

Руководство по интеграции данных для измерения миграции



ЕВРОПЕЙСКАЯ ЭКОНОМИЧЕСКАЯ КОМИССИЯ ООН

**Руководство по
интеграции данных для
измерения миграции**



UNITED NATIONS

New York and Geneva, 2019

Примечание

Использованные в настоящей публикации обозначения и изложение материала не означают высказывания какого-либо мнения со стороны Секретариата Организации Объединенных Наций относительно правового статуса какой-либо страны, территории, города или района, или их властей, или относительно делимитации их границ или рубежей.

ECE/CES/STAT/2018/6
ПУБЛИКАЦИЯ ОРГАНИЗАЦИИ ОБЪЕДИНЕННЫХ НАЦИЙ
Sales No.: E.19.II.E.7 ISBN: 978-92-1-117185-3 e-ISBN: 978-92-1-047575-4

Copyright © United Nations,
2019 All rights reserved
worldwide

United Nations publication issued by the Economic Commission for Europe

Предисловие

Миграция формирует общества. Ее экономические, социальные и демографические последствия являются значительными и продолжают нарастать. Политические деятели, исследователи и другие заинтересованные стороны нуждаются в данных о мигрантах – о их количестве, показателях въезда и выезда, характеристиках и интеграции в общество. Эти данные должны быть исчерпывающими, точными и часто обновляемыми. Не существует единого источника, который мог бы предоставить такие данные, но, объединив несколько источников, можно получить информацию, необходимую пользователям. Некоторые страны разработали методы объединения административных, статистических и других источников данных для составления статистики миграции.

Данная публикация, разработанная рабочей группой экспертов из национальных статистических управлений и одобренная Конференцией европейских статистиков в июне 2018 года, содержит обзор способов использования интеграции данных для составления статистики миграции на основе опроса поставщиков данных о миграции в более чем 50 странах. Тринадцать тематических исследований содержат более подробную информацию об интеграции данных в различных национальных контекстах. В публикации предлагаются принципы передовой практики по интеграции данных для измерения миграции.

Настоящая публикация предназначена для использования национальными статистическими управлениями и другими производителями статистических данных по миграции, а также для предоставления пользователям информации о процессе подготовки данных по миграции, которые они используют.

Благодарность

Настоящее руководство было подготовлено Целевой группой ЕЭК ООН по интеграции данных для измерения миграции при участии следующих членов: Антонио Аргесос (Испания, председатель Целевой группы), Регина Фукс (Австрия), Энрико Туччи (Италия), Моника Перес (Италия), Кирстен Ниссен (Новая Зеландия), Марсель Хайнигер (Швейцария), Джей Линдоп (Соединенное Королевство), Пол Викарс (Соединенное Королевство), Джейсон Шектер (Соединенные Штаты Америки), Стефани Эверт (Соединенные Штаты Америки), Энтони Кнапп (Соединенные Штаты Америки), Джампаоло Ланцери (Евростат), Натан Ментон (ЕЭК ООН), Паоло Валенте (ЕЭК ООН). Следующие национальные эксперты представили тематические исследования по своим странам: Мелани Менье (Канада), Адам Дикманн (Венгрия), Ахмад Хлейхель (Израиль), Йерун Ойеваар (Нидерланды), Байба Зукула (Латвия). Мы благодарим всех экспертов за их вклад.

Содержание

Предисловие	iii
1. Введение	9
2. Определение термина «интеграция»	11
2.1 Общие характеристики	11
2.2 Рабочее определение	13
3. Обследование методов интеграции данных	15
3.1 Обзор основных результатов обследования	15
3.2 Рекомендации	16
3.3 Резюме результатов	16
4. Тематические исследования	17
4.1 Корректировка с учетом неучтенных смертей в реестре населения	18
4.1.1 Австрия	18
4.1.2 Венгрия	19
4.1.3 Израиль	22
4.1.4 Италия	25
4.1.5 Латвия	30
4.1.6 Нидерланды	32
4.1.7 Испания	34
4.1.8 Швейцария	40
4.2 Тематические исследования без реестров населения	43
4.2.1 Австралия	43
4.2.2 Канада	29
4.2.3 Новая Зеландия	48
4.2.4 Великобритания	54
4.2.5 Соединенные Штаты Америки	58
4.3 Резюме	59
5. Метаданные	63

6.	Заключение	65
7.	Рекомендации	67
7.1	Улучшить доступ национальных статистических учреждений к административным данным ..	67
7.2	Использовать административные данные для статистики миграции	67
7.3	Комбинировать данные из различных источников с использованием «сигналов присутствия»	67
7.4	Следить за качеством интегрированных данных	68
7.5	Обеспечить прозрачность применяемых методов интеграции данных и разрабатывать стандарты	68
7.6	Развивать международное сопоставление данных о миграции и обмен ими	68
8.	Дальнейшая работа	69
8.1	Потенциал больших данных	69
8.2	Продольные данные	69
9.	Ссылки	70

Список рисунков

Рисунок 4.1	Источники данных для ежегодного расчета численности населения Венгрии	21
Рисунок 4.2	Схема ежемесячного присутствия моделей непрерывности в трудовой и учебной деятельности..	27
Рисунок 4.3	Иностранное население в Испании по данным реестра населения 1998-2010 гг. (в миллионах).....	36
Рисунок 4.4	Число обычных резидентов Новой Зеландии, родившихся за рубежом	51
Рисунок 4.5	Процент населения Новой Зеландии, родившегося за рубежом	51
Рисунок 4.6	Расчет долгосрочной международной миграции	56

Список таблиц

Таблица 4.1	Схема процесса и численность групп населения (в тысячах) в зависимости от вероятности их обычно проживающий в Италии	28
Таблица 4.2	Сомнительные записи в файле предварительной переписи 2011 года. Выбранные национальности	38
Таблица 4.3	Сомнительная группа населения в файле предварительной переписи по национальности, провинции и возрасту	39
Таблица 4.4	Сбор и обработка данных о международной миграции	50

1. Введение

1. В последние годы тема интеграции данных вызывает большой интерес, особенно в условиях роста потребностей в данных при жестком налоговом климате в сочетании с возросшей нагрузкой на респондентов и требованиями в отношении конфиденциальности. В 2016–2017 годах Группа высокого уровня ЕЭК ООН по модернизации официальной статистики (ГВУ-МОС) занималась реализацией проекта в области интеграции данных. Этот проект предусматривал проведение экспериментов по интеграции данных в различных секторах и завершился разработкой практического электронного руководства по интеграции данных для официальной статистики¹. Евростат поддержал исследования, посвященные методам совершенствования интеграции данных, такие как проект ESSnet в области интеграции результатов обследований и административных данных.² Этот проект стал первой попыткой создания общей методологической базы для интеграции различных источников данных.

2. Однако многие страны заинтересованы в интеграции данных непосредственно для целей измерения международной миграции. Учитывая проблемы в области сбора статистики миграции, зачастую полезно собирать данные из различных источников. Интеграция данных позволяет преодолевать затруднения, связанные с их охватом или точностью, которые могут касаться общих данных о численности и потоках мигрантов, и повышать разнообразие данных о миграции, добавляя к имеющимся данным социально-демографические или экономические параметры.

3. На европейском уровне значение интеграции данных для совершенствования статистики миграции было вновь подтверждено в ходе ежегодной конференции генеральных директоров НСУ 2017 года, программа которой предусматривала статистическое совещание на тему «Вопросы перемещения и интеграции населения: статистика миграции»³. Будапештский меморандум, который обсуждался на конференции и был одобрен Комитетом Европейской статистической системы, включает в себя список мероприятий, среди которых фигурирует «содействие выявлению, оценке и внедрению новых методов и источников данных, в частности расширение использования для статистических целей источников административных данных соответствующего качества, обеспечиваемого благодаря постоянной оценке качества, – либо отдельных регистров, увязанных данных их различных административных источников, либо в сочетании с источниками данных, полученных в результате обследований, а также возможностям новых источников данных (например, больших данных)».

4. Методы интеграции различных источников административных данных внутри той или иной страны (например, регистры населения и регистрационные записи в сферах здравоохранения, трудовой деятельности, налогообложения или образования) для восполнения информации, которую не удается получить из других источников (например, по измерению эмиграции для тех, кто не удаляет регистрационные данные), показали в ряде стран успешные результаты. В других случаях источники неадминистративных данных позволяют получать информацию, недополученную из административных источников (например, применение обследований домохозяйств для сбора информации, не включенной в регистры). Большие данные, поступающие от коммунальных и телефонных компаний, в некоторых странах также рассматриваются в качестве дополнительной информации. Другие виды применения интеграции данных могут заключаться в согласовании различных показателей миграции, поступающих из разных источников, в частности оценочных данных по миграционным группам, которые с трудом поддаются учету, таким как нелегальные мигранты или эмигранты.

5. Учитывая важность и злободневность этой темы, в 2014 году на рабочей сессии ЕЭК ООН-Евростата по статистике миграции состоялось обсуждение вопроса об использовании различных источников административных данных для измерения международной миграции. Обсуждение было сосредоточено на возможностях и затруднениях при использовании административных данных,

¹ <https://statswiki.unecce.org/display/DI/Data+Integration+Home>

² https://ec.europa.eu/eurostat/cros/content/data-integration_en

³ 103-я Конференция ГД НСУ (Будапешт, 20–21 сентября 2017 года). Документы и доклады см. по адресу <http://www.ksh.hu/dgins2017/presentations.html>.

особенно с учетом разных уровней развития систем административных данных на территории региона. Были рассмотрены пути улучшения сотрудничества между национальными миграционными службами, статистическими учреждениями, реестровыми службами и другими органами, готовящими административные данные, а также методы интеграции различных источников административных данных для совершенствования измерений миграции. Кроме того, участники отметили потребность в практических рекомендациях и руководстве по вопросам подготовки метаданных для упрощения сопоставлений между оценочными данными по миграции, составленными в разных странах.

6. В заключение участники рабочей сессии рекомендовали продолжать методологическую работу по теме интеграции множественных источников данных для измерения международной миграции, включая источники данных внутри одной страны и между разными странами, и передовой практики в сфере коммуникации между национальными статистическими учреждениями и производителями административных данных.

7. Опираясь на результаты этих дискуссий, в октябре 2015 года Бюро Конференции европейских статистиков (КЕС)⁴ учредило Целевую группу по интеграции данных для измерения миграции с целью подготовки свода руководящих принципов и описания передовой практики интеграции различных источников данных для совершенствования измерений иммиграции, эмиграции и чистой миграции. В настоящей публикации представлены результаты работы Целевой группы.

8. В настоящей публикации дается общий обзор этой темы наряду с обзором видов интеграции данных, которые уже используются в различных странах, будь то разрозненные регулярно публикуемые данные или элементы пилотных исследований. На основе этого обзора излагаются также принципы передовой практики, призванные служить рекомендациями по совершенствованию интеграции данных для измерения миграции в различных странах. Опыт отдельных стран систематизирован в этой публикации на основе обследования поставщиков данных по миграции почти в 50 странах, а также на основе более подробных тематических исследований по ряду стран.

9. Вопрос об интеграции данных рассматривается в настоящей публикации путем анализа как **интеграции макроданных**, т. е. сопоставления/статистического моделирования на основе данных, представляющих собой конгломераты (статистику) индивидуальных регистрационных систем, так и **интеграции микроданных**, т. е. интеграции данных на основе увязывания/сопоставления отдельных записей. Наряду с этим анализируются различные уровни дублирования переменных и/или индивидуальных данных из разных источников.

10. Последующие детали, связанные в вышеуказанными понятиями, изложены в виде элемента концептуальной основы и определения термина «интеграция данных» в главе II. Глава III содержит анализ результатов обследования поставщиков данных по миграции. Практические примеры интеграции данных по миграции приведены в рамках углубленных тематических исследований, выполненных в ряде стран. Обобщенный результат этих тематических исследований содержится в главе IV (тематические исследования включены в полную версию настоящего документа, которая издана только на английском языке). Предложение по метаданным, которое предстоит рассмотреть в связи со статистикой международной миграции, представлено в главе V. Наконец, выводы и общие рекомендации, сформулированные на основе итогов обследования и тематических исследований, изложены в главе VI.

⁴ В состав Конференции европейских статистиков входят национальные статистические организации региона ЕЭК ООН (по странам – членам ЕЭК ООН см.

www.unecsc.org/oes/nutshell/member_states_representatives.html), а также Австралия, Бразилия, Китай, Колумбия, Мексика, Монголия, Новая Зеландия, Республика Корея, Чили и Япония.

В ее работе также участвуют крупные международные организации, занимающиеся вопросами статистики в регионе ЕЭК ООН, такие как Статистическая служба Европейской комиссии (Евростат), Организация экономического сотрудничества и развития (ОЭСР), Межгосударственный статистический комитет Содружества Независимых Государств (СНГ-СТАТ), Международная организация труда (МОТ), Международный валютный фонд (МВФ), Всемирная торговая организация (ВТО) и Всемирный банк.

2. Определение термина «интеграция данных»

2.1 Общие характеристики

11. Уже давно известно, что измерение миграции является весьма сложной задачей. Хотя в последнее время как в ученых кругах, так и в сфере официальной статистики предпринимались значительные усилия для повышения качества статистики миграции, точность этих данных по-прежнему имеет определенный предел погрешности. Позднее в более широком статистическом контексте внимание обращалось на перспективы, связанные с «интеграцией данных» как с возможным дополнительным методом совершенствования статистики (например, Евростат в 2009 и 2013 годах; ESSnet в 2008, 2011 и 2014 годах; ГВУ-МОС ЕЭК ООН в 2017 году).

12. Интеграция данных зачастую упоминается как один из наиболее эффективных подходов, способствующих обогащению и совершенствованию статистической информации. В последние годы с интеграцией данных был связан целый ряд видов деятельности и проектов, в рамках которых использовались различные определения этого термина. Например, согласно Типовой модели производства статистической информации (ТМПСИ), интеграция данных – это один из методов производства статистической информации, который заключается в объединении данных из одного или нескольких источников. В вышеупомянутом проекте ГВУ-МОС интеграция данных определяется как «метод, при котором данные, поступившие как минимум из двух разных источников, комбинируются в одном наборе данных»⁵. Однако, насколько нам известно, в статистическом сообществе не существует определения интеграции данных, согласованного на международном уровне. Есть целый ряд аспектов, помогающих охарактеризовать интеграцию данных: количество и типология источников данных, методология, своевременность, нагрузка на респондентов и, не в последнюю очередь, цель ее применения.

13. Интуиция подсказывает, что для интеграции требуется наличие как минимум двух исходных материалов. Такими материалами являются наборы данных, т. е. систематизированная информация, взятая из определенных источников данных в рамках статистической операции. Применительно к миграции наборы данных можно распределить по четырем категориям: (данные, источником которых являются) статистические обследования (в том числе сплошные обследования, т. е. переписи), административные реестры, большие данные и географическая информация. Указанные типологии способны сочетаться во всевозможных комбинациях, и эту процедуру можно повторять, формируя наборы данных, которые затем будут интегрированы в другие наборы данных.

14. Метод интеграции наборов данных зависит в первую очередь от того, какой является содержащаяся в них информация – микро- или макроуровневой, а в случае отдельных записей – от наличия единого идентификатора. Для интеграции на микроуровне от этого последнего аспекта зависит выбор между статистическим согласованием и увязкой записей; метод увязки записей обычно применяется при наличии идентификатора во всех интегрируемых наборах данных. При интеграции на макроуровне, т. е. между данными, собранными из отдельных записей, соответствующая процедура проводится в два этапа: во-первых, устранение различий в понятиях и функциональных определениях, а во-вторых – согласование («уравновешивание») данных с использованием различных статистических методов или только на основе экспертного мнения.

15. Применительно к данным о миграции индивидуальным идентификатором является персональный идентификационный номер (ПИН). На практике метод микроуровневой интеграции предпочтителен для составления регистров населения – процесса, производной которого зачастую является статистика миграции. Следует отметить, что ПИН – это не единственный возможный единый идентификатор, поскольку при увязке данных, полученных в результате субнационального агрегирования, от административных образований, из файловых записей и т. д., могут использоваться разные идентифицирующие коды.

16. Источники данных, которые используются в качестве исходных материалов для составления регистра населения, могут быть многочисленны, и в некоторых странах со временем были

⁵ <https://statswiki.unece.org/pages/viewpage.action?pagelId=169018059>

интегрированы дополнительные источники данных, обогащающие статистическую информацию о населении и косвенно – о миграции. Аналогичным образом, хотя обычно в меньшей степени, может обстоять дело с регистрами иностранного населения/registрами иностранцев. Таким образом, интеграцию данных можно рассматривать как расширяющийся процесс, в ходе которого к первым интегрированным наборам данных добавляются другие данные, полученные благодаря приобретенному опыту, совершенствованию методов интеграции и доступу к новым источникам данных.

17. Еще одной отличительной чертой измерения миграции является то, что, учитывая характер этого явления, наблюдающегося между двумя странами (страной приема и страной происхождения, либо страной назначения и страной рождения, либо других переменных для идентификации мигрантов), могут использоваться другие источники данных, кроме национальных. Использование этого характерного международного фактора ограничено в основном обменом агрегированными данными, учитывая серьезную озабоченность соблюдением требований конфиденциальности и национальной безопасности, которыми страны могут объяснять свой отказ на любой запрос об обмене микроданными по международной миграции. Примеры применения этого последнего подхода имеются в северных странах, где международный обмен микроданными используется для повышения качества регистров населения.

18. Добавление новых источников данных может приводить к изменению своевременности статистической информации, т. е. промежутка времени между контрольным периодом и получением результатов. Хотя по общему правилу своевременность того или иного конкретного продукта из комплексных данных определяется источником данных, имеющим более широкий временной охват, статистическое моделирование может в конечном счете преодолеть эту ситуацию, открывая возможность для представления более своевременных оценочных показателей. По-видимому, снижение уровня своевременности будет приемлемым лишь в тех случаях, когда интеграция данных обеспечивает соответствующее повышение охвата и точности статистических измерений.

19. В качестве одной из возможностей, создаваемых благодаря интеграции данных, иногда упоминается сокращение нагрузки на респондентов. На сегодняшний день большинство видов использования альтернативных источников данных для нужд статистики миграции обеспечивается путем замены одного источника данных другим, с меньшей нагрузкой на респондентов/поставщиков данных, или путем совершенствования статистической операции, применяемой к соответствующему источнику данных. С концептуальной точки зрения этот метод можно рассматривать скорее как повышение эффективности, нежели как расширение спектра имеющихся данных, вследствие чего он не вполне характерен/актуален для интеграции данных.

20. Исходные материалы, взятые из альтернативных источников данных, можно использовать для подтверждения статистической информации из отдельного, официального источника данных. В этом случае применяемый метод можно считать **сопоставлением**, а не **интеграцией**. Он также может рассматриваться в качестве первого шага в процессе постепенной интеграции используемых источников данных, особенно когда дополнительные источники данных не отвечают требованиям официальной статистики.

