собственным ребенком» может быть, например, неженатый или незамужний ребенок в возрасте до 18 лет, который является сыном или дочерью по рождению, пасынком или падчерицей (одного из родителей) или приемным ребенком главы домохозяйства. В таблицах можно показать данные о «собственных детях» с дальнейшей разбивкой на живущих с двумя родителями или только с одним родителем. При такой схеме «собственные дети» главы домохозяйства, проживающие с двумя родителями, по определению будут только в семье в составе супружеской пары.

629. В субсемье «собственный ребенок» может означать неженатого или незамужнего ребенка в возрасте до 18 лет, который является сыном, дочерью или пасынком или падчерицей; или приемным ребенком (1) матери в субсемье в составе матери и ребенка, (2) отца в субсемье в составе отца и ребенка или (3) одного из супругов в субсемье в составе супружеской пары.

630. Производная переменная для понятия «собственные дети» может складываться из суммы числа собственных детей конкретного лица, как правило, женского рода, с определениями, отобранными группой редактирования. Иногда пользователям нужны более подробные сведения о собственных детях по возрасту. Например, для Соединенных Штатов производная переменная разрабатывается для числа собственных де-

тей младше 6 лет и для детей в возрасте от 6 до 17 лет. Эти данные указываются в записях о всех лицах женского пола. Такая информация используется, в частности, для определения характеристик лиц женского пола в самодеятельном населении с «собственными детьми».

4. Родители в доме

631. Эти данные показывают характеристики детей в семьях с одним родителем в сопоставлении с жилищными единицами, где проживают оба родителя. Программа редактирования дает эту производную переменную, определяя, сколько родителей конкретного лица есть в доме с использованием кодов степени родства. Программа показывает код степени родства для каждого ребенка и использует эту информацию в сочетании с информацией о субсемьях, чтобы определить, сколько родителей живет в жилищной единице.

5. Текущий год в учебном заведении

632. В некоторых странах задается два вопроса об образовании:

- посещает ли данное лицо в настоящее время учебное заведение;
- высший уровень образовательной подготовки.

633. В этих странах группа редактирования, зачастую, находит несоответствие между двумя элементами, когда лицо фактически посещает учебное заведение во

время регистрации. Иногда это приводит к тому, что высший уровень образовательной подготовки данного лица указывается на год меньше, чем год фактического посещения учебного заведения. Если лицо находится в середине ряда ступеней или уровней, то на статистику это не повлияет. Однако, если данное лицо посещает первую ступень в течение конкретного уровня, то согласование с данными из других источников может оказаться невозможным. Например, лицо, посещающее первый класс, будет зарегистрировано, как посещающее учебное заведение, но без образовательной подготовки. Аналогичным образом лицо, посещающее среднюю школу, будет зарегистрировано, как посещающее учебное заведение, но уровень образовательной подготовки при этом будет последняя ступень (или уровень) начальной школы.

634. Для такого сочетания элементов можно разработать производную переменную «текущий год в учебном заведении». Если данное лицо на текущий момент не посещает учебное заведение, то код будет таким же, как для высшего уровня образовательной подготовки. Если данное лицо посещает учебное заведение на текущий момент, то программа редактирования добавит единицу к ступени (или уровню) образовательной подготовки и отнесет это к показателю «текущий год в учебном заведении».

635. Некоторые страны задают три вопроса об образовании, два из них отмечены выше, а третий вопрос касается того, завершено ли обучение на высшей ступени. Если такие сведения собираются, то их следует использовать также при определении показателя «текущий год в учебном заведении».

Приложение II

ВЗАИМОСВЯЗЬ МЕЖДУ ФОРМАТОМ ПЕРЕПИСНОГО ЛИСТА И ВВОДОМ ДАННЫХ С КЛАВИАТУРЫ

636. Двумя самыми распространенными форматами переписного листа для демографических элементов при проведении переписи или обследовании являются страницы, посвященные населению, и страницы, посвященные домохозяйствам.

637. Страницы, посвященные населению, содержат одну или две титульные страницы с демографической информацией с отдельными страницами для каждого лица. Такой метод полезен потому, что все сведения, касающиеся одного лица, записываются на одной странице, что облегчает сбор данных. Кроме того, такой формат облегчает проверку на внутреннюю согласованность во время регистрации. Персональные страницы можно объединять под одним переплетом для облегчения обработки данных на местах, как показано на рисунке А.П.1.

Рисунок А.П.1. Выборочный бланк переписного листа с персональными страницами

<i>Персональные страницы для лица X</i>		<i>Персональные страницы для лица X + 1</i>	
Элемент 1	Элемент 10	Элемент 1	Элемент 10
Элемент 2	Элемент 11	Элемент 2	Элемент 11
.	.	.	.
.	.	.	.
.	.	.	.

638. Кодирование и введение данных с клавиатуры с персональных страниц — это, в основном, механическая операция, при которой не предполагается, что кодировщик/оператор по вводу данных будет оценивать достоверность представленной информации, а от него требуется лишь присваивать соответствующий код и правильно нажимать на клавиши. На рисунке А.П.2 показан поток информации для данного лица, записанной на одной странице. Гораздо проще вводить данные с одной страницы, чем работать на клавиатуре, одновременно переворачивая страницы. Проверка на достоверность производится позднее во время машинного редактирования.

Рисунок А.П.2. Пример потока в переписном листе с персональными страницами

<i>Персональная страница</i>			
Элемент 1		Элемент 11	
Элемент 2	↓	Элемент 12	↓
Элемент 3		Элемент 13	
и т. д.		и т. д.	
.			
.			
.			

639. В странице, посвященной домохозяйствам, содержатся все сведения о домохозяйстве, по возможности, на одной странице или на нескольких страницах, при том, что данные о каждом члене домохозяйства будут записываться на одной странице. Такой метод перечисления членов домохозяйства полезен, поскольку элементы переписного листа не надо печатать для каждого человека отдельно, в связи с чем экономится место. Кроме того, счетчик может сопоставить записи о членах домохозяйства после сбора данных.

Рисунок А.П.3. Выборочный переписной лист, страница, посвященная домохозяйству, со всеми его членами на одной странице


<i>Страница, посвященная домохозяйству</i>					
<i>Лицо</i>	<i>Элемент 1</i>	<i>Элемент 2</i>	<i>Элемент 3</i>	<i>Элемент 4</i>	<i>и т. д.</i>
1					
2					
3					
4					
5					
.					
.					
.					

640. Третий метод заключается в том, чтобы иметь отдельные бланки для каждого лица, из которых счетчик во время или после регистрации составляет буклет с вкладными страницами. Такой метод является эффективным, поскольку счетчик собирает только точное количество бланков (страниц), необходимое для данного домохозяйства. Недостаток здесь заключается в том, что при передаче или во время каких-то других операций бланки могут отсоединиться, что чревато множеством проблем, которые могут возникнуть при редактировании и охвате информации, если бюро переписи не сможет вновь соединить их для правильного отражения данных о домохозяйстве.

641. Физический размер страниц вопросника также имеет значение не только для регистрации, но и для ввода данных с клавиатуры. Во время кодирования или ввода данных документ должен ровно лежать на поверхности стола, и кодировщики или операторы должны легко находить и обрабатывать элементы бланка.

642. Когда вся информация содержится на одной странице, то сотрудникам будет легко вводить данные, и, несомненно, это будет быстрее, поскольку вводящему данные оператору не надо будет переворачивать страницы. На рисунке А.П.4 показана последовательность информации со страницы, посвященной домохозяйству.

Рисунок А.П.4. Пример последовательности информации для переписного листа со страницами, посвященными домохозяйству, где на одной странице содержатся данные о всех членах домохозяйства

Страницы, посвященные домохозяйству					
Лица	Элемент 1	Элемент 2	Элемент 3	и т. д.	
1	→				
2	→				
3	→				
4					
5					
.					
.					
.					

643. Проблемы могут возникнуть в том случае, если вводится информация о населении или жилом фонде, расположенной более чем на одной странице. Для разрешения этой проблемы национальное статистическое управление может использовать один из двух подходов, изложенных ниже.

644. Данные можно вводить об одном лице за один раз. Оператор по вводу данных может набирать строку информации для данного лица на первой странице из ряда страниц, а затем переходить на вторую и последующую страницы. Когда страницы, посвященные первому лицу заканчиваются, оператор возвращается к первой из посвященных данному домохозяйству страниц и набирает данные о втором лице, потом о третьем и т. д. Такой способ введения данных работает до тех пор, пока оператор может оставаться на соответствующей строке в течение всего набора. Хотя могут быть такие программы машинного редактирования, которые восстановят информацию, если данные о том или ином лице ошибочно внесены в строку, посвященную другому лицу, но подготовить саму программу очень сложно.

645. Данные можно вводить по одной странице за один раз. Оператор может набрать целую страницу информации, а затем перейти на следующую страницу. При этом оператор набирает всю информацию, изложенную на первой странице, независимо от числа лиц. Затем оператор переворачивает страницу и набирает следующую часть информации по всем лицам. При этом схема пропуска может включать или не включать, в зависимости от способа набора данных (с машинным редактированием или без него). В любом случае во время компьютерного редактирования записи из различных наборов вводимых данных необходимо будет собирать, а затем уже рассматривать всю неправильно введенную информацию по числу лиц.

646. В следующем примере (рисунок А.П.5) демографическая информация о домохозяйстве не представляет никаких особых проблем в связи с вводом данных с клавиатуры, поскольку при проведении переписи были получены ответы по всем элементам в отношении всех лиц.

Рисунок А.П.5. Пример страницы, посвященной домохозяйству, где указаны все лица, без проблем, связанных с вводом данных

<i>Страницы, посвященные домохозяйству</i>					
<i>Лица</i>	<i>Степень родства</i>	<i>Пол</i>	<i>Возраст</i>	<i>и т. д.</i>	
1	Глава домохозяйства	М	40		
2	Супруг(а)	Ж	35		
3	Ребенок	Ж	18		
4	Ребенок	М	12		
5	Брат (сестра)	М	35		
6	Ребенок супруги	Ж	30		
7	Ребенок брата (сестры)	М	5		
8	Ребенок брата (сестры)	Ж	3		
и т. д.					