21. Еще один случай – это охват каждым источником данных определенной подгруппы населения (мигрантов). Объединение таких данных для получения общей статистической информации можно охарактеризовать как компиляцию, а не как чистую интеграцию, поскольку фактического объединения на уровне отдельной записи не происходит (либо индивидуальные, либо агрегированные образования).

22. Из вышеизложенных соображений следует, что квалифицирующим признаком интеграции данных, бесспорно, является использование множественных наборов данных, но в зависимости от метода их совокупного использования. На практике существуют статистические операции, которые в некоторой степени представляют собой «пограничные» случаи с интеграцией данных, такие как компиляция данных и сопоставление данных. Что касается других признаков, то степень детализации данных имеет значение скорее для выбора методологии, которая будет применена, нежели для идентификации вида деятельности по интеграции данных. Своевременность также нельзя использовать в качестве критерия идентификации, поскольку интеграция данных в принципе способна как сократить, так и удлинить задержку в получении результатов, причем позитивное влияние на сроки можно оказывать и с помощью применения других статистических операций, кроме интеграции. Аналогичным образом, не всякая типовая ссылка на повышение качества статистической информации (включая снижение нагрузки на респондентов) будет полезной, поскольку в принципе каждая статистическая операция должна быть нацелена на достижение такого результата. Очевидно, это не исключает, что метод интеграции данных может обеспечивать подобный позитивный эффект, а лишь означает, что такие ссылки нельзя использовать для выявления отличительных признаков интеграции данных.

23. Результатом мер по интеграции данных должен являться «обогащенный» или «высококачественный» набор данных. Расчет, производимый в рамках типовой матрицы данных $n \times p$, где n – это записи, а p – переменные, может обеспечить такое обогащение в обоих направлениях, расширяя охват (т. е. увеличивая n) и/или объем информации в тех же записях (т. е. добавляя новые переменные в набор p). Что касается повышения качества, то оно найдет отражение не в изменениях объема получаемого набора данных, а скорее в его содержании (например, в весах выборочного обследования).

2.2 Рабочее определение

24. Имеющееся определение термина «интеграция данных» заимствовано из ИТ-среды (ОСДМ, 2009 год) и сформулировано следующим образом: «процесс комбинирования данных из двух или более источников для подготовки статистической информации». В этой же ссылке уточняется, что «интеграция данных может осуществляться на микроуровне, где она часто квалифицируется как сопоставление, или на макроуровне.

25. В основу этого определения заложены исходные материалы («двух или более источников») и цель («для подготовки статистической информации»). Последний элемент носит довольно типовой характер, поскольку получение статистической информации можно считать итогом любого статистического процесса, не обязательно требующего интеграции данных.

26. Для целей настоящего документа предлагается следующее рабочее определение: «Интеграция данных – это вид статистической деятельности с использованием двух или более наборов данных, результатом которого является получение единого расширенного и/или более качественного набора данных».

27. Интеграция данных может производиться на двух основных уровнях агрегации – микро- и макроинтеграции⁶, которые определяются следующим образом:

- «Интеграция микроданных» – это интеграция данных на основе увязки/статистического сопоставления отдельных записей с использованием ключевых идентифицируемых переменных.
- «Интеграция макроданных» – это комбинирование данных на основе совокупностей (статистики) отдельных записей.

28. Ссылка на типовое понятие «статистическая деятельность» позволяет применять на макро- или микроуровне любую методологию, включая экспертное мнение. Это понятие также охватывает анализ концептуальных различий, являющийся одним из важнейших шагов любого статистического процесса. Прямое упоминание методов статистического согласования и/или увязки записей, хотя они составляют специфику интеграции данных, привело бы к исключению макроуровневой интеграции.

29. Выражение «двух или более источников» заостряет внимание на многоаспектном характере интеграции данных. Оно также указывает на то, что, хотя интеграция завершается достижением единого результата, она представляет собой процедуру, которую необходимо повторять каждый раз, когда искомая цель заключается в достижении такого результата. Иными словами, при регулярной подготовке статистических данных на основе более чем одного набора данных интеграция не является окказиональной, однократной статистической процедурой. «Интегрированный» результат генерируется всякий раз, когда появляются более свежие вводные данные⁷.

30. Употребление термина «наборы данных» вместо «источники данных» указывает на то, что предметом интеграции являются фактические данные, а не те источники данных, из которых они получены. Например, объединение двух имеющихся выборочных обследований в одно

⁶ В некоторых контекстах термин «микро-макро» употребляется при комбинировании микроуровневых и макроуровневых данных. Здесь этот подход может рассматриваться как один из случаев интеграции макроданных

⁷ Возможно, выявление концептуальных различий в контексте интеграции данных является исключением из «правила повторений», поскольку такая процедура потребуется лишь один раз, если концептуальная основа вводимых данных со временем не изменится

всеобъемлющее обследование путем изменения вопросников, плана выборки и т. п. не может рассматриваться в данном контексте как пример **интеграции данных**. Это также подтверждает вышеизложенную точку зрения, согласно которой интеграция не преобразует источники вводных данных в новые, смешанные источники данных; основное внимание в ходе этой процедуры уделяется **результатам**, полученным на основе этих источников данных. Таким образом, процедура интеграции носит, пожалуй, столь же систематический характер, что и подготовка статистической информации на основе данных, к которым она применяется.

31. Результатом интеграции данных является прежде всего «единый» набор данных, т. е. новый набор данных, в котором вся информация из наборов вводных данных реорганизована в единообразном виде. Это не просто объединение нескольких наборов данных, поскольку предполагается, что из них надлежащим образом устранены дублирования, концептуальные различия и любые другие проявления необъективности. Это – «расширенный» набор данных, т. е. структурированный новый набор данных, измерения которого не могут быть меньше самых крупных соответствующих измерений отдельных наборов вводных данных, и/или «высококачественный» набор данных, элементы которого были изменены в результате интеграции данных, обеспечившей (по возможности измеряемое) повышение качества статистической информации.

3. Обследование методов интеграции данных

32. Чтобы собрать информацию, необходимую для выполнения этой работы, Целевая группа провела обследование методов интеграции данных. Был подготовлен специальный вопросник с целью сбора информации о нынешних методах использования комбинированных или комплексных источников данных для измерения иммиграционных и эмиграционных потоков и для подготовки статистики по группам мигрантов и некоренного населения.

33. В сентябре 2016 года этот вопросник был разослан национальным статистическим учреждениям стран – членов ЕЭК ООН⁸. Ответы на те части обследования, которые касались применяемых ими методов сбора данных, представили 56 стран.

34. Анализ итогов обследования был проведен за период с конца 2016 года по первый квартал 2017 года; его основные результаты обобщены в пунктах ниже.

3.1 Обзор основных результатов обследования

35. Для обеспечения подготовки статистической информации о международных миграционных потоках ключевую роль играет та или иная форма интеграции данных. Там, где это возможно, интеграция данных осуществляется на уровне отдельных записей, однако такой подход приемлем в основном для стран, имеющих полные регистры населения, в которых учитываются выезды своих граждан и въезды иностранцев.

36. Во многих странах фактическое миграционное событие не регистрируется в каком-либо одном источнике данных. Статистику миграционных потоков можно получить только путем объединения составных элементов, в которых учитываются изменения, происходящие в постоянном составе населения вследствие миграции.

37. Статистическая информация о международных миграционных потоках на основе одного источника производится при наличии возможности постоянно вести административный учет миграционных событий (фактических пересечений границы) или при наличии системы учета, основанной на полном регистре населения. Однако в ходе контроля качества и дальнейшей разбивки этой информации может потребоваться интеграция со вспомогательными источниками данных. Эта процедура может происходить в форме корректировки миграционных потоков после сопоставления с имеющейся статистикой, полученной из других стран (зеркальная статистика). Объединение лиц в рамках одного источника может потребоваться для продольного исследования трансграничных перемещений как метода определения статуса мигрантов.

38. В результате обследования было выявлено широкое разнообразие способов применения интеграции данных. В некоторых странах регистры учета граждан могут не обеспечивать увязку на уровне регистрационных единиц с какой-либо отдельной административной процедурой учета данных об иностранных гражданах. В таких случаях регистр населения может использоваться в качестве источника для перекрестной проверки и условного расчета отсутствующих данных.

39. Наиболее распространенными источниками для разработки статистических данных о некоренном населении будут являться перепись населения и регистр населения. Перепись населения может обеспечить информацию о населении в целом, а последующая интеграция с административными данными позволяет обновлять статистические данные о некоренном населении.

40. В случаях, когда миграционная статистика составляется на основе данных из нескольких источников, некоторые страны указали на необходимость оценки приоритетности источников с точки

⁸ Кроме 56 стран – членов ЕЭК ООН, вопросник был также направлен другим странам, регулярно участвующим в работе Конференции европейских статистиков, в том числе Австралии, Колумбии, Мексике, Новой Зеландии, Чили и Японии.

зрения их своевременности, надежности, последовательности и других критериев качества.

3.2 Рекомендации

41. Поскольку страны занимаются развитием и стандартизацией регистров наборов статистических данных по социальной сфере и народонаселению, интеграция данных чаще осуществляется на уровне отдельных записей по различным секторам. Одним из неперенных элементов разработки статистики международной миграции наверняка станет уделение повышенного внимания оценке качества интегрированных данных.

42. Все более важным для статистических учреждений становится обеспечение прозрачного представления правил и методов, которые используются в процессе интеграции данных для разработки статистики международных миграционных потоков и других видов статистики международной миграции. Это открывает возможности для разработки в будущем методологических стандартов по способам интеграции данных на уровне национальных статистических учреждений. Отдельные страны всегда будут иметь собственные наборы данных и методы их интеграции, но наряду с этим, скорее всего, возрастет число открытых технических форумов для обсуждения правил и методов, используемых при комбинировании источников для выработки миграционной статистики.

3.3 Резюме результатов

43. Почти все страны-респонденты занимаются выработкой статистики международных миграционных потоков, причем более половины из этих стран используют несколько источников данных. Одни страны опираются на центральный регистр учета всех сведений о регистрациях и дерегистрациях жителей, а последующая интеграция источника на уровне отдельных записей с административными наборами данных по эмиграционным/иммиграционным потокам и случаям смерти способствует уточнению информации о миграционных потоках. Другие страны прибегают к интеграции регистров населения с данными, полученными в результате обследований или других видов административного сбора информации о негражданах, проводимых на макроуровне.

44. В целом около трети стран-респондентов (всего 35 ответов) интегрировали источники данных на макроуровне с применением статистического моделирования или другого метода комбинирования агрегированных данных, а несколько меньшая доля стран использует методы интеграции на уровне регистрационных единиц. При этом следует отметить, что несколько стран пользовались другими источниками только для перекрестной проверки и дополнительной информации. Около половины стран-респондентов отметили, что между единицами наблюдения, взятыми из интегрированных источников данных, наблюдается частичное дублирование; у меньшей доли стран при комбинировании источников оказывались идентичные единицы наблюдения; а несколько стран сообщили об использовании взаимоисключающих единиц наблюдения.

45. Менее половины стран готовят статистическую информацию о международных миграционных потоках, используя только один источник. Подготовка статистики по миграционным потокам на основе единственного источника зачастую выполняется в случаях, когда есть возможность постоянно вести сбор административных данных на всех пунктах пересечения границы или когда миграционные потоки могут вычлняться из регистра населения. Еще один, реже встречающийся вариант – это использование данных переписи населения для косвенной оценки миграционных потоков в показателях численности населения.

46. Порядка 90% стран-респондентов представили статистику или собранные данные по некоренному населению. Лишь немногим более половины этих стран используют более одного источника данных и обычно осуществляют интеграцию данных на уровне регистрационных единиц или на макроуровне. К числу основных видов и признаков используемых источников, будь то одиночные или интегрированные, относились проводимая каждые пять лет перепись населения, регистр населения, обследование или административный сбор данных. Сообщалось, что основной целью интеграции является подготовка статистики некоренного населения.

47. Около 60% стран-респондентов сообщили, что они не вырабатывают другой статистики по мигрантам или некоренному населению. Поэтому несколько стран указали, что интеграция источников данных дает возможность готовить такую статистическую информацию, как данные о поданных и удовлетворенных ходатайствах об убежище, разрешениях на проживание, разрешениях на работу, нелегальных иностранных рабочих, социально-экономических характеристиках мигрантов и причинах их расселения.

4. Тематические исследования

48. С учетом информации, собранной в ходе обследования методов интеграции данных, Целевая группа выявила ряд стран, накопивших значительный опыт в сфере их интеграции. Этим странам было предложено представить краткие записки с указанием первичных источников данных о миграции, причин считать процесс интеграции данных необходимым, применяемых методов и выгод от интеграции.

49. Тексты, представленные странами, приведены в данной главе как собрание значимых практических примеров интеграции данных для статистики миграции. Представленные тематические исследования по различным странам существенно различаются, поскольку национальные контексты отличаются по ряду важных аспектов, таких как: наличие регистров населения или других реестров; существование персонального идентификационного номера; возможность связывания данных между регистрами; правовые рамки; а также общественная чувствительность в отношении использования данных регистров. Поэтому классифицировать тематические исследования или сгруппировать их по категориям весьма затруднительно. Однако очень важным аспектом, касающимся интеграции данных — который в значительной степени влияет на применяемые методы, — является то, используется ли в стране регистр населения. Соответственно, сначала представлены тематические исследования стран, системы которых в основном основаны на регистрах населения или других административных реестрах, а затем — исследования стран, где системы не используют регистры населения и основаны преимущественно на карточках пассажиров или иных административных данных, не являющихся регистрами.

50. Большинство представленных тематических исследований следует структуре, предложенной Рабочей группой, которая включает следующие разделы:

- a) исходный/первичный источник данных для статистики миграции;
- b) ограничения исходного или основного источника. Почему необходима интеграция?;
- c) методы, используемые для интеграции данных;
- d) преимущества интеграции со статистикой из других источников. Некоторые тематические исследования имеют иную структуру, в зависимости от доступной информации или специфического контекста соответствующей страны.

4.1 Тематические исследования с реестрами населения

4.1.1 АВСТРИЯ

51. С 2002 года статистика миграции в Австрии основывается на данных Центрального реестра проживания (Zentrales Melderegister, ZMR). Жители, которые обустраивают свое жилье в частном или институциональном домохозяйстве, имеют юридическое обязательство зарегистрировать и сняться с регистрации по месту жительства в течение трех дней после переезда. Все жители, независимо от их гражданства⁹ или продолжительности пребывания, должны зарегистрироваться, если их пребывание превышает три дня¹⁰. Статистическое управление Австрии ежеквартально получает и обрабатывает все регистрации и снятия с регистрации по месту жительства.

52. Данные из Центрального реестра проживания дополняются путем интеграции данных в двух отдельных процессах. Во-первых, информация о смертности, полученная от Организации социального страхования Австрии (Hauptverband der Sozialversicherungsträger, HV), связывается с реестром проживания на уровне отдельных записей, добавляя информацию об умерших лицах и их отсутствующих снятиях с учета. Во-вторых, с целью корректировки данных с учетом отсутствующих сведений о снятии с регистрации эмигрантов, в ежегодных оценках определяются потенциальные номинальные члены на основе информации из нескольких административных источников данных. Хотя оба этапа связывают данные на микроуровне на заключительном этапе, лежащие в их основе процессы существенно различаются.

4.1.1.1 Корректировка с учетом неучтенных смертей в реестре населения

53. Данные Организации социального обеспечения Австрии служат дополнительным источником информации для улучшения качества статистики населения и миграции Австрии. Организация социального обеспечения Австрии является головной организацией всех фондов социального обеспечения в Австрии. Он собирает информацию обо всех застрахованных в Австрии лицах и их иждивенцах. Охват их базы данных определяется статусом страхования, а не статусом регистрации по месту жительства¹¹. Таким образом, реестр жителей и база данных социального обеспечения частично пересекаются, но также охватывают категории лиц, которые представлены только в одной из баз данных.

54. Регистрация смертей в Центральном реестре гражданского состояния (ZPR) является обязательной для всех смертей, произошедших на территории страны, а также для смертей граждан Австрии, произошедших за рубежом. После регистрации смерти умерший исключается из реестра по месту жительства. Однако смерти австрийских резидентов иностранного гражданства, произошедшие на территории иностранного государства, не подлежат регистрации. Уведомление о смерти может поступить в австрийские органы с задержкой.

55. Таким образом, база данных социального обеспечения предоставляет дополнительную информацию о смертности и используется для уточнения данных об общей численности населения и миграции. Эти данные сопоставляются с миграционными потоками из реестра проживания с помощью bPK, анонимизированного 27-значного ключа, который позволяет соединять информацию из разных административных источников данных.

4.1.1.2 Оценка потенциального числа номинальных членов в реестре населения

56. На дату проведения переписи населения на основе регистрационных данных (последняя перепись была проведена 31 октября 2011 года) лица, фигурирующие только в Центральном реестре проживания, но не в других административных реестрах, идентифицируются как подозрительные случаи, требующие дальнейшего расследования. С ними связываются, чтобы подтвердить их присутствие в Австрии. Те, кто не ответил, считаются номинальными членами и поэтому исключаются

⁹ Включая просителей убежища и беженцев

¹⁰ Иностранцы дипломатического персонала освобождаются от регистрации

¹¹ База данных социального обеспечения охватывает, например, приезжих из-за границы, не имеющих места жительства в Австрии, но не содержит информации о лицах, которые никогда не были застрахованы, не получали пенсионных выплат, пособий на детей или каких-либо других социальных услуг.

из статистики населения. Муниципалитеты также информируются о результатах и просят их привести в порядок свои административные документы.

57. В период между переписями аналогичная работа проводится ежегодно. В данном случае с людьми не связываются напрямую для подтверждения их присутствия, а часть лиц, идентифицированных как уникальные случаи в Центральном реестре проживания, статистически исключаются из общей численности населения. Вероятность быть номинальным членом определяется с помощью логистической регрессионной модели¹².

58. Результаты этого мероприятия включаются в квартальную статистику населения один раз в год (после публикации окончательных результатов). Расхождения между квартальными статистическими данными о численности населения и данными о номинальных членах, полученными в результате так называемой мини-переписи на основе реестра, устраняются двумя способами:

- а) переездом за границу непризнанных зарегистрированных резидентов; и
- б) переездом из-за границы ранее непризнанных резидентов.

59. Таким образом, корректировка численности населения для номинальных членов оказывает прямое влияние на миграционные потоки. Потенциальные номинальные члены, не признанные в популяционных запасах, считаются переехавшими за границу (и, следовательно, статистически учитываются как эмиграция). В свою очередь, ранее не признанные жители могут вновь войти в состав населения, если в следующем году они покажут признаки присутствия в других административных записях, кроме реестра жителей. В этом случае создается статистическая запись о повторном въезде в Австрию (иммиграция).