647. Вместе с тем вторая страница, посвященная тому же домохозяйству (рисунок А.П.6), может вызвать некоторые проблемы, связанные с вводом данных. Например, если страна предпочитает собирать данные о языке по одному и тому же домохозяйству только в отношении лиц в возрасте 5 лет и старше, вместо информации о восьмом лице, (в возрасте 3 года), будет пробел. Оператор по вводу данных должен оставить эту клетку для данного ребенка незаполненной, а программа редактирования попытается позднее ликвидировать этот пробел.

648. Пробелы должны быть и у других элементов, например, в связи с такими признаками, как минимальный возраст для характеристики активности населения, женщины, не достигшие минимального возраста для рождаемости, и рождаемость в отношении всех мужчин. На рисунке А.П.6 показано, что оператор мог неправильно ввести информацию о шестом лице в отношении детей, родившихся живыми (в данном случае 4) и поставить ее по ошибке в гнездо пятого лица. Программа редактирования затем устранил данные о рождаемости мужчин и вычислит рождаемость для женщин, но она может вычислить данные значения и неправильно.

Рисунок А.П.6. Пример страницы, посвященной домохозяйству, где указаны все лица с возможными проблемами ввода данных

<i>Страница 2, посвященная домохозяйству</i>				
<i>Лица</i>	<i>Язык</i>	<i>Активность населения</i>	<i>Дети, родившиеся живыми</i>	<i>и т. д.</i>
1	Язык 1	Да		
2	Язык 1	Нет	3	
3	Язык 1	Нет	0	
4	Язык 1			
5	Язык 1	Да		
6	Язык 1	Нет	4	
7	Язык 1			
8				
и т. д.				

649. Из-за стоимости и в связи с ограниченностью места страна, вынуждена использовать бланк, посвященный домохозяйству, множество раз. Однако, если население страны небольшое или страна может позволить себе дополнительные расходы, то в бланке с данными о населении вероятно будет меньше ошибок, связанных с неправильным вводом данных, чем в бланках, посвященных домохозяйству.

Приложение III

ВВОД ДАННЫХ С КЛАВИАТУРЫ

650. Многие страны используют сканирующие устройства, например, оборудование оптического распознавания меток (ОПМ) или оборудование оптического распознавания символов (ОРС). Каждое из них имеет свои преимущества перед вводом данных с клавиатуры, если операция проходит гладко и эффективно и если затраты при этом небольшие. Однако многие страны, даже те, которые доверяют сканированию, не могут себе позволить первоначальные затраты на освоение этого метода или постоянные расходы по обслуживанию во время и после регистрации. Кроме того, многие страны используют сканеры, приобретенные для проведения переписи, для выполнения ряда других задач, включая другие обследования, ведение таких административных записей, как бланки ввода и вывода информации. Однако основные навыки, связанные с подачей документов на сканер, передаются только если используются те же самые или аналогичные машины.

651. Одно из преимуществ ввода данных с клавиатуры заключается в том, что навыки, приобретенные во время ввода данных с клавиатуры, используются при выполнении других работ в национальных бюро переписи/статистических управлениях и других государственных ведомствах. После того, как операторы по вводу данных приобретут мастерство при проведении переписи, они затем работают с данными и для различных последующих обследований. К таким обследованиям относятся пострегистрационные обследования (ПРО) и другие обследования, например обследования по вопросам рождаемости и доходов и расходов домохозяйств. Сотрудники могут также вводить административные записи, включая записи актов гражданского состояния, и записи, касающиеся торговли, иммиграции и эмиграции и таможенных процедур.

652. В связи с этим, если страна рассматривает вопрос о том, использовать ли сканирующие устройства при проведении своей переписи, она должна также решить вопрос о том, будет ли она и далее использовать эти машины. Многоцелевые машины остаются пригодными еще долго и после проведения переписи. Национальные бюро переписи/статистические управления, которые вводят данные с клавиатуры, увидят, что приобретенные навыки не теряются, а сами машины можно использовать и дальше или в национальном статисти-

ческом управлении, или в других государственных органах. Возможность последующего использования данного оборудования, вероятно, следует учитывать при принятии решения относительно ввода данных с клавиатуры или сканирования.

А. ВВОД ДАННЫХ

1. Сканирование

653. Страны, использующие оптические или другие сканирующие устройства для сбора данных, как правило, их после этого корректируют, хотя изменения могут зависеть от схем пропуска, встроенных в систему. Однако у стран, предпочитающих вводить данные с клавиатуры, есть возможности выбора в зависимости от того, как быстро им нужно ввести данные и какой потребуется объем ручных проверок. Каждая возможность зависит от требований группы редактирования, навыков операторов по вводу данных и степени сложности программы редактирования. При крупных переписных операциях самая большая проблема заключается в том, чтобы вообще использовать клавиатуру для ввода данных. Здесь необходимо использовать любые методы для более быстрого получения результатов. В *Принципах и рекомендациях* отмечается, что страны быстрее достигают конечных результатов, имея достаточно аппаратуры, хорошо подготовленных сотрудников и опытных операторов по вводу данных. Однако все это требует соответствующего финансирования, которое не всегда выделяется.

654. Состав и тип требуемого оборудования ввода данных зависит от выбранного метода сбора данных, времени, выделенного для этой фазы переписи, размеров страны, степени децентрализации операций по сбору данных и ряда других факторов. При вводе данных с клавиатуры средние темпы ввода варьируются обычно в пределах от 5 до 10 тысяч ударов в час. При этом одни операторы работают в гораздо более низком темпе, а другие значительно превышают средние темпы. К числу факторов, которые влияют на скорость работы оператора, относятся: а) вспомогательная программа и программное обеспечение; б) сложность задач оператора;

с) эргономические характеристики, надежность и скорость срабатывания оборудования; d) степень загрузки оператора; e) уровень подготовки и навыки нанятого персонала; и f) мотивация работающих (Организация Объединенных Наций, 1998 год, пункт 1.193).

2. Реверсивный ввод данных

655. Есть две формы реверсивного ввода данных. При использовании первой формы вводятся все встречающиеся элементы данных без схем пропуска. В этом случае ввод данных с клавиатуры происходит более быстро, поскольку операторы по вводу данных не прекращают свою работу, когда встречаются недостоверные или несогласующиеся сведения. Возможно, такая форма и более точная, поскольку операторы выполняют свою работу больше механически. При использовании второй формы реверсивного ввода данных с клавиатуры они вынуждены останавливаться для проверки недостоверных или несогласующихся ответов в переписных листах, поэтому процесс идет более медленно и требует участия гораздо более опытных операторов. Цена здесь с точки зрения скорости очень высокая, и этот вопрос требует серьезного внимания. Как это не парадоксально, при таком методе точность может даже повыситься, если оператор увидит, что данные были внесены правильно, но им был присвоен неправильный код. Неправильный ввод данных с клавиатуры сам по себе может быть иногда незамедлительно компенсирован, поскольку пакет редакторских программ обеспечивает автоматическую проверку.

а) Реверсивный ввод данных без схем пропуска

656. Когда все данные вводятся или пропускаются вручную, можно поддерживать определенный ритм работы, и некоторые схемы пропуска не обойдут достоверные, но временно несогласующиеся сведения. Например, если то или иное лицо зарегистрировано как мужчина, большинство групп редактирования потребует пропустить весь раздел о рождаемости. В этом случае оператор по вводу данных будет набирать насквозь (используя клавишу пробела или стрелку, чтобы пройти через запись, касающуюся мужчины или молодой девушки), поскольку все поля будут пустыми. Однако это требует времени, и промежуток может быть неполностью точным. Например, оператор может пройти слишком далеко или, наоборот, не дойти до нужного места, и тогда другие элементы могут быть введены неправильно, поскольку будут выстроены не так, как нужно. Если все поля вводятся таким образом, то эту информацию можно вводить, не имея схем пропуска. Например, если оператор встречает данные о рождаемости в отношении взрослых женщин (женщины, в отно-

шении которых собраны и закодированы такие данные, как дети, родившиеся живыми, дети, живущие на момент переписи или дети, родившиеся в прошлом году), то вводятся все элементы. Если вводится информация о рождаемости, то компьютерная программа редактирования может определить, какой элемент или набор элементов достоверен, а какой необходимо изменить. Если программа редактирования определяет, что данным лицом является взрослая женщина, но информация о рождаемости отсутствует, то в этом случае следует применить метод динамической интерполяции или другие соответствующие средства для того, чтобы получить информацию о рождаемости для таблиц. Если фактическая информация утеряна из-за схем пропуска, то группа редактирования должна решить, компенсирует ли эта потеря повышение эффективности и скорости. При наличии схем пропуска операторы по вводу данных могут идти назад по экрану до соответствующей позиции, требующей исправления. Хотя оператор по вводу данных теряет какое-то время, проходя через элементы, которые не вводятся с клавиатуры, при такой форме ввода данных несоответствие между полом, возрастом и рождаемостью лучше устранять во время редактирования, чем во время ввода.

б) Реверсивный ввод данных с клавиатуры со схемами пропуска

657. Второй метод реверсивного ввода данных предусматривает наличие схем пропуска. Если группа редактирования настаивает на схемах пропуска, как правило, для того, чтобы представить, как счетчики собирали эти данные, то клавиатурный ввод будет проходить легче и быстрее, если схемы пропуска будут удобны в обращении и если операторы по вводу данных быстро осваивают схемы ввода. Если схемы пропуска очень сложные, операторы по вводу данных могут запутаться и постоянно вводить данные неправильно. Ввод со схемами пропуска будет проходить более эффективно, если использовать ограниченные схемы, которые покрывают большие разделы вводимых записей.

658. Группа редактирования должна определить соответствующие схемы пропуска для проведения переписи или обследования у себя в стране. Например, имеет смысл пропустить все данные о занятости для детей, то есть лиц младше определенного в стране возраста для потенциальной занятости. Зачастую это составляет половину демографических элементов, поэтому целесообразно пропускать их в отношении детей, за исключением особых ситуаций, например, когда возраст детей уже приближается к определенному, или когда страна может быть заинтересована в детском труде.