60. В отличие от прямой увязки информации из одного другого источника данных на микроуровне (как в шаге «(а) Корректировка с учетом неучтенных смертей в реестре населения» выше), используется другой методологический подход: во-первых, обращаются к нескольким источникам данных для выявления зарегистрированных жителей, которые являются потенциальными номинальными членами, а во-вторых, с помощью статистических методов оценки в конечном итоге отбираются лица, которые больше не будут признаваться жителями. Здесь связывание данных на уровне единичных записей, создающее статистически полученные данные об эмиграции и иммиграции, является лишь последующим процессом в оценке номинального числа членов на основе различных административных записей.

4.1.2 Венгрия

4.1.2.1 Первоначальный/основной источник данных для статистики миграции

61. В Центральном статистическом управлении Венгрии (HCSO) Департамент переписи населения и демографической статистики составляет ежегодные данные о численности населения, которые основываются как на реестрах актов гражданского состояния, так и на данных о миграции.

62. Число иностранцев с видом на жительство или документом на права поселения, а также лиц со статусом беженца и находящихся под дополнительной защитой лиц с зарегистрированным адресом в Венгрии, рассчитывается на основе регистраций в территориально-административных органах.

63. 4.1.2. Данные о миграции собираются с использованием различных типов реестров. Как представлено ниже, эти источники данных находятся в ведении различных органов.

4.1.2.1.1 Управление по вопросам иммиграции и предоставления убежища (IAO)

64. В 2009 году был подписан Меморандум о сотрудничестве между Центральным статистическим управлением Венгрии (HCSO) и Управлением по вопросам иммиграции и предоставления убежища (IAO). Управление по вопросам иммиграции и предоставления убежища отвечает за два различных типа реестров:

- а) Реестр ЕЕА: реестр содержит данные о гражданах ЕС и ЕЭЗ и сопровождающих их членах

¹² Подробная документация на немецком языке:
http://www.statistik.at/wcm/idc/idcplg?IdcService=GET_PDF_FILE&RevisionSelectionMethod=LatestReleased&dDocName=073537

семьи;

b) Реестр IDTV: этот реестр содержит данные о гражданах третьих стран.

65. Оба реестра созданы для административных целей. Сбор статистических данных является «побочным продуктом» ведения реестров.

66. На основании реестров Управления по вопросам иммиграции и предоставления убежища (IAO), можно выделить два типа иностранных мигрантов:

- Иностранные иммигранты – это иностранные граждане, которые подали заявление и получили вид на жительство от Управления по вопросам иммиграции и предоставления убежища и имеют регистрационное удостоверение (удостоверение жителя в случае граждан ЕЭЗ или вид на жительство в случае граждан третьих стран);
- Иностранные эмигранты – это лица, которые ранее владели документом, дающим им право на законное пребывание на территории Венгрии, но срок действия этого документа истек или их вид на жительство был аннулирован. С 2012 года их число включает оценочные значения, основанные на данных переписи 2011 года.

4.1.2.1.2 Министерство внутренних дел

67. Министерство внутренних дел ведет центральный реестр населения (CPR). Центральный реестр населения (CPR) содержит данные о лицах, иммигрирующих в Венгрию, и данные о лицах, эмигрирующих из Венгрии, которые обязаны зарегистрировать свой адрес в Центральном реестре населения (CPR) (как о венграх, так и об иностранцах; согласно действующему венгерскому законодательству, около двух третей иностранных граждан в Венгрии находятся в стране на законных основаниях).

68. В базе данных Центрального реестра населения (CPR) хранятся данные о лицах, которым предоставлена международная защита, и данные о лицах, натурализованных в Венгрии. Лицо, натурализованное в Венгрии, — это лицо, которое стало гражданином Венгрии путем натурализации (по рождению оно было иностранным гражданином) или повторной натурализации (его прежнее гражданство было аннулировано). Правила натурализации в Венгрии были изменены Законом XLIV от 2010 года. Закон ввел упрощенную процедуру натурализации с 1 января 2011 года и сделал возможным получение венгерского гражданства без проживания в Венгрии (для лиц, которые могут доказать, что у них есть венгерские предки).

69. Данные, опубликованные Центральным статистическим управлением Венгрии (HCSO), относятся только к тем новым гражданам Венгрии, которые имеют адрес в Венгрии.

4.1.2.1.3 Министерство трудовых ресурсов: Национальный фонд медицинского страхования Венгрии (NHIF)

70. Национальный фонд медицинского страхования Венгрии (NHIF) также является одним из основных источников данных о мигрантах, застрахованных в Венгрии, которые составляют подавляющее большинство мигрантского населения. Для расчета численности миграционного населения используются данные Национального фонда медицинского страхования Венгрии (NHIF), позволяющие установить количество иммигрантов и эмигрантов из числа граждан Венгрии (закон гласит, что граждане Венгрии должны сниматься с учета в Национальном фонде медицинского страхования Венгрии (NHIF), если они выезжают из страны в другое государство-член ЕЭЗ, и должны повторно регистрироваться, если они возвращаются в Венгрию после пребывания в стране ЕЭЗ). Таким образом, мигранты с венгерским гражданством могут быть разделены на следующие категории:

- Национальные эмигранты: лица, которые покинули Венгрию и сообщили об этом в Национальный фонд медицинского страхования (NHIF);
- Иммигранты, принадлежащие к одной из двух подгрупп:
 - граждане Венгрии, которые вернулись из-за границы и сообщили об этом в Национальный фонд медицинского страхования (NHIF);
 - Граждане Венгрии, которые получили венгерский адрес (и зарегистрировались в

реестре населения) после приобретения венгерского гражданства, но продолжая при этом проживать за границей (без предварительного проживания в Венгрии).

Рисунок 4.1 Источники данных для ежегодного расчета численности Венгрии



71. Перепись населения 2011 года выделена на рисунке, поскольку административные данные на 31 декабря 2011 года были согласованы с результатами переписи населения 2011 года, а данные последующих лет основаны на этом согласовании.

4.1.2.2 Ограничения исходного основного источника. Почему интеграция необходима?

72. Описанная выше венгерская система статистики миграции является очень сложной, поскольку соответствующие данные о миграции хранятся в различных реестрах. А в некоторых случаях, даже если реестры содержат данные, связанные с миграцией, то они хранятся в различных подреестрах. Это относится, например, к данным об иностранцах.

73. Для получения данных о количестве иммигрантов и эмигрантов (как граждан Венгрии, так и иностранцев) необходимо объединить данные о миграции из различных источников.

74. Необходимо подчеркнуть, что в случае Венгрии нет источника, который можно было бы определить как «основной» источник данных. Источники данных имеют практически одинаковую значимость для составления статистики, связанной с миграцией: то есть источники данных дополняют друг друга.

4.1.2.2.1 Подготовка данных о миграции (граждан Венгрии и иностранных граждан)

75. Для подготовки данных о потоках и численности мигрантов в Венгрии Центральное статистическое управление (HCSO) использует данные из реестров Управления по вопросам иммиграции и предоставления убежища (реестр ЕЕА и реестр IDTV). В ходе процедуры сначала происходит объединение этих двух баз данных, предоставленных Управлением по вопросам иммиграции и предоставления убежища, чтобы осуществить очистку данных и удаление дублирующих записей. Данные также связаны с Центральным реестром населения (CPR) – важно отметить, что Центральный реестр населения (CPR) содержит данные об иностранцах, имеющих регистрацию по месту жительства [граждане третьих стран, имеющие вид на жительство (краткосрочное пребывание), не обязаны регистрироваться в Центральном реестре населения (CPR)]. Во-вторых, реестры Управления по вопросам иммиграции и предоставления убежища (IAO) объединены с Центральным реестром населения (CPR), поскольку данные лиц, взятых под международную защиту, регистрируются в Центральном реестре населения (CPR).

76. Очистка первичных источников данных является важным этапом процедуры, поскольку первичный источник данных содержит множество дублирующихся записей.

77. В настоящее время Центральное статистическое управление Венгрии (HCSO) не использует никаких процедур интеграции данных для получения данных о миграции граждан Венгрии. Источником данных о потоках граждан Венгрии является только Национальный фонд медицинского страхования Венгрии (NHIF).

4.1.2.3 Методы, используемые для интеграции данных

78. Данные реестра Управления по вопросам иммиграции и предоставления убежища (IAO) и данные Центрального реестра населения (CPR) предоставляются Центральному статистическому управлению Венгрии (HCSO) два раза в год (в начале и в конце года). Данные предоставляются на уровне записей, которые связываются между собой с помощью ключевой переменной, разработанной с использованием программы SAS (от Института системного статистического анализа). Метод связывания представляет собой детерминированное связывание с помощью уникальных идентификаторов, что позволяет получить высококачественные интегрированные наборы данных. Конечно, этот метод интеграции предъявляет высокие требования к инфраструктуре административных данных. Несоответствия возникают из-за опечаток и изменений в именах.

79. Сначала данные Управления по вопросам иммиграции и предоставления убежища (IAO) очищаются для удаления дублирующихся записей, а затем очищенный набор данных соотносится с данными Центрального реестра населения (CPR).

4.1.2.4 Преимущества интеграции со статистическими данными из других источников

80. Интеграция данных о миграции является необходимостью для Центрального статистического управления Венгрии (HCSO), поскольку такого рода данные обычно хранятся в одном органе, а в Венгрии – в разных реестрах. Мотивация для интеграции данных возникает из-за определенного недостатка административных данных, но в результате интеграции создается дополнительная информация.

81. С другой стороны, такая интеграция позволяет получать высококачественные данные. Интеграция данных необходима для получения данных, связанных с миграцией, а также важна для ежегодного расчета численности населения.

4.1.3 Израиль

4.1.3.1 Оригинальный/первичный источник данных для статистики миграции

4.1.3.1.1 Для целей оценки численности населения

82. Израильское управление по вопросам народонаселения и иммиграции (PIA) отвечает за регистрацию в ходе пограничного контроля (BCR) и регистрацию в реестре населения (PR). Израильское управление по вопросам народонаселения и иммиграции (PIA) предоставляет Центральному бюро статистики (CBS) отдельные файлы в отношении израильтян и иностранцев. Эти административные данные позволяют проводить различие между иммигрантами и временными жителями (NPR) и дают возможность измерять потоки и численность эмигрантов и иммигрантов.

83. В отношении временных жителей (NPR) файлы, необходимые для составления их оценок, включают в себя их визу на момент въезда в страну. Хотя файлы включают различные типы виз, мы имеем в виду только визы для иностранных работников. Израильское управление по вопросам народонаселения и иммиграции (PIA) предоставляет Центральному бюро статистики (CBS) специальную регистрацию для временных жителей (NPR), которые въехали в Израиль нелегально, не через официальный пограничный пункт.

84. Для постоянных жителей Центральное бюро статистики (CBS) использует новую регистрацию в реестре населения (PR), проводя различие между новыми гражданами (в основном евреями или лицами еврейского происхождения, иммигрировавшими в Израиль в соответствии с Законом о возвращении) и негражданами (в основном лицами, прибывшими в рамках программы воссоединения семей).

85. Для оценки эмиграции Центральное бюро статистики (CBS) использует пограничный контроль (BCR) для анализа продолжительности пребывания за границей.

86. «Эмигрант» определяется как человек, который уехал за границу и прожил там не менее одного года. «Возвратившийся эмигрант» определяется как лицо, которое вернулось из-за границы после пребывания там не менее одного года и проживало в Израиле не менее трех месяцев подряд после возвращения из-за границы.

4.1.3.1.2 Для аналитических характеристик

87. Данные о постоянных и временных жителях, предоставляемые Израильским управлением по вопросам народонаселения и иммиграции (PIA), являются основным источником административных данных для аналитических исследований и политических нужд. Этот источник данных представляет собой перепись иммигрантского населения и, как таковой, дает возможность детально изучить небольшие группы населения. Этот источник данных о постоянных жителях содержит социально-демографическую информацию об иммигрантах, прибывших в Израиль с 1970 года. Получение информации о характеристиках, связанных с местом назначения в Израиле, страной происхождения, предполагаемой профессией или уровнем образования, происходит в момент въезда в страну. Для получения дополнительной административной информации мы связываем данные Израильского управления по вопросам народонаселения и иммиграции (PIA) с другими административными данными (образование, занятость, доход и т. д.).

88. Перепись населения и все исследования домохозяйств включают в себя модуль вопросов, которые позволяют нам узнать год иммиграции и страну рождения постоянных жителей. Этот модуль позволяет Центральному бюро статистики (CBS) получать информацию по всем темам исследований.

4.1.3.2 Ограничения оригинального или основного источника. Почему интеграция необходима?

4.1.3.2.1 Для целей оценки численности населения

89. Информация, доступная о иностранцах, не проживающих на постоянной основе, является информацией, полученной в день их въезда в страну. Эта информация не включает полные сведения о месте проживания. Информация, идентифицирующая личность в этих файлах, не позволяет связывать записи с другими источниками (административными или исследовательскими). Эти ограничения мешают Центральному бюро статистики (CBS) готовить данные, выходящие за рамки того, что есть в административных файлах. Кроме того, бывают случаи, когда невозможно установить связь между записями о прибытии и отправлении. Это может повлиять на нашу оценку численности. Это ограничение особенно актуально для записей о лицах с туристическими визами и затрудняет оценку числа лиц, превысивших срок пребывания (незарегистрированная миграция).

90. У многих израильтян есть гражданство другой страны. В большинстве случаев пограничный контроль (BCR) умеет выявлять двойное гражданство и регистрирует пересечения границы, даже если человек пересек границу с не израильским паспортом. Однако в некоторых случаях пограничный контроль (BCR) упускает информацию о двойном гражданстве. Это ограничение влияет на оценку численности эмиграции.

4.1.3.3 Методы, используемые для интеграции данных

4.1.3.3.1 Для целей оценки численности населения

91. В случае иммиграции постоянных жителей интеграция информации осуществляется на уровне отдельных записей. Все файлы содержат одинаковые переменные и связаны между собой с помощью уникального идентификатора (персонального идентификатора в реестре населения (PR)). Измерить количество иммигрантов, въезжающих в Израиль за определенный период времени, несложно, и корректировка данных не требуется. Информация о каждом лице, въезжающем в Израиль с иммиграционным статусом, доступна в административных файлах Израильского управления по вопросам народонаселения и иммиграции (PIA). Каждый месяц Израильское управление по вопросам народонаселения и иммиграции (PIA) готовит файл данных, содержащий записи об иммигрантах за предыдущий месяц, и предоставляет его Центральному бюро статистики (CBS), а также любые дополнения или обновления к уже полученным данным.

92. Среди временных жителей (NPR) выделяются три основные подгруппы, которые различаются с административной точки зрения: иностранные работники с разрешением на работу, нелегальные иностранные работники и иммигранты, нелегально пересекающие границу. Оценки по этим трем

группам должны предоставляться отдельно. В случае иностранных работников с разрешениями и нелегальных иностранцев мы получаем список от Израильского управления по вопросам народонаселения и иммиграции (PIA) и используем его без изменений. Для получения дополнительной информации о методологии в случае нелегальных работников см. Sheps M. (2016).

93. Что касается эмиграции, см. Sheps M. и Hleil A. (2006).

4.1.3.3.2 Для аналитических характеристик

94. Для получения полной информации или проведения продольных исследований Центральное бюро статистики (CBS) использует методы точного сопоставления записей, чтобы объединить информацию об иммигрантах с другими административными файлами (такими как налоговые файлы, файлы об образовании и т. д.), а также с переписями и исследованиями населения.

95. Сопоставление также используется для добавления переменных из источников данных, связанных с днем въезда иммигрантов в страну.

4.1.3.4 Преимущества интеграции со статистическими данными из других источников

4.1.3.4.1 Для целей оценки численности населения

96. Использование уникального идентификатора во всех административных файлах и сбор этого идентификатора в большинстве израильских исследований и переписи населения позволяют Центральному бюро статистики (CBS) объединять все эти источники. Основное преимущество этой операции заключается в предоставлении более качественных оценок, которые будут полезны на национальном уровне.

97. Кроме того, интеграция информации позволяет Центральному бюро статистики (CBS) снизить нагрузку на респондентов и сократить расходы на сбор данных.

4.1.3.4.2 Для аналитических характеристик

98. Центральное бюро статистики (CBS) позволяет своим сотрудникам и внешним пользователям проводить обработку данных и исследования с использованием объединенных данных. Исследователи получают доступ к данным в исследовательских комнатах, расположенных в зданиях Центрального бюро статистики (CBS). Не существует стандартного файла данных об иммигрантах, доступного для исследователей. Обычно каждый исследователь определяет свои данные в соответствии с потребностями исследования. Эти данные позволяют пользователям исследовать процесс ассимиляции и интеграции иммигрантов в израильское общество и экономику, а также отслеживать их динамику в составе населения. Семейные отношения, зафиксированные в реестре населения (PR), позволяют пользователям получить информацию и о втором поколении.

99. В случае эмиграции были проведены исследования характеристик эмигрантов до и после отъезда. Связанные данные об эмигрантах из различных источников позволяют, например, изучать феномен «утечки мозгов».

4.1.4 Италия¹³

4.1.4.1 Основной источник данных для статистики миграции: реестры населения (Anagrafi)

100. С начала 1990-х годов Итальянский национальный институт статистики (Istat) пытается максимально эффективно использовать имеющиеся административные источники для составления статистики международной миграции, в частности реестры населения (Anagrafi). Это привело к созданию обширной статистической документации, отражающей основные социально-демографические особенности миграции.

¹³ Подготовлено Управлением социальной статистики и переписи населения Итальянского национального института статистики (Istat).

101. В последние годы произошло много изменений, как в административной сфере, так и в области управления статистическими данными и процедур их сбора. Цифровизация реестров населения и изменения в законодательстве повысили эффективность административных систем и снизили административную нагрузку на граждан. Такие изменения оказали большое влияние на производство статистики и позволили улучшить качество данных с точки зрения своевременности, охвата и согласованности с другими административными источниками.

102. В Италии официальная статистика по международным миграционным потокам основывается на данных местных реестров населения, которые ведет Министерство внутренних дел, осуществляющее надзор за местными органами власти и национальную безопасность. Данные об изменении места жительства позволяют определить размер международных миграционных потоков, получить информацию как о месте отправления, так и о месте назначения отдельного перемещения, а также об основных социально-демографических характеристиках мигрантов. Итальянский национальный институт статистики (Istat) ежегодно собирает индивидуальные данные из более чем 8 000 муниципалитетов. Собранные данные не могут быть использованы в их нынешнем виде, поскольку в них присутствуют проблемы с качеством данных, вызванные ошибками при их вводе, отсутствием информации или другими недействительными данными. Проверенные данные сверяются с данными в агрегированной форме (демографический баланс), и несоответствия устраняются с целью получения уникальных, более точных и согласованных статистических данных в соответствии с требованиями Регламента Европейского парламента и Совета европейской статистике в области миграции и международной защиты¹⁴.