659. Группа редактирования решает по каждому элементу отдельно, какие из них будут включаться в те

или иные возрастные группы. Сотрудники могут сгруппировать эти элементы, с тем чтобы легче было пользоваться схемами пропуска.

660. Не всегда легко принять четкие решения относительно схем пропуска. Например, рассмотрим следующую последовательность:

-
1. Какое гражданство данного лица?
Родился в этой стране (элемент 3)
Натурализованный
Не гражданин
 2. Год въезда данного лица?
 3. СЛЕДУЮЩИЙ ЭЛЕМЕНТ
-

661. Схему пропуска можно составить, чтобы переходить от 1 к 3, то есть, пропускать элемент о годе въезда для тех, кто родился в данной стране. Однако операторы нарушают схему пропуска или потому, что счетчик или кодировщик делает ошибку, или из-за неправильного ввода данных с клавиатуры. Это связано со многими факторами, включая уровень мастерства операторов по вводу данных, уровень культуры, компоновку переписного листа и формат экрана. Группы редактирования часто работают вместе, чтобы определить, обоснованной ли является схема пропуска в таких случаях, как этот.

3. Интерактивный ввод данных с клавиатуры

662. Интерактивный ввод данных можно использовать при проведении переписи, но более целесообразно применять этот метод при проведении обследований, особенно небольших обследований, когда указанные элементы могут повлиять на результаты обследования. Интерактивный ввод данных может включать в себя ручную или автоматическую корректировку в зависимости от наличия информации, которую можно использовать для внесения изменений или исправлений.

663. Рассмотрим пример небольшого обследования. Для небольших обследований значение имеет каждый ответ. Если страна проводит однопроцентное выборочное обследование, например каждый ответ представляет сто человек, жилищных единиц или сельскохозяйственных владений. Несколько недостоверных или несогласующихся ответов могут существенно повлиять на результаты обследования. В этом случае демографы и другие специалисты в области общественных наук, как правило, хотят вести существенный контроль за обработкой данных.

664. Контроль можно устанавливать по-разному. Демографы и другие специалисты могут сами вводить данные с клавиатуры, проверяя посторонние, недостоверные или несогласующиеся ответы по мере их появления, и используя информацию, записанную на бланках для сбора данных. Они могут, зачастую незамедлительно, разрешить противоречия, проблемы, связанные с неправильным присвоением кодов, или другие несоответствия для непосредственно собранной информации. Иногда они могут отправить назад неполностью заполненные или недостоверные переписные листы. Такой вид интерактивного ввода данных дает наилучшие результаты, поскольку демограф работает и как оператор по вводу данных, но этот метод намного дороже и не так много стран могут себе его позволить.

665. Группы редактирования могут разработать очень подробные правила редактирования для того, чтобы определить, что должны делать операторы по вводу данных в каждом конкретном случае, который возникает во время ввода данных с клавиатуры. В отношении каждого недостоверного кода, если этот вопрос не разрешается, они могут решить, что должен вводить оператор. Группа редактирования может рассматривать случаи, которые не охватывают подробными правилами, и могут видоизменять правила (хотя при этом возникает опасность возникновения несоответствий между первой частью и последующими частями процесса ввода данных с клавиатуры).

666. Схема пропуска, которая играет важную роль при реверсивном вводе данных здесь также имеет важное значение. Как и при реверсивном вводе, операторы должны знать и освоить все схемы пропуска, которые применяются в данном случае. Как упоминалось выше, схемы пропуска могут увеличить скорость ввода данных с клавиатуры, но обычно с некоторой потерей качества. Для интерактивного ввода данных общий практический подход состоит в том, что чем меньше пропусков, тем выше качество.

В. ПРОВЕРКА ИНСТРУКЦИЙ ПО ВВОДУ ДАННЫХ С КЛАВИАТУРЫ

667. Разработав инструкции по вводу данных с клавиатуры, национальные бюро переписи/статистические управления должны организовать проверку инструкций по вводу данных с клавиатуры операторами по вводу данных, прежде чем они приступят к практической операции, независимо от того, используется или нет реверсивный ввод данных. После проверки инструкций по вводу данных из системы можно вывести дефекты и добиться оптимального режима ввода данных.

С. ПРОВЕРКА

668. Национальное бюро переписи/статистическое управление должно также решить вопрос о соответствующем уровне проверки. Многие эксперты рекомендуют проводить стопроцентную проверку. В этом случае все элементы снова вводятся с клавиатуры для того, чтобы убедиться, что собранные данные — это те данные, которые вводятся в машину для компьютерной обработки. Однако во многих случаях общая проверка нереальна, потому что у страны нет времени вновь вводить с клавиатуры все данные или не хватает финансовых или людских ресурсов. Кроме того, если проверенный коэффициент ошибок при вводе данных с клавиатуры очень низкий и операторы по вводу данных совершают весьма незначительное количество ошибок, то, вероятно, полная проверка и не обязательна.

1. Зависимая проверка

669. Методы проверки могут быть и зависимыми или независимыми. При зависимой проверке операторы по вводу данных вводят данные, уже введенные ранее

другими сотрудниками. Если количество нажатий клавиш разное, то программный пакет сообщает об этом оператору по вводу данных, и в зависимости от программы оператор или игнорирует предыдущие данные, или отмечает несоответствие. Поскольку данные вводятся с первоначальных переписных листов, то, как правило, оператор может сам на основе имеющейся информации решить, было ли первоначальное введение данных ошибочным.

2. Независимая проверка

670. При независимой проверке операторы могут начать введение данных с нуля; они создают полностью независимый файл вводимых данных, используя первоначальные переписные листы. Два получившихся в результате файла, набор первоначально введенных данных и набор данных, введенных при проверке, затем сопоставляются между собой с помощью компьютерной программы для того, чтобы проверить их на несоответствия. Предположительно при этом могут использоваться некоторые ручные операции для исправления недостоверных или несогласующихся нажатий клавиш.

Приложение IV

ВЫБОРОЧНЫЕ РАБОЧИЕ БЛОК-СХЕМЫ

671. Одной из задач группы редактирования является разработка относительной структуры для переменных, используемых при редактировании. Блок-схемы дополняют письменные спецификации для редактирования, помогая выявлять различные связи между переменными. Эти спецификации для относительных связей помогают специалистам, занимающимся содержанием переписи, и специалистам по обработке данных отчетливо представлять себе процесс редактирования и способствовать общению между двумя группами.

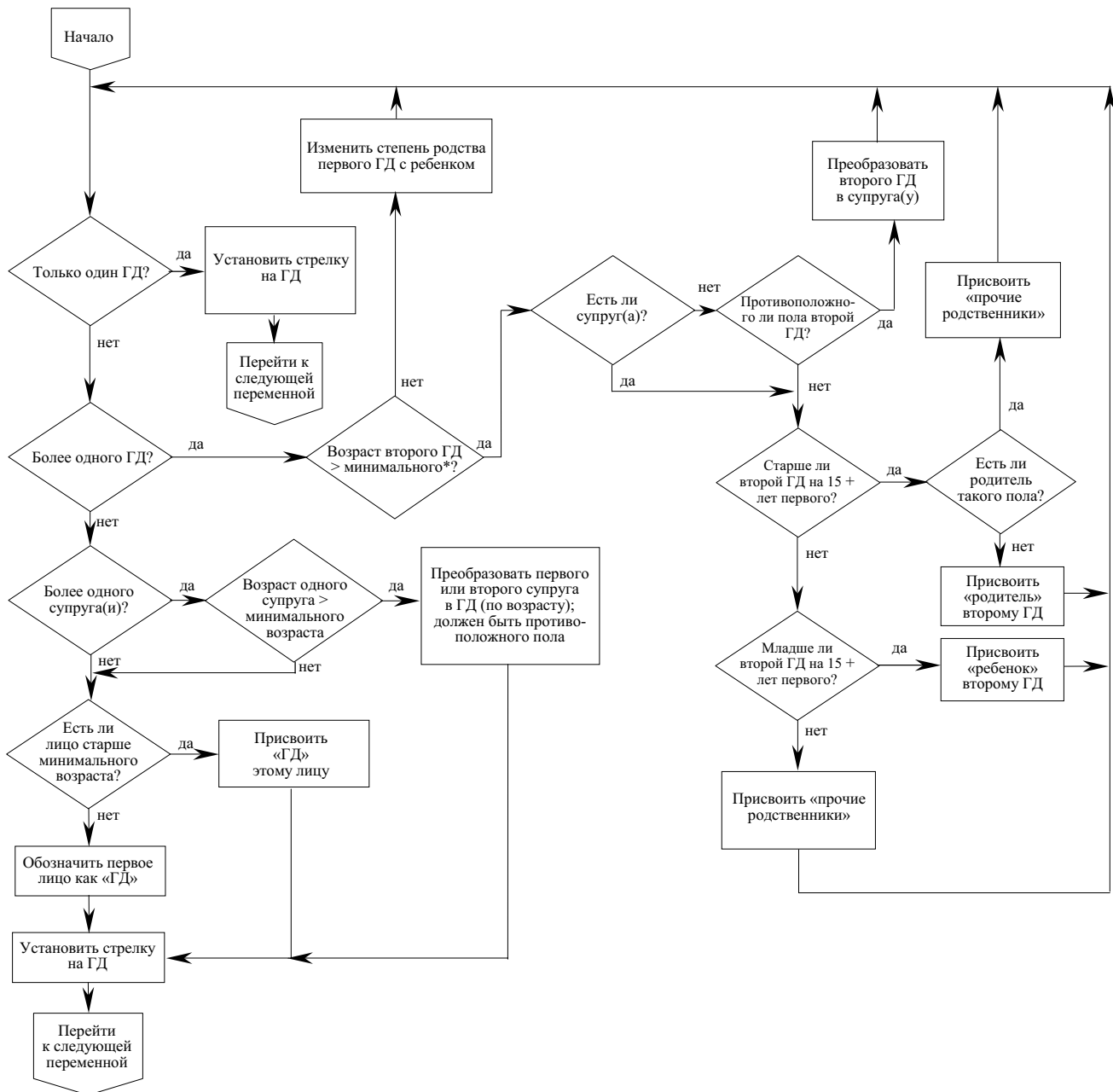
672. На следующих страницах представлены три выборочных блок-схемы:

- a) блок-схема для главы домохозяйства;
- b) блок-схема для определения, есть ли супруга (супруг) в домохозяйстве;
- c) блок-схема для редактирования показателя пола главы домохозяйства и супруги (супруга).