4.1.4.2 Ограничения исходного источника. Почему интеграция необходима?

103. Хотя за последнее десятилетие качество значительно улучшилось, точность этого реестра по-прежнему остается серьезной проблемой, в основном из-за сложности регистрации международной эмиграции, а также из-за расхождений между административными концепциями и международными определениями. Это приводит к трем основным ограничениям исходного источника:

- Чрезмерное покрытие: в Италии граждане, проживающие за границей, и иностранцы, которые покинули страну навсегда или на длительный срок, должны быть исключены из реестра населения. Однако эмигранты не имеют достаточной мотивации для уведомления властей о своем отъезде и могут оставаться в реестрах в течение многих лет, прежде чем их запись будет удалена в порядке «снятия с регистрации по должностному распоряжению»;
- Недостаточное охватывание: статистические данные относятся к «де-юре» проживающему населению, а не к «де-факто»: иностранное население, обычно проживающее в Италии, но не имеющее официального места жительства, исключается из статистики. Хотя охват иностранного населения, обычно проживающего в Италии, контролируется, есть иностранные граждане, которые не хотят быть зарегистрированными в реестре населения, особенно граждане стран-членов Европейского Союза. Кроме того, лица, которые больше не соответствуют требованиям для пребывания в Италии, и иммигранты, въехавшие без разрешения, не могут быть зарегистрированы в реестре населения. Последние, которые не фигурируют ни в одном административном источнике, представляют собой трудно поддающуюся подсчету подгруппу населения, и для их подсчета требуется проведение соответствующего исследования/опроса;
- Временные критерии: источник не содержит информации о предполагаемой продолжительности пребывания, хотя вполне вероятно, что лица, регистрирующиеся (или снимающиеся с регистрации), намерены сменить место жительства как минимум на 12 месяцев.

4.1.4.3 Методы, используемые для интеграции данных

104. Решением проблемы охвата реестрами населения является использование интегрированной

¹⁴ Регламент (ЕС) № 862/2007 Европейского парламента и Совета от 11 июля 2007 г. о статистике Сообщества в области миграции и международной защиты и об отмене Регламента Совета (ЕЭС) № 311/76 о составлении статистики по иностранным работникам. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32007R0862>

системы реестров: официальные реестры населения могут быть связаны с другими тематическими административными источниками (имеющим отношение к таким сферам, как занятость, образование, налогообложение, доходы и т. д.). Эта система может использоваться для выявления групп, которые соответствуют национальному или международному определению «постоянно проживающего населения». Недостаточное охватывание в реестрах населения можно выявить с помощью сигналов о присутствии лиц на территории Италии, поступающих из других реестров; в то же время отсутствие сигналов о лицах в реестре населения может свидетельствовать о чрезмерном охвате.

105. Итальянский национальный институт статистики (Istat) организовал рабочую группу с целью проверки использования административной информации для проведения переписи населения и подсчета численности населения. Здесь приведены результаты исследования, проведенного специалистами Итальянского национального института статистики (Istat).

106. Интеграция данных необходима для определения предварительного рабочего процесса обработки административных данных с целью получения данных о численности обычно проживающего населения. Итальянский национальный институт статистики (Istat) создал интегрированную систему административных источников для управления растущим количеством наборов административных данных, полученных для статистических целей, и для максимального использования преимуществ, которые дает огромный объем доступной информации. Для этого потребовалось централизовать некоторые функции по сбору, хранению, интеграции и административной оценке качества данных в системе под названием SIM (Интегрированная система микроданных). Эта система управляет социальными и экономическими административными данными о характеристиках физических лиц и домохозяйств (демографические аспекты, статус занятости, уровень образования, место жительства, место работы и место обучения). Этап интеграции в системе SIM представляет собой процесс сопоставления и физической интеграции микроданных, поступающих из разных источников: в зависимости от доступных переменных сопоставления применяется подходящая стратегия интеграции и набор алгоритмов.

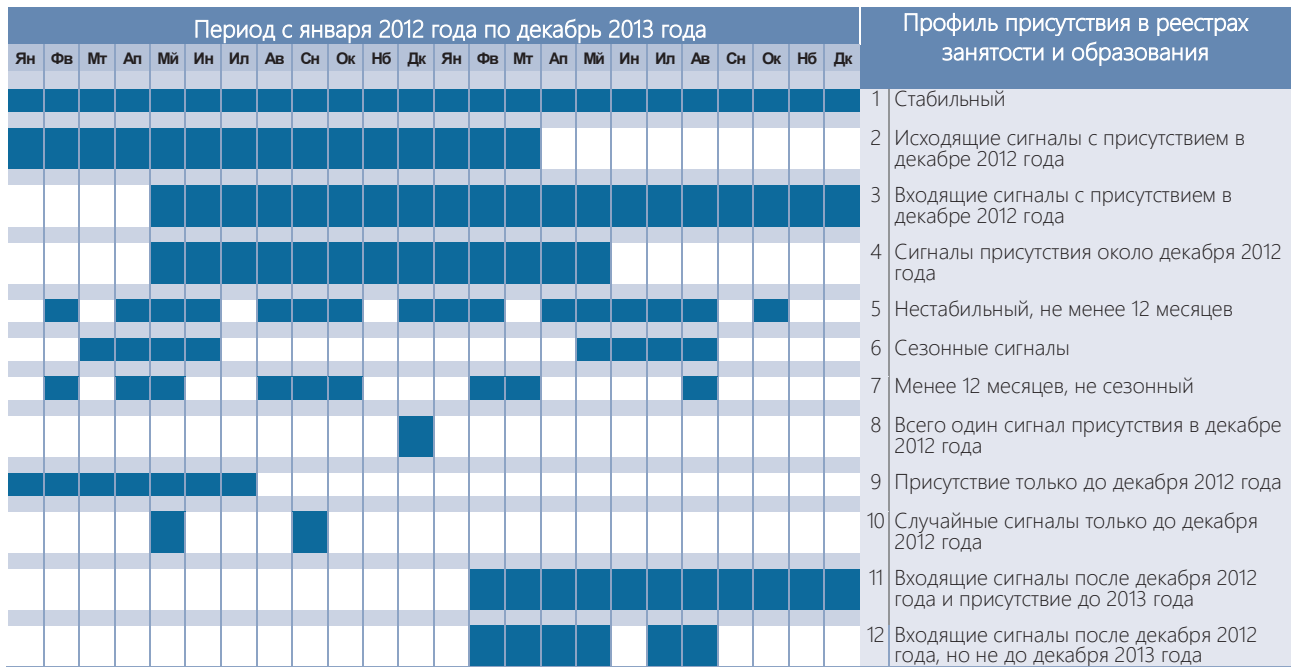
107. Интеграция означает:

- идентификацию каждого объекта (физического лица, экономической единицы) в административных источниках данных с помощью идентификационного номера (ID), который является уникальным и стабильным во времени;
- определение для каждого объекта логических и физических отношений между наборами данных, поступающими из разных источников.

108. Система SIM была создана с целью поддержки процесса производства статистических данных: использование идентификаторов SIM позволяет формировать структуры данных, которые могут служить отправной точкой для статистических процессов, основанных на административных данных. Для оценки того, как административные данные могут улучшить подсчет населения, была создана тематическая база данных «административных сигналов» в поддержку первоначального эксперимента. Данные, использованные в исследовании, были получены из конкретных административных источников, которые уже были сохранены и интегрированы в систему SIM: реестры населения, видов на жительство, наемных работников и самозанятых лиц, лиц, проходящих программу получения обязательного образования, студентов университетов, пенсионеров, непенсионных выплат, доходов и налогов.

109. Благодаря стандартизированной временной шкале продолжительности отдельных процессов деятельности человека удалось разработать схему ежемесячного присутствия. Эта функция состоит из последовательности длиной 24 цифры, каждая из которых указывает, присутствовал ли человек в ходе какого-либо процесса своей деятельности и в каком географическом районе страны. Кластеризация по этим последовательностям, очевидная в большинстве сигналов от процессов деятельности человека, показала наличие паттернов непрерывности (рис. 4.2).

Рисунок 4.2 Ежемесячная схема присутствия моделей непрерывности в трудовой и учебной деятельности



110. Оценка согласованности между исследованными реестрами приводит к классификации всех случаев, зарегистрированных в экспериментальной базе данных, по группам, которые полезны при подсчете численности населения. Процесс выделения этих групп состоял из трех этапов (табл. 4.1). Первый этап включал сопоставление экспериментального реестра населения (PR), сигналов о деятельности человека из реестров труда и образования (LE) и реестра видов на жительство (PS); на втором этапе среди всех лиц, не имеющих сигналов из реестров труда и образования (LE), были выявлены пенсионеры или лица, получающие другие виды пособий; наконец, реестр налоговых деклараций (TR) предоставил косвенные сигналы от лиц, которые могли каким-либо образом подтвердить свое присутствие в Италии.

111. Эти группы можно организовать в соответствии с оценкой обычного места жительства. Лица, относящиеся к группе А, группе Е и группе G, могут считаться «правомочными» в качестве обычно проживающего населения с использованием концепции обычного места жительства с критерием в 12 месяцев. Напротив, группы С, D и H требуют более тщательного анализа (табл. 4.1).

Таблица 4.1 Схема процесса и численность групп населения (в тысячах) в зависимости от вероятности их обычного постоянного проживания в Италии

Шаг I			Шаг II				Шаг III			
PR	LE	PS	Группа	Кол-во	RR-NPR	Группа	Кол-во	TR	Группа	Кол-во
61 068	37 704	3 378			20 764			26 649		
Сигналы			Сигналы				Сигналы			
Да	Да	-	A	36 618						
Да	Нет	-	B	24 450	Да	E	14 485			
					Нет	F	9 965	Да	G	6 939
Нет	Да	-	C	1 086				Нет	H	3 026
Нет	Нет	Да	D	351						

Пояснения к таблице:

- Имеющие право на статус обычного жителя
- Неопределенные жители

Список сокращений административных источников:

- PR Реестры населения
- LE Реестры труда и образования
- PS Реестр видов на жительство
- RR-NPR Реестры пенсионных и непенсионных выплат
- TR Реестр налоговых деклараций

112. В конце 2012 года в группе Н было зарегистрировано около 3 миллионов человек, что представляет собой подгруппу населения, зарегистрированную в реестре населения (PR) без сигналов из других источников. Почти у 75 процентов этих людей было итальянское гражданство. В связи с отсутствием записей в трудовых и образовательных реестрах группа Н представляет высокий риск неучтенной эмиграции, что может привести к искажению данных реестра населения.

4.1.4.4 Преимущества интеграции со статистическими данными из других источников

113. Интеграция итальянских административных источников данных оказалась очень эффективной для использования уже имеющейся обширной информации. Однако, поскольку сигналы могут быть результатом временного или эпизодического присутствия, необходимо провести процесс формирования характеристик путем построения производных переменных, которые затем позволяют выявить случаи постоянного присутствия (в ом или ином процессе деятельности человека), соответствующие международному определению обычного места жительства. В ходе исследования Итальянский национальный институт статистики (Istat) определил предварительный рабочий процесс для интеграции использования административных источников и официальных реестров населения с целью расчета численности постоянного проживающего населения. С помощью этого рабочего процесса можно определить группу лиц, которые на определенную дату могут быть включены в число обычно проживающих в Италии. Эта группа потенциальных постоянных жителей насчитывала по состоянию на 31 декабря 2012 года 62,6 млн человек. Из этой группы лиц можно выделить три основные подгруппы:

- a) Подгруппа населения, присутствующая в реестре населения, без сигналов из других источников (3,0 млн человек в 2012 г.), которые могли эмигрировать, не уведомив о своем отъезде;
- b) Подгруппа населения, присутствующая в реестре населения, которая показала признаки в других источниках (58,1 млн человек в 2012 г.);
- c) Подгруппа населения, не зарегистрированная в реестре населения, но имеющая признаки из других источников (1,5 млн человек в 2012 г.), которые могли иммигрировать, не уведомив о своем прибытии.

114. Анализ силы сигнала на основе его непрерывности во времени является лишь отправной точкой в использовании продольных данных, которые могут быть обработаны с помощью административных источников. Поэтому основной целью будущих исследований должно быть изучение продольных моделей в течение нескольких лет с целью получения оценок субпопуляций, которые будут более стабильными в отношении колебаний, связанных с рынком труда, от которого и приходят сигналы.

4.1.4.5 Список использованной литературы

Tucci E., Marsili M. and Terra Abrami V. "Improving quality of international migration outcomes by incorporating the micro-approach in managing current demographic accounting (MIDEA) and statistical population registers (ANVIS)", UNECE/Eurostat Work Session and Workshop on Migration Statistics, 10-12 September 2014, Chisinau, Republic of Moldova («Повышение качества результатов международной миграции путем внедрения микроподхода в управление текущим демографическим учетом (MIDEA) и статистическими реестрами населения (ANVIS)», Рабочая сессия и семинар ЕЭК ООН/Евростат по статистике миграции, 10-12 сентября 2014 г., Кишинев, Республика Молдова)

Prati, Gallo, Chieppa, Tomeo et al. "Feasibility study of using usual residence concept with 12 month's criteria for all the breakdowns of population, births and deaths requested in the regulation (art. 8)", note prepared by the National Institute of Statistics of Italy for the REGULATION EU 1260/2013 on European Demographic Statistics («Технико-экономическое обоснование использования концепции обычного места жительства с критерием 12 месяцев для всех разбивки населения, рождаемости и смертности, требуемых в регламенте (ст. 8)», записка, подготовленная Национальным институтом статистики Италии для РЕГЛАМЕНТА ЕС 1260/2013 о европейской демографической статистике)

4.1.5 Латвия

4.1.5.1 Первоначальный/основной источник данных для статистики миграции

115. Численность населения оценивается на основе данных, предоставленных Управлением по вопросам гражданства и миграции (ОСМА) из подсистемы «Реестр населения» Единой миграционной информационной системы. Индивидуальные данные из этих файлов включают информацию о поле, дате рождения, стране, гражданстве, юридическом семейном положении, коде административно-территориальной единицы заявленного места жительства в соответствии с Классификацией территорий и территориальных единиц (САТТУ), виде на жительство.

116. С января 2013 года Центральное статистическое бюро (CSB) также получает информацию о регистрации актов гражданского состояния (брака, рождения или смерти) из подсистемы «Единая система регистрации актов гражданского состояния» (CSRCS) Единой миграционной информационной системы (ОСМА). Эта информация поступает в виде ежегодных отдельных файлов данных из каждого источника (ранее информация поступала из районных отделов записи актов гражданского состояния в бумажном виде). С помощью вышеуказанных данных формируется информация о годовом числе рождений и смертей.

117. Анализ данных переписи населения и жилищного фонда 2011 года позволил Центральному статистическому бюро (CSB) определить численность и состав населения страны. На 1 января 2011 года население Латвии составляло 2 074,6 тыс. человек. По сравнению с ранее опубликованной информацией (согласно Реестру), численность населения была на 155 тысяч человек или 7 процентов меньше. Данные подтвердили, что часть населения страны не соответствует требованиям Закона о регистрации населения, а информация, содержащаяся в реестре миграции населения, является неполной.

4.1.5.2 Ограничения оригинального или основного источника. Почему интеграция необходима?

118. В реестре населения были выявлены следующие проблемы:

- Неподача заявления о снятии с регистрационного учета или несвоевременная подача заявления о снятии с регистрационного учета (незарегистрированная миграция). Хотя согласно Закону о регистрации населения люди обязаны сообщать в регистрационный орган об изменениях в своем законном семейном положении, детях, обычном месте жительства и т. д., но они не всегда это делают. Согласно результатам переписи 2011 года 7 % населения Латвии переехали за границу, но не изменили свое обычное место жительства в реестре (согласно реестру они все еще проживали в Латвии). Кроме того, анализируя информацию из реестра населения через несколько лет после переписи, можно увидеть, что довольно много людей с задержкой исправляют информацию в реестре;
- Очень пожилые люди, которые внесены в реестр, но чье обычное место жительства находится не в Латвии. В случаях, когда такие люди умерли за границей, а в реестре населения эта информация не была зарегистрирована, они остаются в реестре как живые. В худших случаях человек, который покинул страну без снятия с регистрационного учета и, следовательно, продолжает числиться в реестре как проживающий в Латвии, мог умереть за границей.

4.1.5.3 Методы, используемые для интеграции данных

119. В 2012 году Центральное статистическое бюро (CSB) разработало новый метод оценки численности населения. Метод основан на реестре населения и индивидуальных данных из других административных реестров (у всех источников данных идентификационные коды, которые используются для объединения данных).

120. Для оценки статуса проживания каждого человека была разработана модель логистической регрессии. С помощью данных административного реестра для каждого человека, зарегистрированного в реестре населения, разрабатывается более 200 переменных характеристик. Цель модели – предсказать вероятность (число, попадающее в интервал от нуля до единицы) того,

что каждый человек является жителем, чтобы оценить статус фактического места жительства на начало года для каждого зарегистрированного жителя (на уровне каждого отдельного человека) в Латвии. Вероятность обязательного включения в число постоянно проживающих людей зависит от возраста и пола.

121. Модель была разработана с использованием данных переписи населения 2011 года о фактическом месте проживания и данных из административных источников за 2010 год по состоянию на 1 января 2011 года или по состоянию на 1 марта 2011 года.

122. Ежегодно составляется база данных по каждому источнику с данными по состоянию на 1 января. База данных позволяет составлять сводные таблицы данных для национальных и международных пользователей данных.

123. Помимо данных Единой миграционной информационной системы (OCMA), статистические данные по населению составляются с использованием индивидуальных данных из других административных реестров, доступных Центральному статистическому бюро (CSB), которые соответствуют рассматриваемому периоду (начиная с 2010 года). Центральное статистическое бюро (CSB) имеет доступ к данным из следующих административных реестров:

- Государственная налоговая служба (SRS);
- Государственное агентство социального страхования (SSIA);
- Министерство образования и науки (MES);
- Центр сельскохозяйственных данных (ADC);
- Служба поддержки сельских районов (LRSS);
- Национальная служба здравоохранения (NHS);
- Государственное агентство по трудоустройству (SEA);
- 32 из 58 высших учебных заведений (89% от общего числа студентов);
- Управление безопасности дорожного движения;
- Служба по исполнению наказания.

124. Центральное статистическое бюро (CSB) также располагает информацией из информационной системы администрации социального обеспечения (SSAIS) о лицах, получавших социальные пособия от местных органов власти. Однако эта информация доступна только начиная с 2012 года. Эти данные используются для корректировки численности населения в возрастной группе 18–30 лет и оценки качества модели.

125. Начиная с 2016 года, численность населения определяется на основе данных Государственного агентства социального страхования (SSIA) о лицах, которым выплачивались пособия или пенсии и которые проживают в учреждениях социального обеспечения. Кроме того, проводится дополнительный анализ в отношении обычного места жительства матери, если обычным местом жительства ребенка (младше 16 лет) и отца, по оценкам, является Латвия.

126. Международная долгосрочная эмиграция из Латвии в другую страну теоретически соответствует данным о международной долгосрочной иммиграции из Латвии в соответствующую страну. Такая «зеркальная статистика» используется для оценки международной долгосрочной эмиграции из Латвии.

127. При оценке международной долгосрочной эмиграции информация об иммиграции из Латвии поступает из других стран, а именно из Дании, Финляндии, Швеции, Норвегии, Испании, Нидерландов, Австрии, Исландии и Германии. Не все страны готовят данные об иммигрантах из Латвии, поскольку статья 3 Регламента Европейского парламента и Совета европейской статистике по вопросам миграции и международной защиты предусматривает, что страны могут предоставлять информацию об иммигрантах с групповой разбивкой по предыдущему месту обычного проживания. Таким образом, разбивка может быть сгруппирована по странам-членам ЕС; странам Европейской ассоциации свободной торговли; странам-кандидатам; другим странам, не являющимся членами ЕС.

128. Для оценки эмиграции латвийского населения в Великобританию и Ирландию используется

следующая информация: количество впервые выданных номеров национального страхования Великобритании и количество впервые выданных личных номеров государственных служб Ирландии. Следует отметить, что эти данные используются только для оценки общих тенденций, поскольку эти идентификационные номера присваиваются жителям Латвии, даже если они пребывают в Великобритании или Ирландии менее одного года.

129. Для оценки качества оценок используются данные исследования домохозяйств (данные на каждому отдельному случаю из исследования трудоспособного населения (LFS), исследования «Статистика ЕС по доходам и условиям жизни» (EU-SILC), Европейского исследования здоровья и социальной интеграции (EHSIS) и Европейского исследования здоровья (EHIS). Оценка результатов этого метода была одной из задач микропереписи 2015 года.

130. Подробная информация о методологии доступна на интернет-сайте Центрального статистического бюро (CSB) (http://www.csb.gov.lv/sites/default/files/dati/demstat_metodologija_eng.pdf).

4.1.5.4 Преимущества интеграции со статистическими данными из других источников

131. Основное преимущество интеграции различных административных источников данных заключается в предоставлении более точных оценок численности населения, которые будут полезны не только для статистического управления, но и для национальных и международных пользователей данных. Разработанный метод будет использоваться для переписи 2021 года, которая будет проводиться на основе внесенных в реестр данных.

132. Представленный метод получил положительные отзывы от Латвийской статистической ассоциации, экспертов Центрального банка Латвии и демографов. Основным выводом заключался в том, что если информация о месте обычного проживания в реестре населения имеет столь низкое качество, то разработанный метод является наилучшим решением.

4.1.6 Нидерланды

4.1.6.1 Основные источники для измерения статистики миграции

133. Демографическая статистика Нидерландов полностью основана на данных нидерландских реестров населения. Таким образом, описание статистических данных по иммиграции в Нидерландах сводится в основном к описанию определений и практик, используемых в реестрах населения. Однако существуют некоторые ограничения. Статистическое управление Нидерландов сталкивается с этими ограничениями, интегрируя административные данные из многих источников в Систему социальных статистических наборов данных (SSD).

134. Каждый, кто въезжает в Нидерланды, регистрируется в качестве жителя (иммигранта) при условии, что:

- Его/ее пребывание является законным в соответствии с Законом об иммиграции (для лиц, не имеющих голландского гражданства);
- Предполагаемый срок пребывания составляет не менее двух третей предстоящих шести месяцев;
- Лицо идентифицировано надлежащим образом. Последнее означает, что для идентификации предъявляется действующий паспорт или другой официальный документ.

135. Каждый ребенок, родившийся в Нидерландах, чья мать зарегистрирована в качестве жителя, также регистрируется в качестве жителя. Дети, родившиеся за границей у матери, которая сама зарегистрирована как жительница Нидерландов, регистрируются (как иммигранты) при условии, что эти дети будут проживать в стране.

136. Эмиграция касается лиц, которые покидают Нидерланды и намереваются прожить за границей не менее двух третей из предстоящих двенадцати месяцев, а также сообщают о своем отъезде в муниципальные органы власти. Они включены в статистику эмиграции из Нидерландов.

4.1.6.2 Ограничения оригинального или основного источника. Почему интеграция

необходима?

137. Существует два исключения из правил регистрации иммигрантов, в результате чего регистрация не является абсолютно полной.

138. Первое относится к так называемым «привилегированным лицам», включая иностранцев, работающих на территории Нидерландов в качестве дипломатических, консульских или военных должностных лиц или в международных организациях. Поскольку они пользуются особым «привилегированным» статусом и не считаются иностранцами в соответствии с Законом об иммиграции, им предоставляется право выбора, вносить ли их в реестр населения.

139. Просители убежища составляют собой еще одно исключение. Их регистрация происходит только через шесть месяцев после их прибытия в Нидерланды, независимо от предполагаемой продолжительности их пребывания, если им только не был выдан вид на жительство в течение этих шести месяцев. В этом случае они регистрируются при выдаче вида на жительство. Однако дети, рожденные от просителей убежища, которые еще не зарегистрированы, регистрируются сразу после рождения. Это приводит к несколько странной ситуации, когда новорожденный ребенок зарегистрирован, а родители и братья и сестры (пока) нет.

140. Что касается эмиграции, то примерно каждый третий человек, покидающий страну, не уведомляет муниципальные власти о своем отъезде. Когда власти узнают, что кто-то «пропал», закон обязывает их выяснять его или ее местонахождение (проводить расследование). Если в результате расследования делается вывод о том, что данное лицо по-прежнему считается пропавшим без вести, оно регистрируется в административном порядке как эмигрировавшее в страну, квалифицируемое как «неизвестная». В итоге они включаются в статистику административных выбытий. Если они возвращаются в Нидерланды, они включаются в статистику административных прибытий.

4.1.6.3 Описание методов, используемых для интеграции данных

141. Все рабочие процессы в Статистическом управлении Нидерландов, касающиеся социальной или пространственной статистики, основаны на Системе социальных статистических баз данных (SSD). Последняя содержит обширную информацию о лицах, домохозяйствах, рабочих местах, пособиях, пенсиях, образовании, госпитализациях, сообщениях о преступлениях, жилых помещениях, транспортных средствах и многом другом. До внедрения Системы социальных статистических баз данных (SSD) различные статистические реестры были разбросаны по всему Статистическому управлению Нидерландов и не были стандартизированы. В результате процесса микроинтеграции Система социальных статистических баз данных (SSD) создала библиотеку стандартизированных и связанных между собой статистических реестров, а также меры, которые были введены для контроля ее различных аспектов.¹⁵

142. Когда новые реестры поступают в Статистическое управление Нидерландов, они также объединяются и в Системе социальных статистических баз данных (SSD) с использованием того самого подхода относительно стандартизации, связывания и микроинтеграции. Недавно Статистическое управление Нидерландов добавило в Систему социальных статистических баз данных (SSD) данные Центрального агентства по приему просителей убежища (COA). Это решит описанную выше проблему, поскольку просители убежища будут регистрироваться с первого дня их прибытия в Нидерланды. Есть надежда, что в ближайшее время к Системе социальных статистических баз данных (SSD) будут добавлены и другие источники, такие как PROBAS (база данных, содержащая информацию о привилегированных лицах) или RNI (муниципальный реестр лиц без права постоянного или временного жительства).

4.1.6.4 Преимущества интеграции со статистическими данными из других источников

143. Преимущества интеграции со статистическими данными из других источников очевидны. Прежде всего, доступны интегральные данные, что снижает необходимость в проведении исследований. Это снижает затраты на сбор данных и увеличивает количество респондентов. Преимущества стандартизации и микроинтеграции данных заключаются в том, что информация из разных

¹⁵ Для получения дополнительной информации о Системе социальных статистических баз данных (SSD) см.: Bakker, B., Van Rooijen, J., & Van Toor, L. (2014). Система социальных статистических наборов данных Статистического управления Нидерландов: Комплексный подход к производству регистровой социальной статистики. *Journal of the International Association for Official Statistics (Журнал Международной ассоциации официальной статистики)*, 30: 411–424.

источников может быть связана с одними и теми же лицами. В результате получается библиотека данных, содержащая более пятидесяти административных реестров. Таким образом, интегрированные данные Статистического управления Нидерландов представляют собой богатый источник информации, который позволяет исследователям не только измерять миграцию, но и дает возможность изучать социально-экономическую интеграцию с помощью широкого спектра переменных.

4.1.7 Испания¹⁶

4.1.7.1 Введение

144. Национальный институт статистики (INE-Spain) в настоящее время не применяет методы интеграции данных для составления статистики миграции. Единственным источником, из которого получают эти статистические данные, является реестр населения, хотя для улучшения совокупных показателей по потокам проводится статистическая обработка исходных данных, содержащихся в реестре населения.

145. Однако перепись населения 2011 года в Испании стала ярким примером интеграции различных источников для улучшения оценки численности населения, и в особенности численности иностранных жителей в Испании. В результате были пересмотрены миграционные потоки за период 2008–2011 гг. Ниже описывается метод, использованный для интеграции данных в переписи 2011 года.

4.1.7.2 Основной источник данных для статистики миграции: реестр населения Padrón

146. Основным источником данных в Испании как по численности населения, так и по статистике миграции является реестр населения Padrón. Padrón — это официальный список жителей каждого из 8 100 муниципальных образований Испании. Padrón, как список жителей каждого муниципального образования, является продолжением очень старой испанской традиции, которая восходит к средневековью.

147. В Испании существует столько же реестров, сколько и муниципальных образований. Однако с 1996 года действует закон, который объединяет эти муниципальные списки в единую национальную базу данных. Существуют также правовые процедуры, обеспечивающие взаимосвязь между этой базой данных и муниципальными архивами, а также их ежемесячное обновление.

148. Закон обязывает муниципальные образования этими записями. Это осуществляется через Национальный институт статистики (INE-Spain). Таким образом, в отличие от других стран, где за ведение реестров населения отвечают полиция или другие административные органы, в Испании Национальный институт статистики (INE-Spain) является национальным учреждением, координирующим этот единый национальный реестр населения. Каждый месяц Национальный институт статистики (INE-Spain) получает информацию обо всех изменениях, произошедших в каждом муниципалитете, проводит проверку и направляет эти результаты во все муниципалитеты, чтобы избежать дублирования, а также включить данные о смертях, рождаемости или получении испанского гражданства, которые он (Национальный институт статистики (INE-Spain)) получает ежемесячно (или чаще) из Гражданского реестра.

149. Все консульские учреждения Испании по всему миру (около 250) также подключены к системе Padrón таким же образом, как и муниципалитеты. Любой гражданин Испании, покидающий муниципалитет в Испании для проживания за границей, должен заявить о своем новом месте жительства в ближайшем консульстве. Консульские учреждения направляют эту информацию в Национальный институт статистики (INE-Spain), а он уже передает эти данные муниципалитетам в сформированных ежемесячных координационных файлах.

150. Согласно закону с юридической точки зрения нет никаких ограничений для регистрации в Испании. Все люди, проживающие или желающие проживать в Испании, независимо от их правового статуса, имеют право на регистрацию (фактически это является обязательным требованием) и, как правило, регистрируются, поскольку регистрация дает много преимуществ и не имеет никаких

¹⁶ Подготовлено Национальным институтом статистики (INE-Spain).

недостатков. Преимущества регистрации могут включать в себя бесплатный доступ к системе общественного здравоохранения, государственным школам, услугам и правам, предоставляемым муниципалитетами, а также более дешевый общественный транспорт и многое другое.

151. В отношении каждого человека в реестре указаны пол, дата рождения, место рождения (страна, в случае иностранцев), гражданство, уровень образования и национальный идентификационный номер. Иностранцы, проживающие на законных основаниях, имеют идентификационный номер. Для лиц с нелегальным статусом сохраняется номер паспорта.

152. Padrón также является продольной базой данных: в ней хранятся все предыдущие места проживания (в пределах Испании), что позволяет проводить продольный анализ и очень тщательно и точно отслеживать внутреннюю миграцию.

4.1.7.3 Ограничения исходного источника. Почему интеграция необходима?

153. Как уже отмечалось, система Padrón является источником данных о численности населения и статистики миграции. Это не означает, что численность населения напрямую соответствует количеству людей, зарегистрированных на данный момент. Для получения данных о численности населения к исходным данным, содержащимся в реестре населения, применяются некоторые статистические методы обработки.

154. Основным недостатком системы, основанной на реестре населения, подобной этой, является сложность измерения эмиграции. Что касается иммиграции, то зарегистрированные притоки населения, вероятно, ближе к реальным значениям. Люди могут найти много причин для регистрации, но не так много для снятия с регистрации.

155. Когда в 1996 году был создан реестр населения, число иностранных граждан составляло 540 000 человек (1,4% населения Испании, которое на тот момент насчитывало 39,7 млн человек). Несколькими годами ранее, в переписи 1991 года, этот показатель был почти таким же (470 000). В первой половине 1990-х годов в Испания была очень небольшая чистая международная миграция.

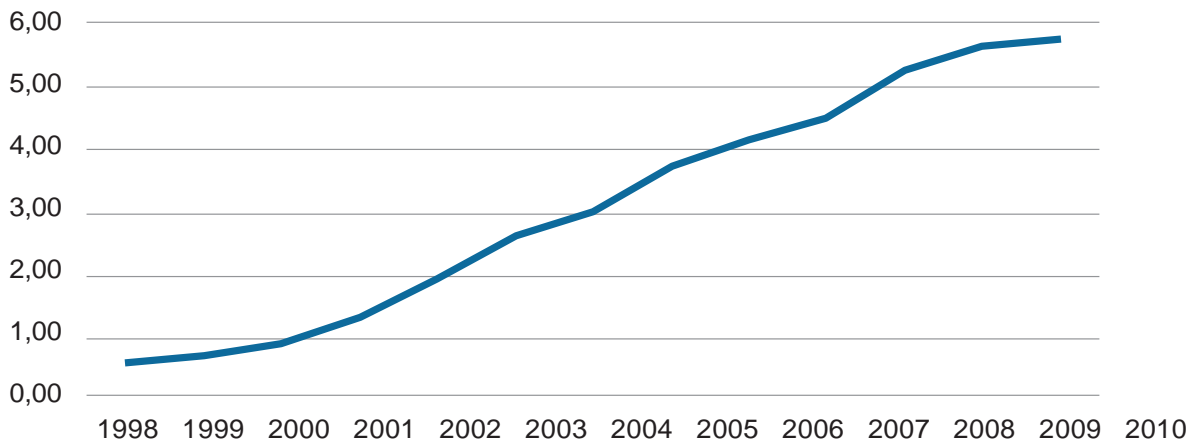
156. Вторая половина 1990-х годов и особенно первое десятилетие XXI века были совсем другими. В связи с ростом экономики количество иностранных жителей резко возросло. Всего за десять лет (с 2000 по 2010 год) число иностранных жителей, согласно данным реестра населения, выросло с 1 миллиона до почти 6 миллионов (с 1,4% до 12,2% населения в 2010 году).

157. Реестр должен был решить важную задачу, для которой он не был предназначен: как подсчитывать иностранцев, а точнее, как измерить эмиграцию. Не существовало никакой процедуры для снятия с регистрации иностранных граждан, не заявивших о своем отъезде.

158. В 2006 году была утверждена новая юридическая процедура. В зависимости от своего правового статуса иностранные граждане могут быть разделены на две группы: Граждане ЕС или граждане стран, не входящих в ЕС, но имеющие вид на жительство (назовем их «иностранцами-гражданами ЕС»), и остальные, т. е. граждане стран, не входящих в ЕС, с временным видом на жительство или без вида на жительство («иностранцы-неграждане ЕС»).

159. С 2006 года все иностранцы, не являющиеся гражданами ЕС (примерно 50 % от общего числа иностранцев на тот момент), обязаны каждые два года продлевать свою регистрацию в системе Padrón. Администрация каждого города должна отправлять уведомления всем иностранцам, не являющимся гражданами ЕС, чья регистрация скоро истекает. Если они не продлевают свою регистрацию, после административной процедуры, которая может занять несколько месяцев, они исключаются из муниципальных реестров.

Рисунок 4.3 Иностранное население в Испании по данным реестра населения за 1998-2010 гг. (в млн человек)



Источник: Реестр населения (Padrón). Национальный институт статистики (INE-Spain)

160. Аналогичный процесс был налажен для граждан ЕС, но он был создан позднее, в 2009 году, и в полной мере применялся только с 2016 года.

161. В то время, когда необходимо было принимать решения о проведении переписи населения, примерно в 2008-2009 годах, существовала значительная неопределенность в отношении реального числа иностранных граждан, проживающих в Испании. Сколько из почти 6 миллионов иностранцев, зарегистрированных с 1996 года, покинули страну, не сообщив об этом государственным органам? Сколько испанцев покинули страну, не заявив о новом месте жительства за границей в консульских учреждениях?

162. В конце концов, в 2009 году была разработана методология проведения переписи 2011 года. Он был задуман как сочетание реестров (связывающих Padrón с другими) и адресованного 10 % населения исследования, которое в основном использовалось для сбора всех демографических переменных, не включенных в реестры.

163. Основной задачей при разработке методологии переписи было определение способа объединения реестров для более точного измерения численности населения, как испанцев, так и иностранцев, учитывая, что в реестре населения было мало данных об оттоке населения, особенно иностранного. В случае испанцев, по крайней мере, есть источник информации об оттоке населения, а именно постановка на учет в консульских учреждениях.

4.1.7.4 Методы, используемые для интеграции данных

164. Перепись населения 2011 года включала в себя сложную процедуру интеграции данных для повышения точности подсчета численности населения. Эта процедура использовалась для подсчета так называемого «сомнительного» населения, т. е. людей, зарегистрированных в системе Padrón, но не имеющих подтверждения проживания в Испании в других административных реестрах (базах данных налоговых органов и социального обеспечения, среди прочих).

165. Во время переписи населения 2011 года для того, чтобы подсчитать численность населения, данные по зарегистрированному населению (как испанцам, так и иностранным жителям) были сопоставлены с множеством различных административных реестров, чтобы определить вероятность проживания каждого человека, т.е. количественный параметр, называемый «коэффициентом подсчета». В результате этого процесса был создан файл «предварительной переписи». Основной принцип прост. Если человек правильно зарегистрирован и получает заработную плату или пенсию, либо платит налоги в течение 2011 года, то, скорее всего, этот человек проживает в Испании. Люди по этим проверенным записям называются «достоверным» населением. А те, которых нет в других записях, считаются «сомнительным» населением.

166. Результатом этого процесса стал файл, содержащий данные, собранные до проведения переписи, с примерно 47,4 миллионами записей.