Эти выборочные блок-схемы представлены лишь для примера и к ним нужно и относиться соответственно. Группа редактирования может пожелать внести изменения в выборочные блок-схемы по мере необходимости исходя из ситуации в стране.

673. В целях редактирования блок-схемы должны быть составлены для каждой переменной в переписи. Группа редактирования должна вести совместную работу с другими специалистами по составлению блок-схем, а специалисты по обработке данных должны использовать их вместе со спецификациями редактирования в целях разработки компьютерных программ для редактирования данных переписи. Блок-схемы и спецификации для редактирования должны быть должным образом задокументированы с тем, чтобы их можно было использовать при обработке данных в ходе дальнейших переписей и обследований.

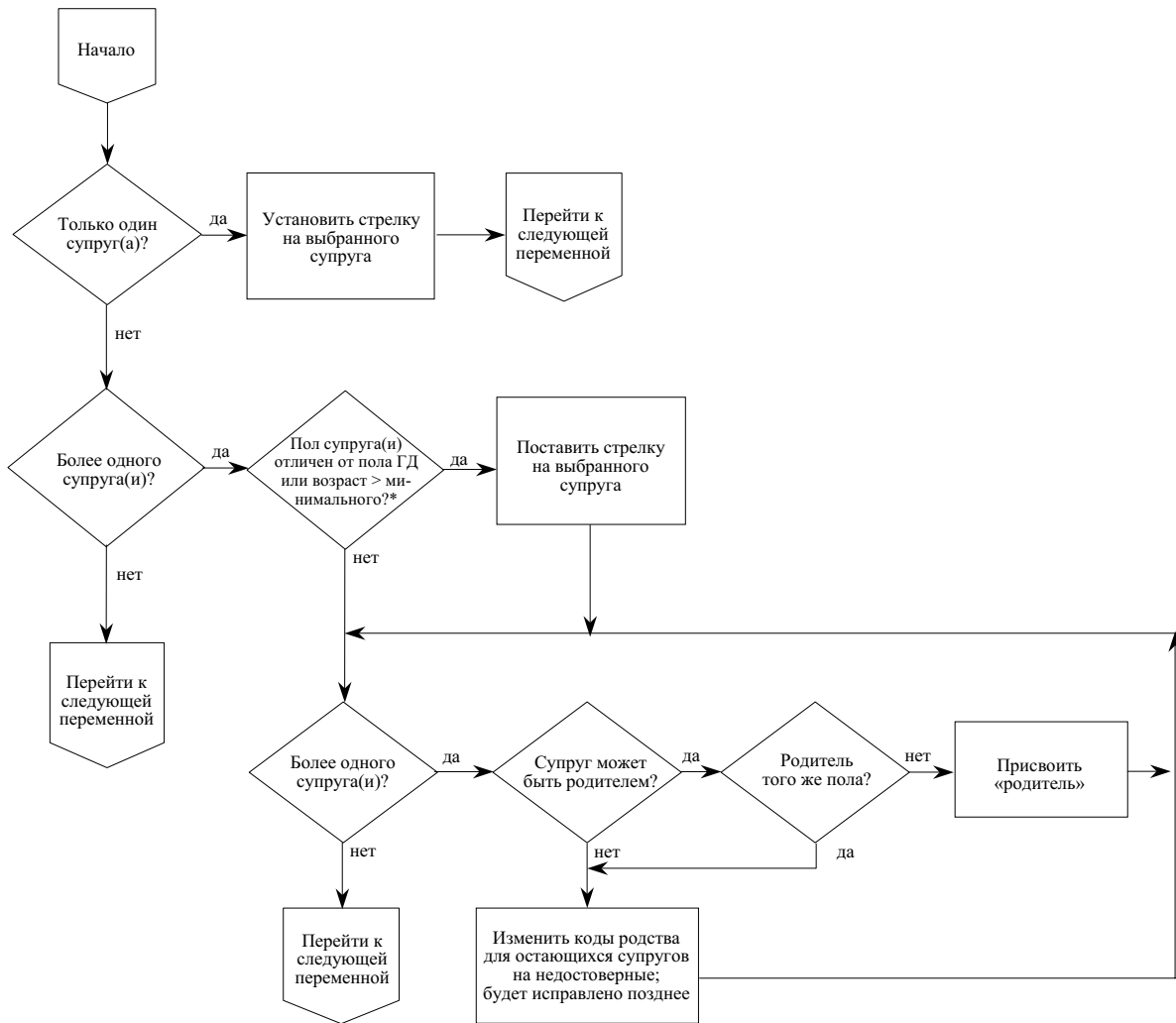
Рисунок А.IV.1. Выборочная блок-схема для определения главы домохозяйства (ГД)



ПРИМЕЧАНИЯ: ГД — глава домохозяйства.

* минимальный возраст должен быть определен группой редактирования.

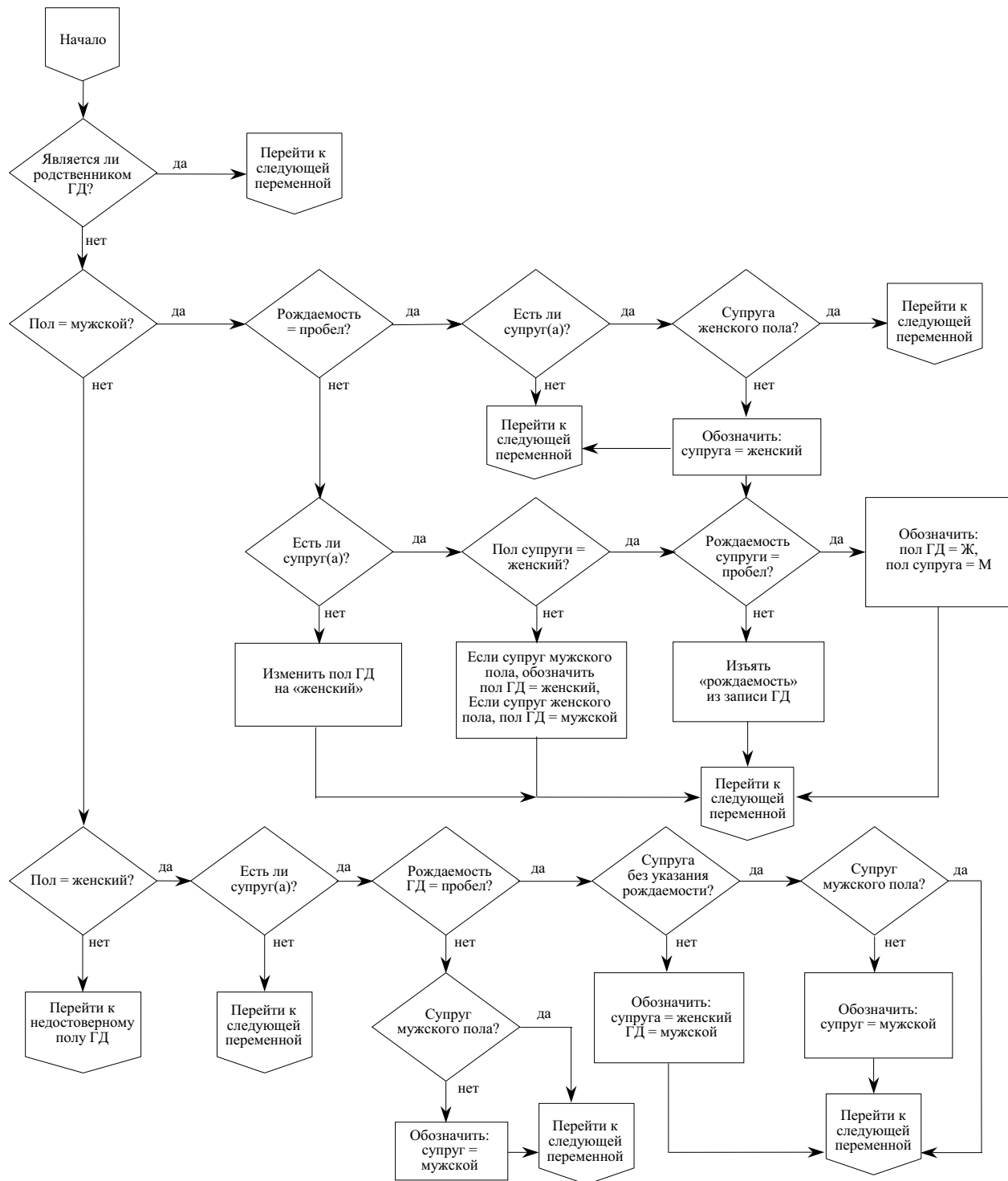
Рисунок А.IV.2. Выборочная блок-схема для определения наличия супруга(и) в домохозяйстве



ПРИМЕЧАНИЯ: ГД — глава домохозяйства.

* минимальный возраст должен быть определен группой редактирования.