167. Было определено около 15 различных критериев, по которым запись может быть признана сомнительной. После этой связи результаты были следующими:

- 97,7% записей в файле, составленном до переписи (46 372 000), были признаны «достоверными» данными о численности населения (каждому лицу в этой группе был присвоен коэффициент подсчета, равный 1);
- Около 40 000 записей (0,1 % населения) были удалены из файла, подготовленного до переписи, после вероятностного сопоставления записей со свидетельствами о смерти. Они были признаны ошибками в реестре населения и им был присвоен коэффициент подсчета, равный 0;
- В отношении остальных 2,2% записей доказательства были неубедительными, поэтому они были признаны сомнительными. Таких записей было 1 046 000, и 87 % из них касались иностранцев. Поэтому их «коэффициент подсчета» был неизвестен.

168. Первым выводом этого процесса было то, что сомнительная группа населения состояла в основном из мигрантов, но не из экономических мигрантов. В течение десятилетия, предшествовавшего переписи 2011 года, Испания принимала большие потоки мигрантов из Эквадора, Марокко и Румынии, но также наблюдался значительный и непрерывный приток немцев и британцев, которые приезжали в Испанию, чтобы жить там после выхода на пенсию. Этих «мигрантов в поисках хорошей погоды», которых часто называют «любителями солнца», сложнее найти в реестрах: они могут быть зарегистрированы в системе Padrón, но не получать пенсию от испанского правительства и не платить налоги в Испании.

169. «Коэффициент подсчета» для сомнительных записей был рассчитан с помощью «проведенного в рамках переписи исследования» (как уже упоминалось, перепись 2011 года в Испании включала в себя адресованное 10 % населения исследование).

170. Для этой цели сомнительные записи были сгруппированы по классам населения, определенным по социально-демографическим характеристикам: возрастные группы, гражданство и провинция проживания (в Испании 52 провинции). Файл, подготовленный до переписи, был разбит на эти классы. Каждый класс содержал достоверные и сомнительные записи. Если класс не содержал сомнительных записей, то оценка не требовалась, и численность населения для этого класса была равна численности из файла, составленного до переписи.

171. Все классы, содержащие сомнительные записи, были настроены так, чтобы включать не менее 1 000 таких записей. Классы с менее чем 1 000 сомнительных записей были сгруппированы таким образом, чтобы достичь этого порога. Стратегия группирования заключалась в том, чтобы сначала выбрать более широкие возрастные группы, затем более широкие территориальные области и, наконец, сгруппировать по государственной принадлежности. Было создано 724 кластера или класса,

в каждом из которых было не менее 1 000 сомнительных записей.

Таблица 4.2 Сомнительные записи в файле предварительной переписи 2011 года. Избранные страны гражданства

Гражданство	Общая численность населения в файле, подготовленном до переписи, с CF* ≠ 0	Сомнительные записи	% сомнительных записей
Итого	47 418 916	1 046 433	2,2
Испания	41 589 484	136 452	0,3
Румыния	886 631	120 216	13,6
Великобритания	394 998	112 100	28,4
Марокко	806 012	61 459	7,6
Германия	195 667	58 690	30,0
Боливия	197 113	37 492	19,0
Парагвай	90 788	32 107	35,4
Болгария	173 061	31 269	18,1
Бразилия	105 216	28 352	26,9
Франция	120 780	27 482	22,8

*CF = коэффициент подсчета

172. Для расчета этих коэффициентов подсчета нам необходимо было оценить долю достоверной популяции в исследовании. В данном классе i файла предварительной переписи общее количество записей T_i было разделено на количество записей о достоверном S_i и сомнительном населении D_i ,

$$T_i = S_i + D_i$$

173. Для уточнения оценки общей численности населения и с целью создания новой оценки общей численности населения для этого класса к сомнительным записям в классе i был применен коэффициент подсчета T , как показано ниже:

$$T_i^* = S_i + CF_i * D_i$$

174. Население, участвовавшее в исследовании, было сгруппировано в те же 724 класса в соответствии с теми же характеристиками (возраст, территория, гражданство). После этого выборка и записи в файле предварительной переписи были сопоставлены на индивидуальном уровне, чтобы определить, какие записи из выборки были ранее классифицированы как сомнительные в файле предварительной переписи.

175. С учетом этих коэффициентов в каждой сомнительной записи файл, подготовленный до переписи, превратился во взвешенный файл переписи (WCF) и позволил получить данные о численности населения по переписи. Население данной географической области, T_i^* , составляло сумму коэффициентов подсчета записей во взвешенном файле переписи (WCF) в этой области (при условии, что коэффициент подсчета (CF) был равен 1 для всех достоверных записей).

176. Эта процедура обеспечила статистические данные для переписи, но, конечно, она не может быть использована в административных целях для обновления реестра населения, поскольку невозможно определить, каких отдельных граждан следует учитывать, а каких нет.

177. Наиболее примечательной группой с точки зрения доли сомнительного населения была группа немецких граждан в возрасте 80 лет и старше в провинции Аликанте (место на побережье Средиземного моря), насчитывающая почти 3 000 человек, 70% из которых были признаны сомнительными. Наибольшую группу по абсолютному числу сомнительных лиц составляют граждане Великобритании, также проживающие в провинции Аликанте, в возрасте от 60 до 65 лет. Она включает в себя 5 963 сомнительных записей из 19 317 граждан Великобритании, включенных в эту группу в файле, составленном до переписи.

178. Некоторые избранные случаи приведены в таблице 4.3.

4.1.7.5 Преимущества интеграции со статистическими данными из других источников

179. Как указано выше, количество сомнительных записей составило 1 046 000 в округленных цифрах. Средний коэффициент подсчета для всех классов составил 0,424, что означает, что эти сомнительные записи были подсчитаны как население примерно в 440 000 человек.

180. Коэффициент подсчета (CF) был выше 1 только в 22 классах, что означает, что эти люди считались «недозарегистрированными». Это касалось граждан Пакистана в возрасте от 25 до 29 лет с коэффициентом подсчета 1,37, что затронуло 1 055 сомнительных записей (впоследствии подсчитанных как население в 1 445 человек).

Таблица 4.3 «Сомнительная» группа населения в файле предварительной переписи по гражданству, провинции и возрасту

Группа (гражданство, провинция, возраст)	Население в файле, подготовленном до переписи (T_i)	Достоверные записи (S_i)	Сомнительные записи (D_i)	% сомнительных записей	Коэффициент подсчета (CF)	Расчетная численность населения $T_i^* = S_i + (D_i * CF)$
Великобритания, Аликанте, 60-64 лет	19 317	13 354	5 963	30,9%	0,347	15 426,1
Великобритания, Аликанте, 65-69 лет	21 180	15 287	5 893	27,8%	0,222	16 598,1
Германия, Аликанте, 70-74 лет	6 703	2 633	4 070	60,7%	0,269	3 729,1
Великобритания, Аликанте, 55-59 лет	11 732	7 663	4 069	34,7%	0,254	8 695,3
Румыния, Мадрид, 25-29 лет	32 909	29 867	3 042	9,2%	0,257	30 649,8
Пакистан, Испания, 25-29 лет	4 384	3 329	1 055	24,1%	1,370	4 774,5
Марокко, Мадрид, 0-9 лет	17 162	16 158	1 004	5,9%	0,115	17 046,6

181. В результате этих расчетов население Испании на 1 ноября 2011 года составило 46 815 916 человек, что примерно на 450 000 человек меньше, чем показатель, указанный в реестре населения на ту же дату. В то время как число испанцев практически не изменилось, согласно переписи, численность иностранцев сократилась на 480 000 человек по сравнению с данными реестра населения и составила 5 252 473 человека (11,2 % населения Испании, что на 8 % меньше, чем данные реестра населения).

182. Наиболее заметным случаем сокращения численности населения стало сокращение числа граждан Великобритании. Согласно реестру, в стране проживало 391 194 гражданина Великобритании, однако по данным переписи их число составляло всего 312 098 человек.

183. В заключение следует отметить, что использование методов интеграции данных, сочетающих административные данные и данные исследований, позволило улучшить показатели численности населения, главным образом в отношении иностранных граждан. Кроме того, с 2011 года качество реестра населения стало значительно лучше. В городских администрациях были введены процедуры для проверки фактического места жительства иностранного населения, особенно иностранцев-граждан ЕС, а также для улучшения данных о численности населения, предоставляемых реестром населения.

4.1.7.6 Список использованной литературы

http://www.ine.es/en/censos2011_datos/cen11_datos_metodologia_en.htm

Argüeso, A. and Vega, Jorge L. (2013) "A population census based on registers and a "10% survey" methodological challenges and conclusions" («Перепись населения на основе реестров и «10-процентное исследование»: методологические проблемы и выводы»). Statistical Journal of the IAOS – Vol. 30

4.1.8 Швейцария

184. Швейцарское федеральное статистическое управление (FSO) в настоящее время не применяет никаких методов интеграции данных в строгом смысле этого слова для составления статистики миграции. Единственным источником таких статистических данных являются реестры населения.

185. Подготовка ежегодных данных о миграции является неотъемлемой частью статистики населения и домохозяйств (STATPOP). Она основана на одном из нескольких различных исследований, проведенных в рамках федеральной переписи населения. Статистические данные содержат информацию о численности и составе населения на конец определенного года, а также об изменении численности населения и его демографических компонентах в течение того же года.

186. В 2010 году традиционная десятилетняя перепись населения была заменена интегрированной статистической системой, которая предоставляет ежегодные данные. Новая система переписи сочетает использование административных реестров с выборочными исследованиями. Она состоит из четырех различных ежегодных исследований, среди которых особое значение для статистики миграции имеет анализ самого реестра.

187. Именно он является источником данных для статистики населения и домохозяйств (STATPOP). Она использует существующие и гармонизированные административные данные из централизованных и местных реестров населения и поэтому особенно хорошо подходит для получения данных о миграционных потоках. Реестры зависят от своевременной регистрации населения и отслеживают все изменения места жительства, обеспечивая таким образом постоянное обновление записей по каждому человеку. Административные источники часто способны обеспечить это, несмотря на некоторые свои хорошо известные недостатки (например, сбор статистических данных не является приоритетом, определения и охват зависят от законодательства и административных правил).

188. Национальный номер социального страхования, который однозначно идентифицирует человека (универсальный PIN-код), играет важную роль как в ведении административных реестров, так и в их статистическом использовании, особенно при сопоставлении информации из реестров и данных исследований.

4.1.8.1 Основные источники данных для измерения миграционных потоков и миграционной численности: реестры населения

189. Основными источниками данных для официальной ежегодной статистики миграции (демографические запасы и потоки) являются административные реестры на федеральном, кантональном и муниципальном уровнях.

190. Существует три различных типа реестров:

- a) 2 300 местных реестров (ведутся муниципалитетами или, в некоторых случаях, кантонами¹⁷) собственных и иностранных граждан;
- b) Центральная миграционная информационная система (реестр иностранцев, который ведет Государственный секретариат по миграции на федеральном уровне) для законно проживающих иностранцев (т. е. лиц, имеющих официальный вид на жительство);
- c) Информационная система Ordipro (реестр иностранных граждан, ведомый Федеральным

¹⁷ В Швейцарии нет центрального реестра населения, охватывающего все население страны (собственных граждан и иностранцев). Кроме того, Швейцарское федеральное статистическое управление (FSO) не ведет статистический реестр населения. Оно получает выдержки данных, содержащие заранее определенные переменные в заранее определенном стандартизированном формате данных, из каждого реестра через регулярные промежутки времени.

министерством иностранных дел) для иностранных граждан, которые имеют право на привилегии и иммунитеты и не подпадают под действие швейцарского иммиграционного законодательства, т. е. сотрудники дипломатических миссий, консульских учреждений, постоянных представительств и межправительственных организаций (включая супругов, партнеров, не состоящих в браке детей, а также частных домашних работников).

191. В то время как граждане Швейцарии регистрируются только в местном реестре, иностранные граждане могут фигурировать в нескольких реестрах (в одном или нескольких местных реестрах населения, а также в одном из двух федеральных реестров). Поскольку не существует единого источника данных, который давал бы полное представление о миграции иностранных граждан, рекомендуется использовать несколько источников. В одних случаях эти источники приводят к одинаковым выводам, а в других случаях подробности и тенденции могут различаться. Однако это не обязательно означает, что один источник «правильный», а другой «неправильный» или что один источник «лучше» другого.

192. Поэтому для всех категорий иностранных граждан (например, постоянных жителей, просителей убежища и т.д.) необходимо выбрать источник, на котором будет основываться количество мигрантов (миграционная численность) или миграционные потоки, которые будут учитываться в официальной статистике. Своевременность, полнота, надежность и другие факторы имеют решающее значение для определения приоритетности источников.

193. Для подготовки ежегодных данных о миграционных потоках и миграционной численности были приняты следующие решения:

- Реестром для «постоянно проживающего иностранного населения» (долгосрочные мигранты, т. е. лица, проживающие в Швейцарии не менее 12 месяцев, а также лица, прибывающие в Швейцарию и покидающие ее, которым выданы разрешения на легальное проживание или пребывание со сроком действия не менее 12 месяцев) являются местные реестры населения;
- Реестром для «временных иностранных граждан» (краткосрочных мигрантов, т. е. лиц, проживающих в Швейцарии менее 12 месяцев, и лиц, прибывающих в Швейцарию и покидающих ее с законными разрешения, действительные менее 12 месяцев) является Центральная миграционная информационная система;
- Реестром для иностранных граждан, которым Федеральное министерство иностранных дел выдал удостоверения личности, является информационная система Ordiprog.

194. Следовательно, «интеграция данных для измерения миграции» в Швейцарии означает сознательный выбор одного административного реестра вместо другого для определенных подгрупп населения, не являющихся гражданами страны.

195. Местные и федеральные источники данных не являются взаимоисключающими. У них частично пересекаются единицы наблюдения, а также частично пересекаются и переменные параметры. Это пересечение объясняется тем, что в местные реестры населения включаются все проживающие иностранные граждане, которые к тому же включаются и в один из двух федеральных реестров. Однако не предпринимается никаких попыток систематически сравнивать реестры с целью выявления лиц (их численности) и миграционных событий (потоков), которые зарегистрированы в нескольких реестрах или только в одном реестре (и поэтому отсутствуют в других). Число мигрантов и миграционных потоков не увеличивается за счет данных, содержащихся в административной базе данных, которая не считается «базовым реестром» для конкретной подгруппы населения, и данные «базового реестра» не игнорируются и не удаляются, если они не содержатся в другом реестре.

196. Кроме того, характеристики, относящиеся к отдельному лицу, могут различаться в разных реестрах. В зависимости от задачи или положения реестра в административном процессе, со статистической точки зрения можно предположить, что качество некоторых характеристик или переменных (таких как пол, возраст, семейное положение, гражданство, страна последнего/предыдущего проживания и т. д.) в некоторых реестрах лучше или более целостно, чем в других. Поэтому может возникнуть необходимость в приоритизации источников для определенных характеристик. Для составления ежегодной статистики миграции снова используется «справочный реестр», который определяет, какой вид миграции будет учитываться.

197. Однако Центральная миграционная информационная система предоставляет некоторые эксклюзивные переменные, которых нет на местном уровне. Они добавляются в записи местного реестра населения, относящиеся к тем же наблюдениям (отдельные лица или миграционные события). Эти переменные включают «гражданство супруга/партнера», «цель пребывания» или «мотив миграции», «дату первого выдачи разрешения», «дату истечения срока действия вида на жительство» и, для лиц, ищущих убежища, подробную информацию о различных административных этапах процедуры предоставления убежища.

198. Эта процедура «улучшения» качества данных была внедрена в процесс подготовки ежегодных данных о миграции и может рассматриваться как «скромная» форма интеграции данных. Данные, собранные на федеральном уровне, сопоставляются с данными, собранными на уровне кантонов или муниципалитетов. Наличие номера социального страхования в каждом из соответствующих реестров позволяет однозначно связывать записи об одном и том же лице в разных источниках.

199. В области миграционной статистики Швейцарское федеральное статистическое управление (FSO) проводит четкое разграничение между статистическими данными и оценками: первые в целом представляют собой результат (насколько это возможно) исчерпывающего сбора записей из первичных источников данных, вторые — результат применения вероятностных/статистических моделей, которые могут включать объединение информации из различных источников. Официальная статистика Швейцарии по миграции не содержит никаких оценок — ни на индивидуальном (на уровне отдельного реестра), ни на агрегированном уровне.

4.1.8.2 Другие источники данных для измерения миграционной численности и миграционных потоков, не включенные в составление ежегодной миграционной статистики

200. В Швейцарии существуют дополнительные источники данных о миграции, два наиболее важных из которых — это Швейцарское исследование трудовых ресурсов (LFS) и Структурное исследование (ежегодное выборочное исследование 200 000 человек, которое является частью новой системы переписи и в ходе которого собираются переменные, в настоящее время недоступные в реестрах, например, родной язык, вероисповедание, уровень образования).

201. Ключевым аспектом новой швейцарской системы переписи населения является возможность сопоставления данных из реестров населения с данными ежегодного выборочного исследования с помощью номера социального страхования. Таким образом, информационное содержание обоих источников может быть объединено, чтобы дополнить друг друга. Однако в настоящее время в рамках ежегодной миграционной статистики такое сопоставление или объединение данных не проводится.

202. Информация, относящаяся к миграции, полученная в результате выборочных опросов, представлена отдельно и дает дополнительные сведения о численности и потоках мигрантов. Кроме того, данные исследований и данные реестров различаются не только по географическому охвату и методам сбора данных, но и по составу населения. Как Швейцарское исследование трудовых ресурсов (LFS), так и структурное исследование охватывают только лиц в возрасте 15 лет и старше, проживающих в частных домохозяйствах.

4.2 Тематические исследования без реестров населения

4.2.1 Австралия

4.2.1.1 Источники статистических данных о миграции

203. Административная информация о лицах, прибывающих в Австралию или выезжающих из нее, собирается с помощью различных систем обработки данных, паспортных документов, визовой информации, а также карточек прибывающих и выходящих пассажиров (см. Приложение 1 к справочному материалу ниже). Прибывающие лица предоставляют информацию в своих визовых анкетах, за исключением тех, кто путешествует как граждане Австралии или Новой Зеландии. Эти

административные данные собираются Департаментом иммиграции и пограничной охраны Австралии (DIBP) в соответствии с положениями Миграционного регламента (Закон о миграции, 1958 г.).

204. Статистика Австралийского бюро статистики (ABS) по въезду и выезду иностранных граждан (OAD) составляется на основе информации из источников Департамента иммиграции и пограничной охраны Австралии (DIBP). Данные о перемещениях за границей собираются и сопоставляются (где это возможно) также Департаментом иммиграции и пограничной охраны Австралии (DIBP), а затем хранятся вместе с записями о перемещениях в Системе обработки данных о поездках и иммиграции (TRIPS). Каждый месяц сопоставленные записи по въезду и выезду иностранных граждан (OAD) передаются в Австралийское бюро статистики (ABS), а затем обрабатываются. Уникальный персональный идентификатор — это механизм, используемый для связывания всех наборов данных в Департаменте иммиграции и пограничной охраны Австралии (DIBP), а затем используемый Австралийским бюро статистики для создания истории поездок каждого отдельного путешественника.

205. Ежеквартальные оценки чистой миграции за рубеж (NOM) получают из этих обработанных ежемесячных данных по въезду и выезду иностранных граждан (OAD), а затем объединяют с ежемесячными выписками из несопоставленных записей по въезду и выезду иностранных граждан (OAD). Несопоставимые записи по въезду и выезду иностранных граждан (OAD) — это записи, в которых Департамент иммиграции и пограничной охраны Австралии (DIBP) зарегистрировал въезд или выезд в Системе обработки данных о поездках и иммиграции (TRIPS), но не удалось сопоставить данные с соответствующей пассажирской карточкой. Статистика по миграции за границу не включает в себя: многократные въезды/выезды собственных граждан и иностранцев; въезды/выезды оперативного персонала авиации и судов; транзитных пассажиров, которые проходят через Австралию, но не получают разрешения на въезд; пассажиров круизных лайнеров, которые начинают и заканчивают свой рейс в Австралии, а также прибывших без документов. С 1 июля 2006 года исключаются также иностранные дипломатические сотрудники и члены их семей.

206. Справочную информацию о методах, используемых Австралией для измерения международной миграции, см. на сайте <http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/3412.0Explanatory%20Notes12016-17?OpenDocument>

4.2.1.2 Источники для измерения миграционной численности

207. Как упоминалось выше, для международных мигрантов источником информации о стране рождения являются данные, предоставленные Департаментом иммиграции и пограничной охраны Австралии (DIBP), которые являются комбинацией информации из паспортов и виз. В случае смерти источником информации о стране рождения являются данные, предоставленные органами регистрации актов рождения и смерти в каждом штате и каждой территории Австралии. В Австралии все рождения обязательно регистрируются как акты рождения. Кроме того, раз в пять лет проводится перепись населения Австралии, которая используется для пересчета численности населения Австралии, в том числе с учетом страны рождения.

4.2.1.3 Использование интеграции данных с другими источниками

208. Одной из наиболее полезных переменных стал уникальный персональный идентификатор (PID) для каждого человека, пересекающего границу Австралии, будь то въезд или выезд. Это позволяет использовать данные для многих дополнительных целей, в том числе для установления связи между отдельными лицами в динамике и в рамках различных систем.

209. Австралийское бюро статистики в сотрудничестве с Департаментом иммиграции и пограничной охраны Австралии (DIBP) связывает наборы данных о мигрантах с данными переписи населения. Кроме того, в недавних работах эти же данные были сопоставлены с налоговой информацией — см. соответствующий документ ниже.

210. См. следующие документы, начиная с самых последних:

Migrant Data Matrices, 2017 (Матрицы данных о мигрантах, 2017)¹⁸

Understanding Migrant Outcomes - Insights from the Australian Census and Migrants Integrated Dataset, Australia, 2016 (Понимание результатов миграции — выводы на основе данных австралийской переписи

¹⁸ <http://www.abs.gov.au/AUSSTATS/abs@.nsf/productsbyCatalogue/70AA0E84BE9D586ACA2575400017B0F3?OpenDocument>

населения и интегрированного набора данных о мигрантах, Австралия, 2016)¹⁹

Microdata: Australian Census and Migrants Integrated Dataset, 2016 (Микроданные: австралийская перепись населения и интегрированный набор данных о мигрантах, 2016)²⁰

Perspectives on Migrants, Mar 2013 (Взгляд на мигрантов, март 2013 г.)²¹

См. также Guide to all the other Migrant Statistical Sources (Руководство по всем другим источникам статистических данных о мигрантах)²²

4.2.2 Канада

4.2.2.1 Первоначальный/основной источник данных для статистики миграции

4.2.2.1.1 Для целей оценки численности населения

211. Иммиграционная служба Канады (IRCC) собирает и обрабатывает административные файлы иммигрантов и временных жителей (NPR). Затем она предоставляет эту информацию Статистическому управлению Канады (STC).

212. В сфере иммиграции эти файлы используются для оценки количества и характеристик лиц, которым федеральное правительство предоставило статус постоянного жительства на определенный период.

213. Для временных жителей (NPR) файлы, необходимые для подготовки их оценок, включают разрешения на въезд, разрешения на работу, разрешения на обучение, разрешения министров, заявления о предоставлении статуса беженца, разрешения на высадку, заявления о высадке и депортации. Эта информация используется для оценки количества и характеристик лиц, которым федеральное правительство предоставило статус временного жительства.

214. При оценке эмиграции необходимо проводить различие между лицами, которые приобретают постоянное место жительства в другой стране (т. е. эмигрантами), лицами, которые временно проживают за границей, не сохраняя обычного места жительства в Канаде; другими лицами, которые временно покидают Канаду, а затем возвращаются (т. е. чистыми временными эмигрантами), и, наконец, частью эмигрантов, которые вернулись в Канаду (т. е. возвращающимися эмигрантами).

215. Число эмигрантов оценивается на основе административных источников с учетом валового оттока мигрантов из Канады. Управление иммиграционной статистики Министерства внутренней безопасности США предоставляет данные о канадцах, получивших статус постоянных иммигрантов в США. Этот источник данных используется для оценки эмиграции в Соединенные Штаты. Для оценки эмиграции в другие страны используется информация об уведомлениях об уезде из программы по ежемесячной налоговой выплате семьям с детьми до 18 лет (Canada Child Tax Benefit, CCTB) и налоговые данные Канадского агентства по доходам (CRA). В данные вносятся различные корректировки, например, для исправления неполного охвата.

216. Число временных эмигрантов оценивается по нескольким источникам. Данные из исследования охвата переписью «Обратная проверка записей» (Reverse Record Check), предназначенного для измерения недоохвата, используются для оценки числа лиц, временно покидающих страну; в то время как данные из Национального исследования домохозяйств (NHS) в сочетании с оценками Управления демографии (Demography Division) по возвращающимся эмигрантам используются для оценки числа временно эмигрировавших лиц, возвращающихся в страну.

217. Возвращающиеся эмигранты — это граждане Канады или иммигранты, которые ранее эмигрировали из Канады, а затем вернулись в Канаду, чтобы вновь установить постоянное место жительства. Опять же, для оценки числа возвращающихся эмигрантов используются данные программы по ежемесячной налоговой выплате семьям с детьми до 18 лет (Canada Child Tax Benefit, CCTB) и налоговые данные.

¹⁹ <http://www.abs.gov.au/AUSSTATS/abs@.nsf/productsbyCatalogue/706907E56F9F5128CA257BEA00111584?OpenDocument>

²⁰ <http://www.abs.gov.au/AUSSTATS/abs@.nsf/productsbyCatalogue/C916B16440BF9B60CA257C7E000FC851?OpenDocument>

²¹ <http://www.abs.gov.au/AUSSTATS/abs@.nsf/productsbyCatalogue/3B2787D4377D2D84CA2573F7000DDE5B?OpenDocument>

²² [http://www.abs.gov.au/AUSSTATS/abs@.nsf/Latestproducts/3414.0Main%20Features42011%20\(Edition%202\)?opendocument&tabname=Summary&prodno=3414.0&issue=2011%20\(Edition%202\)&num=&view=](http://www.abs.gov.au/AUSSTATS/abs@.nsf/Latestproducts/3414.0Main%20Features42011%20(Edition%202)?opendocument&tabname=Summary&prodno=3414.0&issue=2011%20(Edition%202)&num=&view=)

4.2.2.1.2 Для аналитических характеристик

218. Данные о постоянных и временных жителях, предоставленные Иммиграционной службой Канады (IRCC), содержат информацию, необходимую для аналитических исследований и разработки политики. Этот источник данных представляет собой перепись иммигрантского населения и, как таковой, дает возможность детально изучить небольшие группы населения. Источник данных о постоянных жителях содержит социально-демографическую информацию об иммигрантах, прибывших в страну с 1980 года. Получение характеристик, связанных со знанием языков, местом назначения в Канаде, страной происхождения, предполагаемой профессией или уровнем образования, происходит в момент въезда в страну

219. В отношении временных жителей (NPR) файлы, созданные до их въезда в страну, содержат подробную географическую информацию и сведения об основной деятельности человека, такой как работа или учеба, в определенные периоды после 1980 года.

220. Перепись населения дает статистическую картину страны каждые пять лет. Перепись 2016 года определяет постоянных и временных жителей, проживающих в Канаде. Оно содержит много социально-экономических характеристик иммигрантов, таких как состав семьи, знание языков, статус трудовой деятельности и доход.

4.2.2.2 Ограничения оригинального или основного источника. Почему интеграция необходима?

4.2.2.2.1 Для целей оценки численности населения

221. Данные об иммиграции являются простыми административными данными. Однако в отношении оценок по провинциям и территориям в файле, полученном от Иммиграционной службы Канады (IRCC), указана провинция или территория, в которую иммигрант планирует прибыть, а не провинция или территория, в которой он фактически поселится. В небольшом числе случаев отсутствует информация о провинции назначения. Эти случаи распределяются пропорционально между провинциями и территориями в соответствии с наблюдаемым распределением иммигрантов, по которым имеется такая информация.

222. В отношении данных о временных жителях (NPR) существует такое же ограничение, касающееся предполагаемой провинции поселения, как и в отношении данных об иммиграции. К тому же отсутствует информация о датах выезда временных жителей (NPR), поскольку некоторые из них покидают страну до истечения срока действия своих разрешений. Поскольку информирование федерального правительства перед выездом из страны не является обязательным, Статистическое управление Канады (STC) не имеет дополнительных источников для получения этой информации.

223. Кроме того, иммигрантам и временным жителям (NPR) присваивается идентификационный номер клиента (CID) IRCC. Этот идентификационный номер является ключевой переменной, поскольку позволяет идентифицировать каждого человека только один раз и навсегда. Идентификационный номер клиента (CID) используется при составлении оценок численности населения, поскольку позволяет Статистическому управлению Канады (STC) создать биографию каждого человека, даже если он имеет более одного вида на жительство в один и тот же период времени.

224. В отличие от иммиграционного учета, в Канаде нет законодательных положений, обязывающих вести учет лиц, покидающих страну на временной или постоянной основе. Поэтому оценки числа эмигрантов и лиц, временно проживающих за границей, а также их характеристик необходимо получать из вторичных источников, таких как канадские административные файлы или иммиграционная статистика США. Кроме того, лишь немногие источники данных позволяют оценить масштабы эмиграции, и эти источники часто используют разные определения эмиграции. Временную эмиграцию особенно трудно измерить с учетом ее характера. В результате различные оценки показывают, что Канада недооценивает уровни эмиграции и временной эмиграции. В этом контексте интеграция необходима для максимально эффективного использования преимуществ различных наборов данных. Например, поскольку США являются страной назначения большинства эмигрантов из Канады, использование данных об иммиграции в США, как считается, улучшает качество наших оценок числа канадцев, переезжающих в США.

4.2.2.2.2 Для аналитических характеристик

225. Административные данные Иммиграционной службы Канады (IRCC) не содержат информации о каких-либо результатах после въезда в страну. Информация, касающаяся социально-экономических результатов, изменений в структуре семьи и изменений в месте жительства, может быть получена только путем сопоставления с внешними источниками данных, такими как ежегодные налоговые декларации.

226. Перепись населения действительно предоставляет актуальные данные о иммигрантах после их прибытия, но не содержит никакой информации об условиях въезда в страну или критериях отбора.

4.2.2.3 Методы, используемые для интеграции данных

4.2.2.3.1 Для целей оценки численности населения

227. В отношении иммиграции и временных жителей интеграция информации осуществляется на уровне отдельных записей. Все файлы содержат одинаковые переменные и связаны между собой с помощью уникального идентификационного номера клиента (CID). Измерение количества иммигрантов, прибывших в Канаду за определенный период, несложно, и корректировка данных не

требуется. Информация о каждом лице, въезжающем в Канаду в качестве иммигранта, доступна в административном файле Иммиграционной службы Канады (IRCC). Каждый месяц Иммиграционная служба Канады (IRCC) предоставляет Статистическому управлению Канады (STC) файл данных, содержащий записи об иммигрантах за предыдущий месяц, а также любые дополнения или обновления к уже полученным данным.

228. В отношении временных жителей существуют две основные подгруппы, которые различаются с административной точки зрения (лица, имеющие разрешение на пребывание (РН), и лица, ходатайствующие о предоставлении статуса беженца (RSC)); поэтому оценки их численности должны представляться отдельно. Для получения более подробной информации о методологии оценки количества временных жителей, пожалуйста, см. <http://www.staging.statcan.gc.ca/pub/91-528-x/2015001/ch/ch5-eng.htm#n1>

229. Что касается эмиграции (эмиграция, чистая временная эмиграция и возвращающиеся эмигранты), то см. методологический документ: <http://www.staging.statcan.gc.ca/pub/91-528-x/2015001/ch/ch6-eng.htm>

4.2.2.3.2 Для аналитических характеристик

230. Для того чтобы обойти ограничения административных иммиграционных данных, Статистическое управление Канады использует методы точного сопоставления записей для объединения информации с другими источниками данных, такими как налоговые файлы. Продольная база данных по иммиграции (IMDB) обновляется ежегодно и включает всех иммигрантов, прибывших в Канаду с 1980 года, с результатами из ежегодных налоговых деклараций с 1982 года. За 2014 год записи 89% иммигрантов были связаны с налоговыми файлами. Таким образом, база данных представляет собой административную перепись с продольным «дизайном», и поэтому выборка не проводится.

231. Кроме того, данные Национального исследования домохозяйств 2011 года были сопоставлены с административными данными по иммиграции, чтобы связать условия для въезда в страну с социально-экономическими результатами, доступными в исследовании. Для переписи населения 2016 года было использовано сопоставление записей для добавления в базу данных переписи переменных, связанных с категорией разрешения на въезд в страну.

4.2.2.4 Преимущества интеграции со статистическими данными из других источников

4.2.2.4.1 Для целей оценки численности населения

232. Основное преимущество интеграции различных источников данных заключается в предоставлении более качественных оценок, которые будут полезны не только для Статистического управления Канады (STC), но и для многих внешних пользователей. И наконец, Статистическое управление Канады (STC) в настоящее время работает над созданием статистического реестра населения. Тогда можно было бы связать данные об иммиграции и эмиграции с помощью уникального личного идентификатора.

4.2.2.4.2 Для аналитических характеристик

233. Связь между данными переписи и административными файлами по иммиграции обеспечивает дополнительные аналитические возможности и одновременно снижает нагрузку на респондентов. Этот подход является хорошей альтернативой для получения подробной информации о категории иммиграционного отбора, которую респонденты могут игнорировать, упускать из виду или со временем забывать.

234. Будучи в основном транзакционными по своему характеру, административные файлы Иммиграционной службы Канады (IRCC) по иммиграции имеют некоторые ограничения в отношении показателей эффективности и исследований. Продольная база данных по иммиграции (IMDB) создается на основе сопоставления данных иммиграционных и налоговых файлов с целью обеспечения более полного источника данных о социально-экономических результатах деятельности иммигрантов, подающих налоговые декларации. Этот источник данных содержит подробную информацию о поведении на рынке труда и охватывает достаточно длительный период, чтобы оценить влияние таких характеристик выдачи разрешения на въезд, как образование или знание французского или английского языка. Такие элементы, как опыт иммигрантов до приезда (т. е. опыт

обучения в Канаде), мобильность внутри Канады или изменения в составе семьи, используют преимущества продольного аспекта, который становится возможным с каждым новым налоговым отчетом.

4.2.3 Новая Зеландия

235. Являясь основным ведомством, ответственным за обработку данных о международных перевозках в Новую Зеландию и из нее, Статистическое управление Новой Зеландии предоставляет своевременные и точные статистические данные по трем типам пассажиров: иностранные пребывающие лица, путешественники, проживающие в Новой Зеландии, а также постоянные жители и долгосрочные мигранты. Благодаря выгодному географическому положению, когда въезд в страну осуществлялся только через морские порты и аэропорты, основным способом сбора данных являлись заполняемые самими путешественниками карточки пассажира.

236. Со временем содержание карточек изменилось, и в долгосрочной перспективе ожидается дальнейшее сокращение объема информации, которую пассажиры предоставляют в бумажных формах. В качестве источника для проверки также широко используются интегрированные электронные файлы международного пассажиропотока. Статистическое управление Новой Зеландии изучает возможность использования альтернативных источников и статистических методов для получения информации об отправлении с целью уменьшения или устранения зависимости от заполняемых при выезде карточек отправления.

4.2.3.1 Сбор данных

4.2.3.1.1 Источники для измерения миграционных потоков

237. Публикации статистических данных о международной миграции в Новой Зеландии основаны на электронном учете всех прибытий и отправок пассажиров, предоставляемых Статистическому управлению Новой Зеландии Таможенной службой Новой Зеландии. Записи включают в себя данные о поездке, такие как информация о перевозчике, вид транспорта и порт, а также основные данные паспорта, такие как дата рождения, пол и страна гражданства. Записи включают в себя данные об обработке пассажиров, которым требуется виза для въезда в Новую Зеландию. Типы виз классифицируются в соответствии с политикой и требованиями Иммиграционной службы Новой Зеландии. Все данные о перемещениях через границу и пассажирах, собираемые Таможенной службой Новой Зеландии, регулярно (ежедневно) передаются в Статистическое управление Новой Зеландии.

238. Таможенная служба Новой Зеландии также предоставляет Статистическому управлению Новой Зеландии физические карточки пассажиров, которые были заполнены путешественниками во время поездки. Статистическая информация, собираемая с помощью карточек пассажиров, включает в себя дополнительные сведения о путешественнике, такие как страна последнего или следующего места жительства, цель поездки и предполагаемая продолжительность пребывания в Новой Зеландии или отсутствия в Новой Зеландии. Она также позволяет классифицировать категории пассажиров (долгосрочные мигранты, иностранные пребывающие лица гости, постоянные жители).

239. Статистическое управление Новой Зеландии осуществляет сканирование карточек пассажиров с использованием технологии распознавания изображений для автоматического кодирования текстовых ответов. Затем следует процесс ручного заполнения записей, которые не были четко распознаны и закодированы программным обеспечением для обработки изображений (примерно одной из десяти карточек). Некоторые поля ответов являются обязательными для заполнения в каждой карточке пассажира с целью сопоставления с электронными записями, предоставленными Таможенной службой Новой Зеландии. Другие поля карточки пассажира заполняются только для целей выборки и варьируются в зависимости от типа пассажира и направления движения.