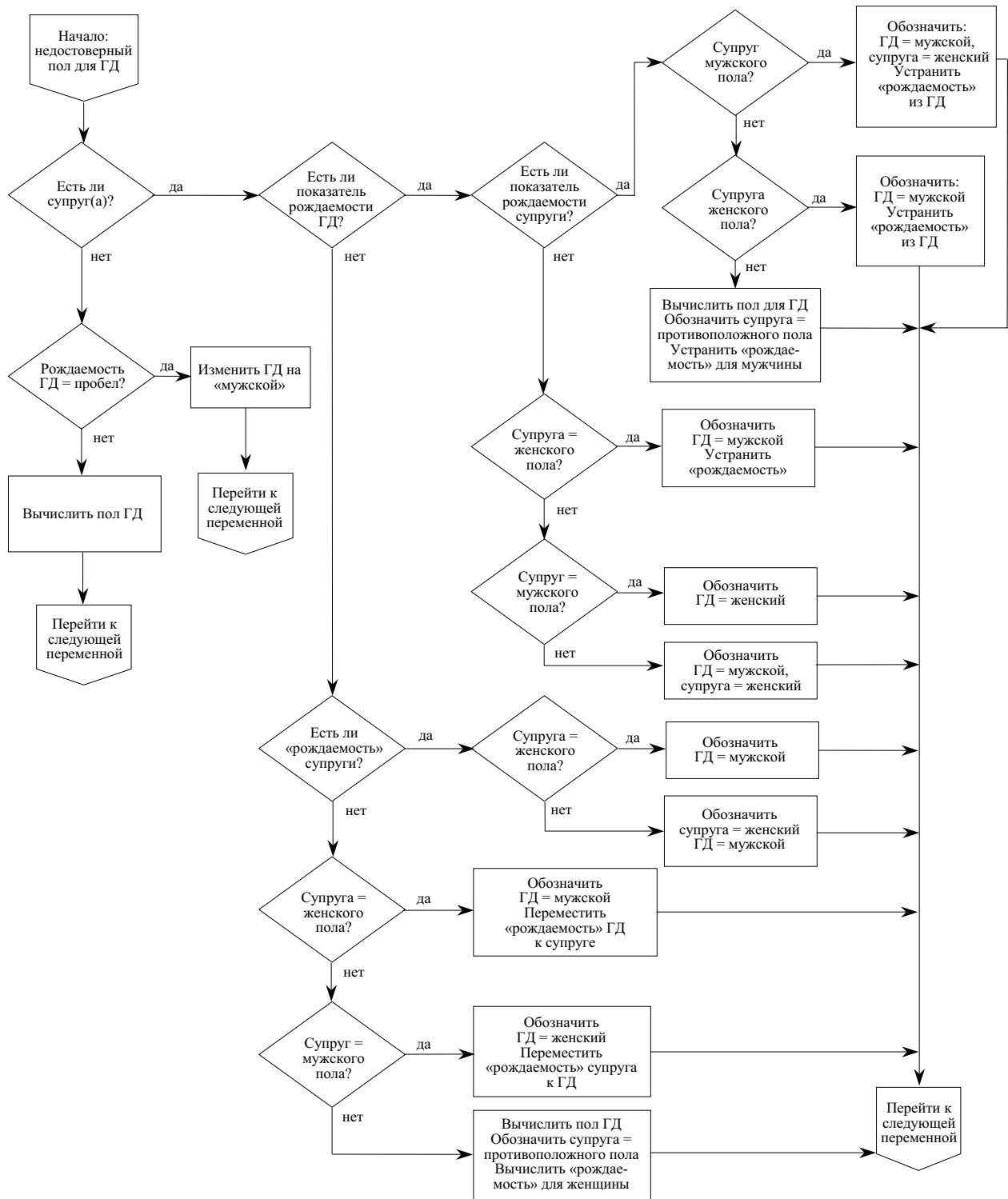
Рисунок А.IV.3. Выборочная блок-схема для редактирования переменной для пола главы домохозяйства и супруга(и)



ПРИМЕЧАНИЯ: ГД — глава домохозяйства.

* минимальный возраст должен быть определен группой редактирования.

Рисунок А.IV.3. (продолжение)



Приложение V

МЕТОДЫ ИНТЕРПОЛЯЦИИ

674. Сейчас разработан целый ряд методов интерполяции. Большинство методов, описанных ниже, рассматриваются в работах Кальтона и Каспржыка (Kalton and Kasprzyk, 1982, 1986); Санде (Sande, 1982); и Сёрндаля, Свенссона и Вретмана (Särndal, Swensson and Wretman, 1992).

675. Методы интерполяции можно подразделить на стохастические или детерминистические в зависимости от степени случайности в исчисляемых данных.

676. **Детерминистический метод интерполяции** включает в себя дедуктивную интерполяцию; методы интерполяции, основанные на модели, такие как средняя интерполяция и регрессионная интерполяция; и, в необходимых случаях, интерполяцию по ближайшим соседним величинам.

677. **Метод дедуктивной интерполяции** позволяет вывести с определенностью недостающую или несогласующуюся величину. В его основе часто лежит схема ответов, даваемых на другие вопросы в переписном листе.

678. В большинстве случаев метод интерполяции применяется для замены величины, которая не выглядит достоверной. Некоторые общие процедуры интерполяции излагаются ниже.

679. За исключением алгоритмов динамической интерполяции по единому донору, описанные ниже методы связаны с интерполяцией одного элемента за один раз. Так, в каждом классе интерполяции содержащиеся в записи элементы рассматриваются последовательно один за другим. Как правило, при этом рассматриваются только те подлежащие редактированию сведения, которые относятся исключительно к данному элементу или к небольшому набору тесно взаимосвязанных переменных. Поскольку могут быть как явные, так и косвенные редакционные изменения, увязывающие данный элемент (данные элементы) с другими элементами, которые предстоит рассмотреть позднее, эта процедура может дать исчисленную величину при обработке тех показателей, которые редактируются в данный момент, без редактирования других показателей, которые должны рассматриваться позднее. Только когда учтен весь набор редактируемых показателей, включая все подразумеваемые показатели, можно быть уверенным в том,

что все исчисленные величины пройдут все этапы редактирования. Подразумеваемый показатель, подлежащий редактированию, — это величина, которую можно вывести логическим путем, совмещая две или более редактируемых величины.

680. В последующем описании понятие «записи, прошедшие редактирование» относится к тем записям, которые прошли все этапы редактирования данного элемента (данных элементов). Понятие «записи, не прошедшие редактирование» относятся к тем записям, которые не прошли, по крайней мере, один этап редактирования в связи с данным элементом (данными элементами).

681. **При суммарной средней интерполяции** среднее значение элемента присваивается по всем прошедшим редактирование записям недостающему или несогласующемуся элементу по всем записям, отредактировать которые не удалось. При таком методе можно получать приемлемые точечные оценки, но он менее привлекателен, если оценки дисперсии вычисляются с использованием стандартной формулы оценки дисперсии. Оценки дисперсии могут быть серьезно занижены, если только не брать очень низкий коэффициент интерполяции или не изменить формулу оценки дисперсии, чтобы учесть производимые исчисления.

682. **При сведении интерполяции по классам** используются классы интерполяции, определяемые для создания групп записей с определенной степенью сходства. В каждом классе средняя элемента вычисляется по прошедшим редактирование записям для недостающего или несогласующегося элемента по всем записям, отредактировать которые не удалось. Этот метод очень похож на метод суммарной средней интерполяции, однако последствия для распределения данных и проблемы с оценками дисперсии здесь, скорее всего, будут менее серьезными.

683. **При регрессионной интерполяции**, более известной как **интерполяция на основе моделей**, используются данные из прошедших редактирование записей для вычисления путем регрессии переменной, которую необходимо интерполировать по ряду предикторных переменных. Предикторами в регрессии могут быть элементы из переписного листа или вспомогательные переменные. Затем используется регрессионное урав-

нение для исчисления значения для недостающих или несогласующихся элементов. Это особый случай, основанный на моделях интерполяции. Такой метод часто используется в связи с непрерывными переменными при обследовании деловой активности, когда данные о предыдущем событии зачастую могут предсказать удовлетворительным образом значения, касающиеся текущего события.

684. При **интерполяции по ближайшим соседним единицам** или **сочетании по функции расстояния** значение для не прошедшей редактирование записи присваивается по «ближайшей» записи, прошедшей редактирование, где понятие «ближайшая» определяется с использованием функции расстояния в исчислении по другим известным переменным. Этот метод можно применять в классах интерполяции. Он обычно считается приемлемым для непрерывных переменных, но его можно применять и с нечисловыми переменными.

685. **Методы стохастической интерполяции** включают в себя регрессию или любой другой детерминистический метод с прибавлением случайных остатков и использованием методов статической или динамической интерполяции.

686. Для каждого детерминистического метода есть свой стохастический аналог. Этого можно добиться, прибавляя случайный остаток от соответствующего распределения к исчисленному значению из детерминистической интерполяции. Такая процедура поможет лучше сохранить структуру частотности файла данных. Калтон и Каспржык (Kalton and Kasprzyk, 1986) рассматривают некоторые подходы к применению такого метода.

687. При **динамической и статической интерполяции** делается попытка создать более реальную изменчивость исчисляемых значений, чем это можно сделать при детерминистических методах. При динамической интерполяции недостающие или несогласующиеся величины заменяются значениями, отобранными (наугад) из прошедших редактирование записей текущего обследования или переписи. При статической интерполяции исчисления производятся по другим источникам, часто по прошлым данным, таким как более ранние события того же обследования или переписи. Есть целый ряд различных форм динамической и статической интерполяции.

688. Самой простой формой динамической интерполяции является **случайная суммарная интерполяция**. Для каждой не прошедшей редактирование записи наугад выбирается одна прошедшая редактирование запись из всех записей такого рода, и ее указанное значение для данного элемента приписывается вместо не прошедшей редактирование записи.

689. При **случайной интерполяции** для ограничения случайного отбора донорских записей в набор, который имеет некоторое сходство с записью, подлежащей интерполяции, опять используются классы интерполяции.

690. При **логической динамической интерполяции** также используются классы интерполяции, и преимущество этого метода заключается в том, что одного прохода через файл данных достаточно для завершения процесса интерполяции. Процедура начинается со статического значения для каждого класса интерполяции, и записи в файле данных рассматриваются по очереди. При обнаружении не прошедшей редактирование записи ее значение для данного элемента заменяет хранящееся в памяти значение для класса интерполяции. При обнаружении не прошедшей редактирование записи ее недостающее или несогласующееся значение заменяется хранящейся в памяти величиной. Число классов интерполяции не может быть чрезмерно большим, поскольку необходимо, чтобы доноры имелись в каждом классе интерполяции. Если записи в файле данных расположены в случайном порядке, этот метод будет почти эквивалентным случайной интерполяции в классах. Недостаток такой процедуры заключается в том, что часто приходится многократно использовать доноров, и это может отрицательно повлиять на распределение элемента и оценки дисперсии.

691. **Иерархическая динамическая интерполяция** — это усиление логической динамической интерполяции, где используется большое количество классов интерполяции. Когда в первоначальном классе интерполяции нельзя найти донора, классы стягиваются в иерархическом порядке, пока донор не находится.

692. Задача алгоритмов **динамической интерполяции по одному донору** заключается в том, чтобы исчислять данные для не прошедшей редактирование записи по одному донору. Следовательно, они позволяют производить совместную интерполяцию всех значений элемента в записи, которая при редактировании была определена как сомнительная. На практике, зачастую, ставится цель использовать одного донора для каждого раздела тесно взаимосвязанных переменных в одной записи. Существенное преимущество такого подхода заключается в том, что он позволяет лучше сохранять не только маргинальные распределения, как в случае вышеупомянутой динамической интерполяции, но и эмпирическую плотность совместных распределений. Еще одно преимущество методов динамической интерполяции по одному донору заключается в том, что при их использовании становится меньше проблем с исчислением величин, не поддающихся редактированию в последующих разделах переменных. С точки зрения методики динамической интерполяции по одному донору прошедшая редактирование запись — это запись,

которая прошла все этапы редактирования по отношению к данному разделу. Не прошедшая редактирование запись — это запись, которая не прошла, по крайней мере, один из этих этапов.

693. **Метод интерполяции и редактирования Фелледжи–Холта** (Fellegi and Holt, 1976) учитывает одновременно все приемы редактирования. Главная особенность метода интерполяции и редактирования Фелледжи–Холта заключается в том, что правила интерполяции выводятся из соответствующих приемов редактирования без явной спецификации. В отношении каждой не прошедшей редактирование записи сначала локализуется ошибка с определением минимального набора переменных для исчисления, а также приемлемый диапазон величин для исчисления, а затем производится интерполяция. В большинстве операций единый донор выбирается из прошедших редактирование записей путем согласования по другим переменным, используемым программой редактирования, но не требующим интерполяции. При такой методике отыскивается один точный элемент согласования, который можно расширить с тем, чтобы учесть другие переменные, не задействованные явным образом в редактировании. Иногда подходящего донора найти не удается, и тогда необходимо использовать метод интерполяции по умолчанию.

694. **Новая методология интерполяции (НМИ)** (Bankier, Luc, Nadeau and Newcombe, 1996; Bankier, Lachance and Poirier, 1999) похож на метод Фелледжи–Холта тем, что в нем учитываются одновременно все этапы редактирования, не уточняются конкретно приемы интерполяции и исчисления происходят по одному донору. При этом для каждой не прошедшей редактирование записи определяются действия по интерполяции с минимальными изменениями в зависимости от имеющихся потенциальных доноров. Такой подход гарантирует наличие донора. В отличие от метода Фелледжи–Холта при использовании НМИ сначала ведется поиск доноров, а затем определяются действия по интерполяции с минимальными изменениями. При такой методологии поиск доноров ведется путем согласования с использованием всех переменных (включая те, которые, возможно, придется исчислять), задействованных в редактировании, и он может быть успешным, если производить согласование числовых переменных при согласовании большинства, но не обязательно всех других переменных. При этом определяются действия с учетом каждого потенциального донора и такие приемы, которые позволяют производить интерполяцию с минимальными изменениями. При такой методике учитываются также ближайшие действия по интерполяции с минимальными изменениями; иногда они могут дать более достоверные исчисленные записи. И, наконец,

наугад выбирается одно из действий по интерполяции с минимальными изменениями и по ближайшей интерполяции с минимальными изменениями и производится исчисление.

695. Хотя метод Фелледжи–Холта и НМИ требуют основательной компьютеризации, эффективный алгоритм для этого имеется, и поэтому с современными компьютерами их вполне можно применять. Особенно это касается НМИ, при которой можно легко решать более или менее крупные проблемы интерполяции и редактирования по сравнению с методикой Фелледжи–Холта.

696. Все вышеперечисленные методы интерполяции дают одно исчисленное значение для каждой недостающей или несогласующейся величины. Все эти методы в какой-то степени искажают обычное распределение значений для данного элемента и могут привести к неправильным оценкам дисперсии, если используются стандартные формулы оценки дисперсии. Степень искажения может быть совершенно разной в зависимости от количества произведенных исчислений и используемого метода.

697. **Многомерная интерполяция** — это метод, предложенный Рубином (Rubin, 1987), при котором эта проблема решается путем исчисления несколько раз (m) каждого значения, требующего интерполяции. Затем из собранного набора данных для соответствующего элемента можно получить оценки m . При этом получается одна комбинированная оценка вместе с совокупной оценкой дисперсии, которые с неопределенностью указывают на то, какую величину исчислять. Недостаток метода многомерной интерполяции заключается в том, что он требует больше усилий по обработке данных и исчисления оценок.

698. В большинстве систем интерполяции используется сразу несколько методов интерполяции; как правило, по возможности, используется дедуктивная интерполяция, а за ней применяется одна или более других процедур. Большинство национальных статистических управлений используют некоторую форму динамической интерполяции при редактировании и исчислении данных переписи. В настоящее время наиболее широкое применение находит логическая динамическая интерполяция и метод Фелледжи–Холта. Из тех национальных статистических управлений, которые в настоящее время пользуются методом Фелледжи–Холта, одно переходит на НМИ, а ряд других рассматривают этот вопрос. Однако с учетом предполагаемого круга основных читателей в данном *Руководстве* говорится, главным образом, о форме логической динамической интерполяции.

Приложение VI

ПАКЕТЫ КОМПЬЮТЕРНОГО РЕДАКТИРОВАНИЯ

699. При наличии относительно недорогих микрокомпьютеров страны могут тщательно и своевременно редактировать данные переписей и обследований. До недавнего времени каждая страна вынуждена была сама составлять собственную программу редактирования, что требует столь ценного времени на отладку и обработку. Однако с появлением стандартного пакета компьютерного редактирования странам сейчас значительно легче удовлетворять свои потребности в редактировании, и это требует меньше знаний в области обработки данных.

700. Одно из преимуществ применения пакета компьютерного редактирования заключается в том, что при правильном использовании данные будут согласовываться между собой и будут достоверными, поэтому появится возможность составлять таблицы более своевременно. Для написания программ редактирования можно использовать многие пакеты программного обеспечения, например, ССА и СПОН, а также другие языки высокого уровня. Кроме того, страна может использовать один из пакетов программного обеспечения, составленных исключительно для целей редактирования данных переписей и обследований. Для большинства стран общее редактирование идет быстрее с пакетом, чем со специально составленными программами, поскольку пакет не требует такого уровня знаний в области обработки данных, как при работе со специально составленными программами.

701. Хороший пакет компьютерного редактирования обеспечит общение между специалистами, занимающимися содержанием переписи, и программистами. Хороший пакет редактирования должен позволить поставить описательный код или символический код рядом с кодом программирования, если сам код программирования не будет очевиден или прозрачен для специалистов, занимающихся содержанием переписи. Демографы и другие специалисты должны иметь возможность проходить через программу строка за строкой и четко понимать, что делает программа.

702. Любой программный пакет редактирования, который страна могла бы использовать, должен давать возможность производить и составлять отчеты по различным проверкам, испытаниям и исчислениям, необходимым для редактирования данных переписи. Эти

требования действуют даже тогда, когда специалисты по обработке данных составляют программы редактирования по заказу. Пакет должен отвечать следующим требованиям:

- a) давать возможность вводить данные с клавиатуры и/или проверять вводимые данные. Этот пакет должен обеспечивать возможность добавлять схемы пропуска. Например, группа редактирования может решить, что сведения о рождаемости в отношении лиц мужского пола должны быть пропущены;
- b) выполнять структурное редактирование, которое позволит определить, есть ли фактически те типы записей, которые должны быть, в том числе, например, запись, касающаяся жилищ для каждого порядкового номера;
- c) производить записи, если они пропущены, и/или добавлять весовые коэффициенты к существующим записям;
- d) определять, что каждая переменная имеет достоверное значение;
- e) хранить все или часть уже отредактированных записей;
- f) проверять на согласованность две или более характеристик в одной и той же записи или между записями. При сокращенном варианте должна производиться проверка на согласованность внутри домохозяйства, сверка ответов с ответами членов предыдущего домохозяйства. Исчислять значения методом динамической интерполяции, если страна предпочтет использовать этот метод;
- g) использовать несколько значений в записи или из многократных записей для построения производной переменной и вводить производную переменную в соответствующую запись;
- h) выявлять и устранять двойные записи;
- i) составлять дневник ошибок и изменений по небольшому географическому району.

Обычно с помощью пакетов или программ редактируется одна запись за раз, другие (не текущие) пакеты позволяют также производить проверку между записями, особенно с данными о жилищных единицах.

ГЛОССАРИЙ

Массив	Набор числовых величин. Иногда называется матрица и может использоваться для хранения периодически повторяющихся числовых данных.
Журнал редактирования	Метод отслеживания изменений в значениях поля и причин и источников каждого изменения. Журнал редактирования, как правило, начинают вести после завершения первоначального опроса.
Автоматизированное редактирование	Корректирование ошибок в данных с помощью компьютера без вмешательства человека. Один из аспектов автоматизированного редактирования данных (Pierzchala, 1995).
Редактирование между записями	Редактирование, производимое на полях, где есть более одной записи обследования. Статистически редактируемые данные являются примером редактирования между записями, потому что распределения составляются на совокупности полей по всем записям обследования (Pierzchala, 1995).
Средняя интерполяция по классам	Используются классы интерполяции, определяемые для создания групп записей с определенной степенью сходства.
Достоверная запись	Запись, в которой нет недостающих величин и которая проходит все этапы редактирования (Pierzchala, 1995).
Перечень кодов	Перечень всех разрешенных (допустимых) величин элемента данных.
Статическая интерполяция	Первоначальная статическая матрица; база корректирования, для которой все элементы даются до начала корректирования и не меняются в процессе корректирования. Например, базой корректирования может быть указание предыдущего года. С модифицированной статической матрицей можно корректировать статические величины в соответствии с текущей (возможно агрегированной) информацией.
Полный набор редактируемых данных	Объединение явных и косвенных редактируемых данных. Необходим для создания возможных участков для интерполяции (если кто-то захочет произвести интерполяцию в целях редактирования) (Pierzchala, 1995).
Редактирование для проверки на согласованность	Проверка определяющих отношений, например, частей, прибавляемых к общему количеству или «земли с урожаем», всегда меньше, чем «засаженные земли» (Pierzchala, 1995).
Сбор данных	Процесс, когда собранные данные переводятся в машинно-считываемую форму. В подмодулях программного обеспечения, куда вводятся данные, часто проводятся элементарные контрольные проверки.
Дедуктивная интерполяция	Метод, когда можно с определенностью вывести недостающую или несогласующуюся величину, часто на основе структуры ответов, даваемых на другие вопросы переписного листа.

Детерминистическое редактирование	Редактирование, которое в случае нарушения указывает на ошибку в данных с вероятностью до одного. Например: возраст = 5 и состояние = мать. В отличие от стохастического редактирования (Pierzchala, 1995).
Детерминистическая интерполяция	Такая ситуация возникает, когда только одна величина в поле дает запись, удовлетворяющую всем редактируемым данным. Возникает в некоторых случаях (когда часть от целого не прибавляется к целому). Первое решение необходимо проверить при автоматизированном редактировании и интерполяции данных обследования (Pierzchala, 1995).
Функция расстояния	Для числовых данных функция, определяемая по согласующимся переменным в записях получателя (кандидата) и донора и используемая для квантификации понятия сходства. Используется для нахождения согласующихся записей при динамической интерполяции (Pierzchala, 1995).
Согласование функции расстояния	Приписывает значение элемента для не прошедшей редактирование записи и из «ближайшей» прошедшей редактирование записи, где «ближайшая» определяется с использованием функции расстояния в исчислении на основе других известных переменных.
Донорская интерполяция	Метод, при котором каждая запись, требующая интерполяции, соединяется попарно, запись о получателе или кандидате с одной записью из определенной донорской совокупности, как, например, при динамической интерполяции (Pierzchala, 1995).
Редактирование (определение 1)	Логические ограничения величин, которые может приобретать каждая переменная (Pierzchala, 1995).
Редактирование (определение 2)	Приемы, с помощью которых определяются запрещенные комбинации ответов (Pierzchala, 1995).
Контрольный журнал	См. «журнал редактирования».
Явные приемы редактирования	Приемы, явно написанные специалистом по содержанию переписи (в отличие от явных приемов редактирования с косвенными приемами) (Pierzchala, 1995).
Не прошедшие редактирование записи	При редактировании и интерполяции записи, которые не прошли, по меньшей мере, одно редактирование того или иного элемента.
Метод Фелледжи–Холта для автоматического редактирования	Метод автоматического редактирования, при котором меняется наименьшее число элементов данных и модель Фелледжи–Холта используется для определения приемлемых наборов величин или диапазонов для исчисляемых элементов. Может применяться последовательная или одновременная интерполяция методом статического или динамического исчисления.
Система Фелледжи–Холта	Касается допущений и целей редактирования и интерполяции, изложенных Фелледжи и Холтом в их работе 1976 года <i>Journal of the American Statistical Association</i> . Основная особенность модели Фелледжи–Холта заключается в том, что она показывает, что косвенные приемы редактирования необходимы для того, чтобы набор величин в полях данных, которые не всегда интерполируются, давал конечные (исчисленные) записи, которые удовлетворяли бы всех.

Нисходящий ввод данных	Стиль ввода данных, когда компьютер не обнаруживает ошибки при вводе данных, позволяя оператору вводить данные быстро и эффективно.
Признак	Признак — это переменная, используемая для того, чтобы отмечать полезную информацию от другой переменной или переменных. Например, если меняется элемент с недостоверного на достоверный, признак можно использовать для того, чтобы отметить первоначальную информацию или отметить, что значение элемента изменилось.
Блок-схема	Изображение в виде диаграммы всех функций, которые должны быть выполнены.
Ручное редактирование	Редактирование, осуществляемое людьми до ввода данных в компьютер (см. «редактирование вручную») (Pierzchala, 1995).
Восходящий ввод данных	Стиль ввода данных, при котором компьютер обнаруживает ошибки в данных по мере их ввода, позволяя оператору немедленно вносить исправления (Pierzchala, 1995). См. «интерактивный ввод данных с клавиатуры».
Динамическая интерполяция	Метод исчисления, при котором донорские записи берутся из текущей колоды выборочных данных (в отличие от метода статической интерполяции, при котором донорская запись берется из данных прошлой переписи) (Pierzchala, 1995).
Косвенное редактирование	Неуказанные редактируемые данные, выводимые логическим путем из явных приемов редактирования, написанных специалистом, занимающимся содержанием переписи (Pierzchala, 1995).
Интерполяция	Присвоение значения полю в связи с отсутствием ответа или замена указанного значения, которое по определению не согласуется с набором редактируемых данных (Pierzchala, 1995).
Интерактивный ввод данных с клавиатуры	Стиль ввода данных, при котором компьютер обнаруживает ошибки в данных, по мере их ввода, позволяя оператору тут же вносить исправления. См. «восходящий ввод данных».
Внутренняя согласованность	Этот термин относится к отношениям между переменными для данной выборочной единицы и является основанием для редактирования в большинстве процедур обследования (Pierzchala, 1995).
Макроредактирование	Обнаружение отдельных ошибок путем (1) проверки агрегированных данных или (2) проверок, используемых по отношению ко всей совокупности записей. Эти проверки основаны на оценках (Granquist, 1987; Pierzchala, 1995).
Редактирование вручную	Редактирование, осуществляемое людьми до ввода данных в компьютер (см. «ручное редактирование») (Pierzchala, 1995).
Согласование	При проведении динамической интерполяции акт согласования донорской записи с записью получателя (кандидата) (Pierzchala, 1995).
Согласующие переменные	Переменные, используемые для согласования записи получателя (кандидата) и записи донора (Pierzchala, 1995).

Микроредактирование	Традиционное редактирование данных на уровне записи. Логический антоним макроредактирования (Pierzchala, 1995).
Микро-макро редактирование	Процедура редактирования, когда подробные микроредактируемые данные заменяются комбинацией микроредактируемых и макро/статистических редактируемых данных. Микроредактируемые данные при комбинированной процедуре представлены менее подробно, чем в первом случае. Идея заключается в том, чтобы выработать приемы редактирования данных обследований не на основе подхода по принципу «уловить все несоответствия в данных», а по принципу «воздействовать на оценки» (Granquist, разные даты, Pierzchala, 1995).
Минимальный набор	Наименьший набор полей, требующих интерполяции, гарантирует, что будут пройдены все этапы редактирования (Pierzchala, 1995).
Интерполяция на основе модели	Используются данные из прошедших редактирование записей для вычисления переменной, которую необходимо интерполировать по набору предикторных переменных.
Многомерная интерполяция	Исчисляется несколько раз каждая величина, требующая интерполяции, а затем выводится оценка для данного элемента.
Многомерное редактирование	Тип статистического редактирования, где многомерное распределение используется для оценки данных и нахождения отклоняющихся значений (Pierzchala, 1995).
Интерполяция по ближайшим соседним единицам	Приписывает значение элемента для не прошедшей редактирование записи и из «ближайшей» прошедшей редактирование записи, где понятие «ближайшая» определяется с использованием функции расстояния, исчисляемой в других известных переменных.
Новая методология интерполяции	Похожа на метод Фелледжи–Холта тем, что учитывает одновременно все редактируемые данные, не уточняет явно действия по интерполяции и проводит исчисления по одному донору. НМИ просматривает каждую не прошедшую редактирование запись для определения действий по интерполяции с минимальными изменениями в зависимости от наличия потенциальных доноров.
Отсутствие ответа	Не полностью заполненный переписной лист или отсутствующий переписной лист (Pierzchala, 1995).
Отклоняющиеся значения	Величина элементов, которые выходят за некоторые границы, в зависимости от определения таких границ (Pierzchala, 1995).
Суммарная средняя интерполяция	Приписывается среднее элемента для прошедших редактирование записей вместо недостающего или несогласующегося элемента по всем не прошедшим редактирование записям.
Прошедшие редактирование записи	При редактировании или интерполяции записи, прошедшие все этапы редактирования по данным элементам.

Стрелка	Стрелка означает переменную, используемую для отметки элемента или другую переменную для последующих ссылок. Например, стрелки используются для того, чтобы отметить номера строки «главы» и «супруги(а)» для последующего использования с тем, чтобы быть уверенными, что супруги противоположного пола и оба состоят в браке.
Эксплуатация	Акт обработки больших объемов данных после устранения первоначальных «дефектов» из программы редактирования или табулирования.
Символический код	Письменные инструкции или спецификации редактирования.
Ошибки качества	Ошибки, которые могут исказить качество данных: например, систематические ошибки, которые ведут к погрешностям (Granquist 1984; Pierzchala, 1995).
Количественное редактирование	Этапы редактирования полей, измеряемых по непрерывной шкале (Pierzchala, 1995).
Случайная интерполяция по классам	Используются классы интерполяции для ограничения случайного отбора донорской записи в набор, имеющий некоторое сходство с записью, требующей интерполяции.
Случайная суммарная интерполяция	Для каждой не прошедшей редактирование записи одна прошедшая редактирование запись выбирается наугад из набора всех прошедших редактирование записей и ее указанное значение для данного элемента исчисляется для не прошедшей редактирование записи.
Запись	Магнитно-удерживаемое машинно-считываемое представление данных обследования. Обычно имеется одна запись для каждого переписного листа, хотя можно разбить данные из одного переписного листа на более, чем одну запись, например, по населению или по жилому фонду (Pierzchala, 1995).
Регрессионная интерполяция	Используются данные из прошедших редактирование записей для вычисления переменной, требующей интерполяции по ряду предикторных переменных.
Поиск	При динамической интерполяции — акт поиска донорской записи (Pierzchala, 1995).
Логическая динамическая интерполяция	Исчисление, при котором ряд переменных редактируются последовательно, и только обработанные величины используются в качестве последующих переменных динамической интерполяции.
Сходство	В числовых данных — понятие близости двух записей, основанных на предписанных согласующихся переменных. Здесь используется функция расстояния для квантифицирования этого понятия в соответствии с некоторыми критериями (Pierzchala, 1995).
Динамическая интерполяция по одному донору	Исчисляются данные для не прошедшей редактирование записи по одному донору, что позволяет проводить совместную интерполяцию всех значений элемента в записи, которая при редактировании определена как сомнительная.

Статистическое редактирование	Набор проверок, основанных на статистическом анализе данных респондента: например, соотношение двух полей лежит в пределах, определенных путем статистического анализа этого соотношения для предположительно достоверных генераторов отчетов (Greenberg and Surdi, 1984; Pierzchala, 1995).
Статистическая интерполяция	Примером статистической интерполяции может быть применение регрессионной модели, где нужно вычислить зависимую переменную, а значение независимых переменных выводятся из предположительно достоверных ответов (Pierzchala, 1995).
Статическое согласование (при динамической интерполяции)	Акт согласования донорской записи с записью получателя (кандидата) в соответствии с определенными статистическими критериями с тем, чтобы перевести данные от донора к получателю (Pierzchala, 1995).
Стохастическое редактирование	Редактирование, которое в случае нарушения указывает на ошибку в данных с вероятностью меньше единицы (Pierzchala, 1995).
Структурное редактирование	Проверки на основе логической взаимосвязи между двумя или более полями редактирования. Например, итог должен равняться сумме его составляющих: или из-за предусмотренной в переписном листе схемы пропуска, где переменные, лежащие на непересекающихся дорожках, не могут оба равняться не нулю. Структурное редактирование обеспечивает сохранение структуры переписного листа в записи данных (Pierzchala, 1995).
Структурная интерполяция	Структурная интерполяция используется, когда между несколькими переменными есть структурная взаимосвязь. Например, итог должен равняться сумме его составляющих: поэтому, если вопрос касается матери, количество детей, родившихся живыми, должно равняться количеству живых детей плюс детей умерших (Pierzchala, 1995).
Редактирование на достоверность	Контрольное редактирование, которое производится между полями в конкретной записи. Сюда входит проверка каждого поля каждой записи, чтобы удостовериться, что она содержит достоверные данные, и проверка записей в определенной заранее комбинации полей, чтобы удостовериться, что данные согласуются между собой (Pierzchala, 1995).
Весовые коэффициенты	По методу редактирования и интерполяции Фелледжи–Холта весовые коэффициенты присваиваются полям на основе надежности. Чем выше весовой коэффициент, тем больше вероятность интерполяции поля (при прочих равных вещах). Весовые коэффициенты могут также присваиваться редактируемому данным (Pierzchala, 1995).
Редактирование в рамках записи	Другое название редактирования на достоверность (Pierzchala, 1995).

БИБЛИОГРАФИЯ

- Banister, J. (1980). Use and abuse of census editing and imputation. *Asian and Pacific Census Forum*, vol. 6, No. 3, pp. 1–20.
- Bankier, M., M. Lachance and P. Poirer (1999). A generic implementation of the new imputation methodology. In *Proceedings of the Section on Survey Research Methods*. Alexandria, Virginia: American Statistical Association (forthcoming).
- Bankier, M., A.-M. Houle and M. Luc (n. d.). Canadian census demographic variables imputation. Manuscript.
- Bankier, M., M. Luc, C. Nadeau and P. Newcombe (1996). Imputing numeric and qualitative census variables simultaneously. In *Proceedings of the Section of Survey Research Methods*. Alexandria, Virginia: American Statistical Association, pp. 287–292.
- Boucher, L. (1991). Micro-editing for the annual salary of manufacturers: what is the value added? In *Proceedings of the Annual Research Conference*. Washington D.C.: United States Bureau of the Census, pp. 765–781.
- Fellegi, I. P., and D. Holt (1976). A systematic approach to automatic edit and imputation. *Journal of the American Statistical Association*, vol. 71, No. 353 (March), pp. 17–35.
- Ford, Bany L. (1983). An overview of hot deck procedures. In *Incomplete Data in Sample Surveys*, vol. 2, *Theory and Bibliographies*. William G. Madow, Ingram Olkin and Donald B. Rubin, eds.
- Granquist, L. (1984). Data editing and its impact on the further processing of statistical data. Paper presented at the Workshop on Statistical Computing, Budapest, 12–17 November 1984.
- Granquist, L. (1987). The short-term developing program for computer-supported editing at Statistics Sweden. Report presented at the Data Editing Joint Group Meeting, Madrid, 22–24 April 1987. Stockholm: Statistics Sweden.
- Granquist, L. (1997). The new view on editing. *International Statistical Review*, vol. 65, No. 3, New York: Academic Press, pp. 381–387.
- Granquist, L. and J. G. Kovar (1997). Editing of survey data: how much is enough? In *Survey Measurement and Process Quality*, L. Lyberg et al (eds), New York: John Wiley and Sons, pp. 415–435.
- Greenberg, Brian, and Rita Surdi (1984). A flexible and interactive edit and imputation system for ratio edits. In *Proceedings of the American Statistical Association, Section on Survey Research Methods*. Alexandria, Virginia: American Statistical Association, pp. 421–426.
- Kalton, G. and D. Kasprzyk (1982). Imputing for missing survey responses. In *Proceedings of the Section on Survey Research Methods*. American Statistical Association, pp. 23–31.
- Kalton, G. and D. Kasprzyk (1986). The Treatment of Missing Survey Data. *Survey Methodology*, vol. 12, pp. 1–16.
- Naus, J. I. (1975). *Data Quality Control and Editing*. New York: Marcel Dekker.
- Nordbotten, S. (1963). Automatic editing of individual statistical observations. Conference of European Statisticians, Statistical Standards and Studies, No. 2. New York: United Nations.

- Pierzchala, M. (1995). Editing systems and software. In *Business Survey Methods*, B.G. Cox, and others, eds. New York: John Wiley and Sons, pp. 411–425.
- Pullum, T. W., T. Harpham and N. Ozsever (1986). The machine editing of large-sample surveys: the experience of the World Fertility Survey. *International Statistical Review*, vol. 54, pp. 311–326.
- Rubin, D. B. (1987). *Multiple Imputation for Non-response in Surveys*. New York: Wiley.
- Sande, I. G. (1982). Imputation in surveys: coping with reality. *American Statistician*, vol. 36, pp. 145–152.
- Särndal, C. E., B. Swensson and J. Wretman (1992). *Model Assisted Survey Sampling*. New York: Springer-Verlag.
- Statistics Canada (1998). *Statistics Canada Quality Guidelines*, 3rd Edition. Ottawa: Statistics Canada.
- Организация Объединенных Наций (1992 год). *Справочник по проведению переписи населения и жилого фонда. Часть I: Планирование, организация и управление переписями населения и жилого фонда*. Методологические исследования, серия F, № 54, в продаже под № R.92.XVII.8.
- Организация Объединенных Наций (1992 год). *Справочник по проведению переписи населения и жилого фонда. Часть II: Демографические и социальные характеристики*. Методологические исследования, серия F, № 54, в продаже под № R.91.XVII.9.
- Организация Объединенных Наций (1998 год). *Принципы и рекомендации в отношении переписи населения и жилого фонда, первое пересмотренное издание*. Статистические документы, серия M, № 67/Rev. 1, в продаже под № R.98.XVII.8.
- United Nations (1999). *Standard Country of Area Codes for Statistical Use*, Statistical Papers, Series M, No. 49/Rev. 4. Sales No. M.98.XVII.9.
- United Nations Statistical Commission and Economic Commission for Europe (1994). *Statistical Data Editing*, vol. No. 1: *Methods and Techniques*. Conference of European Statisticians, Statistical Standards and Studies Series, No. 44. Sales No. E.94.II.E.36.
- United Nations Statistical Commission, and Economic Commission for Europe (1997). *Statistical Data Editing*, vol. 2: *Methods and Techniques* Conference of European Statisticians, Statistical Standards and Studies Series, No. 48. Sales No. E.96.II.E.30.

كيفية الحصول على منشورات الأمم المتحدة

يمكن الحصول على منشورات الأمم المتحدة من المكتبات ودور التوزيع في جميع أنحاء العالم . استعلم عنها من المكتبة التي تتعامل معها أو اكتب إلى : الأمم المتحدة ، قسم البيع في نيويورك أو في جنيف .

如何 购取 联合国 出版物

联合国出版物在全世界各地的书店和经售处均有发售。 请向书店询问或写信到纽约或日内瓦的联合国销售组。

HOW TO OBTAIN UNITED NATIONS PUBLICATIONS

United Nations publications may be obtained from bookstores and distributors throughout the world. Consult your bookstore or write to: United Nations, Sales Section, New York or Geneva.

COMMENT SE PROCURER LES PUBLICATIONS DES NATIONS UNIES

Les publications des Nations Unies sont en vente dans les librairies et les agences dépositaires du monde entier. Informez-vous auprès de votre libraire ou adressez-vous à : Nations Unies, Section des ventes, New York ou Genève.

КАК ПОЛУЧИТЬ ИЗДАНИЯ ОРГАНИЗАЦИИ ОБЪЕДИНЕННЫХ НАЦИЙ

Издавания Организации Объединенных Наций можно купить в книжных магазинах и агентствах во всех районах мира. Наводите справки об изданиях в вашем книжном магазине или пишите по адресу: Организация Объединенных Наций, Секция по продаже изданий, Нью-Йорк или Женева.

COMO CONSEGUIR PUBLICACIONES DE LAS NACIONES UNIDAS

Las publicaciones de las Naciones Unidas están en venta en librerías y casas distribuidoras en todas partes del mundo. Consulte a su librero o diríjase a: Naciones Unidas, Sección de Ventas, Nueva York o Ginebra.