4.2.3.1.2 Источники для измерения миграционной численности

240. Пятилетняя перепись населения позволяет оценить численность населения, родившегося за рубежом, путем сбора данных о стране рождения и годе прибытия в Новую Зеландию для населения, обычно проживающего в стране. По данным переписи 2013 года, численность населения, родившегося за рубежом, составляла около 1 миллиона человек, что представляет почти четверть от

общей численности обычно постоянного проживающего в Новой Зеландии населения (рисунок 4.4). В 1986 году доля населения Новой Зеландии, родившегося за рубежом, составляла менее 15 процентов, и до 2006 года наблюдался заметный рост численности и доли населения, проживающего в стране и родившегося за рубежом (рис. 4.4 и рис. 4.5).

241. Проводимое после переписи населения исследование (PES) для оценки полноты охвата переписью. Оценки недоучета и переучета производятся в рамках проводимого после переписи населения исследования (PES) для групп населения, определенных по возрасту, полу, этнической принадлежности и географическому региону. Пятилетние исторические данные проводимого после переписи населения исследования (PES) (1996-2013) не содержат оценок чистого недоучета населения, родившегося за границей, в переписи (Статистическое управление Новой Зеландии, 2014). Однако для проводимого после переписи населения исследования (PES) 2018 года Статистическое управление Новой Зеландии планирует включить вопрос о стране рождения. Это признание необходимости повышения точности и расширения статистики населения, родившегося за рубежом.

242. В таблице 4.4 обобщены данные, содержащиеся в сборниках миграционной информации, объединенной для составления ежемесячной статистики международного пассажиропотока.

243. Пассажир может изменить свои планы относительно продолжительности пребывания или отсутствия после прибытия или отправления, в результате чего тип пассажира, зарегистрированный на момент поездки, может оказаться неверным. Статистическое управление Новой Зеландии не пересматривает опубликованные статистические данные с целью учета таких изменений. Существует необходимость в альтернативном методе классификации пассажиров после наблюдения за их фактическими вариантами пребывания в Новой Зеландии. Временная серия связанных между собой последовательностей поездок для индивидуальных путешественников позволяет окончательно классифицировать их статус мигранта.

4.2.3.1.3 Источники для измерения миграционной численности

244. Пятилетняя перепись населения позволяет оценить численность населения, родившегося за рубежом, путем сбора данных о стране рождения и годе прибытия в Новую Зеландию для населения, обычно проживающего в стране. По данным переписи 2013 года, численность населения, родившегося за рубежом, составляла около 1 миллиона человек, что представляет почти четверть от общей численности обычно постоянного проживающего в Новой Зеландии населения (рисунок 4.4). В 1986 году доля населения Новой Зеландии, родившегося за рубежом, составляла менее 15 процентов, и до 2006 года наблюдался заметный рост численности и доли населения, проживающего в стране и родившегося за рубежом (рис. 4.4 и рис. 4.5).

245. Проводимое после переписи населения исследование (PES) для оценки полноты охвата переписью. Оценки недоучета и переучета производятся в рамках проводимого после переписи населения исследования (PES) для групп населения, определенных по возрасту, полу, этнической принадлежности и географическому региону. Пятилетние исторические данные проводимого после переписи населения исследования (PES) (1996-2013) не содержат оценок чистого недоучета населения, родившегося за границей, в переписи (Статистическое управление Новой Зеландии, 2014). Однако для проводимого после переписи населения исследования (PES) 2018 года Статистическое управление Новой Зеландии планирует включить вопрос о стране рождения. Это признание необходимости повышения точности и расширения статистики населения, родившегося за рубежом.

Таблица 4.4 Сбор и обработка данных о международной миграции

Сбор данных		
Записи о перемещениях	Записи о пассажирах	Отсканированные карточки пассажиров
Фактическая дата	Таможенный порт	Дата поездки
Запланированная дата	Номер паспорта	Номер паспорта Название перевозчика/судна
Название основного перевозчика/судна	Гражданство	Направление
Код совместного использования (код-шеринг)	Дата рождения	Дата рождения
Направление	Имена (текст)	Предполагаемая продолжительность пребывания/отсутствия
Порт Новой Зеландии	Фамилии (текст)	Цель поездки
Вид транспорта	Пол	Страна основного назначения
	Код типа визы	Страна последнего/следующего постоянного места жительства
	Дата визы	Зарубежное государство
	Страна рождения	Сфера деятельности
	Код порта маршрута	Адрес в Новой Зеландии (город/район)
	Код совместного использования (код-шеринг)	
Обработка для целей статистики		
<p>Загрузка данных о перемещениях и пассажирах</p> <p>Валидация</p> <p>Редактирование/корректировка</p> <p>Связывание данных со сканами карточек пассажиров</p> <p>Кодирование</p> <p>Связь с историей классификации пассажиров в прошлом</p> <p>Производные переменные:</p> <ul style="list-style-type: none"> – Возраст – Продолжительность пребывания/отсутствия (кол-во дней) – Тип пассажира (иностранцы, проживающие в Новой Зеландии, а также постоянные жители и долгосрочные мигранты) <p>Условное исчисление</p> <p>Выборка и взвешивание данных</p> <p>Подготовка ежемесячной статистики пассажиропотока</p>		
Серия связанных между собой историй путешествий для окончательного определения статуса мигранта		
<p>Идентификатор пассажира</p> <p>Идентификатор, который можно связать с другими записями о перемещении и пассажирах (Идентификатор, который можно связать с интегрированной инфраструктурой данных Статистического управления Новой Зеландии)</p> <p>Направление передвижения</p> <p>Дата и время передвижения</p> <p>Дата рождения</p> <p>Статус долгосрочного мигранта, полученный по правилу «12/16 месяцев»</p>		

Рисунок 4.4 Число постоянных жителей Новой Зеландии, родившихся за рубежом

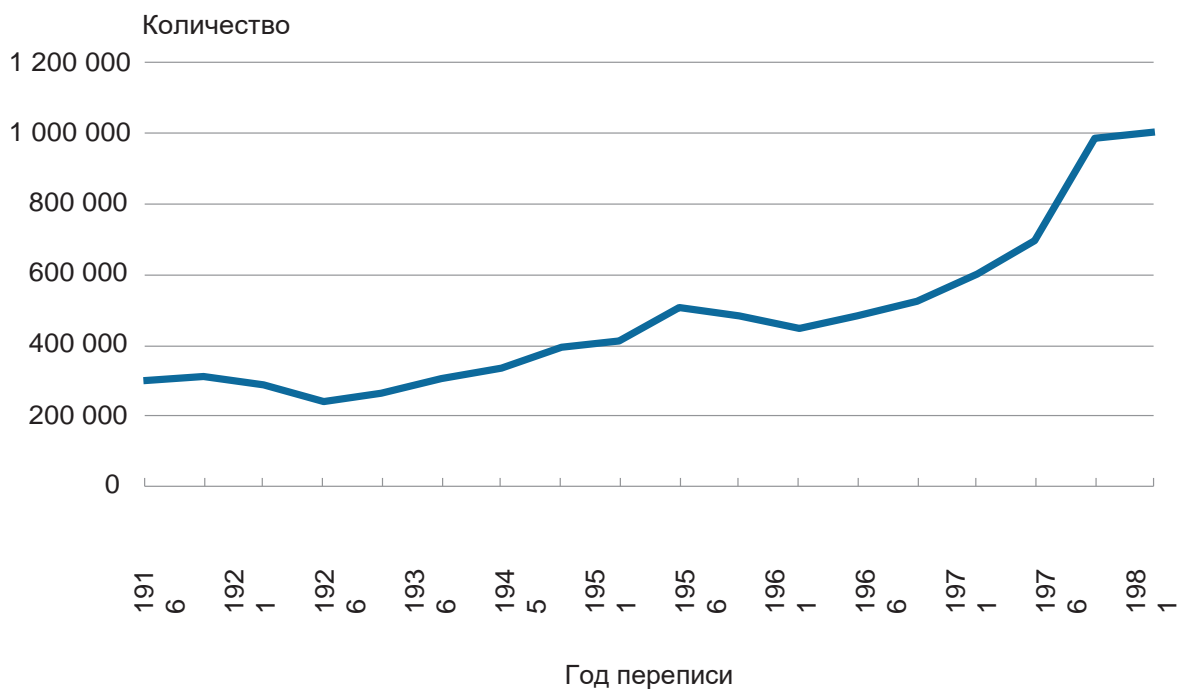
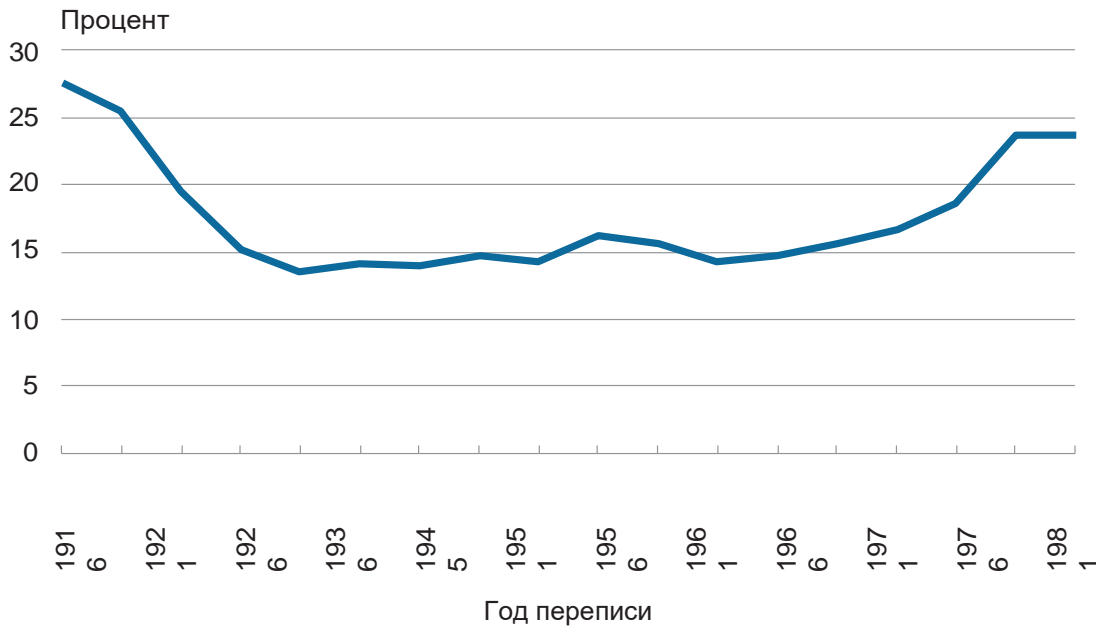


Рисунок 4.5 Процент населения Новой Зеландии, родившегося за рубежом



4.2.3.2 Цель интеграции данных

4.2.3.2.1 Миграционные потоки

246. Связывание записей с изображениями карточек пассажиров с электронными записями о пересечении границы и пассажирах осуществляется в первую очередь для включения дополнительной информации о паспортах и перевозчиках, доступной в электронных записях о пересечении границы. Интеграция данных о пересечении границы с данными о пассажирах также служит средством проверки фактических пассажиров, пересекающих границу в данный момент времени. Для некоторых пассажиров не будет заполненных карточек, но будут доступны электронные записи, или могут быть заполненные карточки, но не будет электронных записей. Для оптимизации охвата международного пассажиропотока предусмотрен процесс ручного связывания данных.

247. Для путешественников, которые заявили о своем намерении остаться в Новой Зеландии или покинуть Новую Зеландию на длительный срок, создаются связанные последовательности классификации поездок и пассажиров за 16 месяцев, предшествующих месяцу заявления о намерении. Цель состоит в том, чтобы подтвердить статус жителя Новой Зеландии на момент поездки, используя информацию на карте о том, проживают ли они в Новой Зеландии 12 месяцев или более. С учетом интегрированных источников данных и истории классификации пассажира оценка служит проверкой качества данных при первоначальном подсчете прибытий и отправлений на долгий срок.

248. Пассажир может изменить свое решение о пребывании или отъезде после прибытия в Новую Зеландию или после выезда из Новой Зеландии. Дополнительная мера по подтверждению статуса мигранта учитывает, когда международный пассажир включается в состав постоянного населения или исключается из него после учета времени пребывания в Новой Зеландии или отсутствия в Новой Зеландии. Эта мера использует связанные истории поездок за период наблюдения в 16 месяцев и применяет операционные правила для определения изменения статуса жителя. «Правило 12/16 месяцев» не зависит от статуса законного проживания лица и от информации, указанной в карточке пассажира.

249. Наблюдение за последовательностью поездок отдельных лиц в течение определенного периода времени и анализ фактической продолжительности пребывания или отсутствия могут дать результаты, значительно отличающиеся от информации, первоначально записанной в карточках пассажиров. Человек может перейти из категории иностранных туристов в категорию жителей Новой Зеландии при следующем выезде за границу, пройдя этапы получения временной визы на территории страны и постоянного вида на жительство. Переезды в пределах страны, приводящие к изменению статуса жителя, могут привести к завышению или занижению первоначальных оценок числа прибывших и выбывших мигрантов.

250. Использование данных об истории поездок также открывает дополнительные возможности для расширения статистики международной миграции. Например, в будущем может возникнуть интерес к показателям краткосрочной миграции. Кроме того, история классификации пассажиров представляет собой источник продольных данных, позволяющий проводить статистические измерения возвратной и маятниковой миграции. Интеграция с интегрированной инфраструктурой данных Статистического управления Новой Зеландии (Statistics New Zealand, 2016) также облегчает проведение других анализов миграционных траекторий и результатов.

4.2.3.3 Методология интеграции данных

251. Файлы с данными о перемещениях через границу и пассажирах могут быть автоматически связаны через систему хранения записей, предоставляемую Таможенной службой Новой Зеландии. Детерминированная интеграция этих электронных файлов с изображениями карточек пассажиров включает в себя автоматический шаг, который связывает номер паспорта, имя и дату рождения, название перевозчика/судна и направление поездки. По крайней мере два поля (номер паспорта, имя или дата рождения) должны полностью совпадать, чтобы запись считалась совпадающей. Добавлен ручной шаг для исправления несвязанных записей путем просмотра изображений имен и неверных ответов на связывающие переменные.

252. Небольшое количество электронных записей о перемещениях пассажиров не имеет

соответствующих заполненных карточек пассажиров. Вместо отсутствующих ответов на карточке используются данные о прошлых поездках и характеристиках пассажиров. С другой стороны, для некоторых пассажиров карточка пассажира может быть единственным источником информации о пересечении границы. Они в основном касаются пассажиров, прибывающих или убывающих на корабле, и для этих пассажиров создаются соответствующие электронные записи.

253. Объединение этих записей с последовательностями классификации пассажиров за последние 16 месяцев осуществляется путем детерминированной привязки по следующим полым: номер паспорта, имя или дата рождения. Опять же, записи считаются совпадающими, если имеется полное совпадение по двум из трех связывающих переменных.

254. Исключение избыточных данных о перемещении через границу для создания серии уникальных идентификаторов пассажиров является вероятностным процессом, в котором используется программное обеспечение IBM QualityStage. В нем применяется методология Fellegi-Sunter (Statistics NZ, 2015) с использованием ключевых идентификационных данных пассажиров, например, имени, даты рождения, номера паспорта. Оно присваивает уникальный идентификационный номер пассажира, что позволяет пассажирам менять гражданство в паспорте, использовать более одного паспорта или женщинам менять фамилию. Для получения более подробной информации о методологии интеграции данных и применении правила «12/16 месяцев» см. ссылку (Statistics NZ, 2017).

4.2.3.4 Использование интеграции данных с другими источниками

255. Исторические наборы данных о перемещениях через границу и классификации пассажиров, собранные с помощью описанных выше процессов, хранятся в интегрированной инфраструктуре данных (IDI) Статистического управления Новой Зеландии (Statistics New Zealand, 2016). Данные о поездках и миграции в интегрированной инфраструктуре данных (IDI) также включают в себя наборы данных, содержащие информацию о решениях и критериях утверждения иммиграции, а также о категориях временных виз (например, студенческих, рабочих), выдаваемых при въезде в Новую Зеландию. База исследований интегрированной инфраструктуры данных (IDI) содержит интегрированные данные на уровне отдельных лиц, собранные от ряда государственных учреждений, статистических исследований Новой Зеландии, включая перепись 2013 года, и неправительственных организаций.

256. Интеграция данных о поездках и миграции с другими источниками, такими как налоговые данные, позволяет проводить продольные исследования влияния мигрантов на рынок труда. В частности, результаты рынка труда могут быть сопоставлены между конкретными группами мигрантов и для различных географических регионов.

4.2.3.5 Список использованной литературы

Statistics New Zealand, International Travel and Migration, Data Collection Methodology (Статистическое управление Новой Зеландии, Международные поездки и миграция, Методология сбора данных), <http://datainfolplus.stats.govt.nz/Item/nz.govt.stats/f705ca38-ea6e-453f-b1d9-a95dc0caf59#/nz.govt.stats/d83affa2-c0e1-4629-9aff-cc8452eb41c6#>

Statistics New Zealand, Defining migrants using travel histories and the '12/16-month' rule (Статистическое управление Новой Зеландии, Определение мигрантов с использованием истории поездок и правила «12/16 месяцев») http://www.stats.govt.nz/browse_for_stats/population/Migration/define-migrants-travel-rule.aspx

Statistics New Zealand, Integrated Data Infrastructure (Статистическое управление Новой Зеландии, Интегрированная инфраструктура данных) http://www.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure.aspx (2016)

Statistics New Zealand, Data integration manual (Статистическое управление Новой Зеландии, Руководство по интеграции данных) <http://www.stats.govt.nz/methods/data-integration/data-integration-manual-2edn.aspx> (2015)

Statistics New Zealand, Coverage in the 2013 Census based on the New Zealand 2013 post-enumeration Survey (Статистическое управление Новой Зеландии, Охват переписи 2013 года на основе

проводимого после переписи населения исследования) http://www.stats.govt.nz/browse_for_stats/population/census_counts/report-on-2013-post-enumeration-survey.aspx (2014)

4.2.4 Великобритания

257. Управление национальной статистики (ONS) составляет статистику по международной миграции в Великобританию и из Великобритании, которая помогает нам понять состав общества и изменения в структуре населения. Наши основные статистические данные по миграции:

- Оценки долгосрочной международной миграции (потоки);
- Оценки краткосрочной международной миграции (потоки);
- Население Соединенного Королевства по стране рождения и гражданству (численность населения).

258. В данном тематическом исследовании основное внимание уделяется оценкам Управления национальной статистики (ONS) по долгосрочной международной миграции (LTIM).

4.2.4.1 Источники данных для измерения статистики миграции

259. Данные Управления национальной статистики (ONS) о потоках долгосрочной международной миграции (LTIM) основаны на результатах Международного опроса пассажиров (IPS) с несколькими корректировками, внесенными на основе других источников данных исследований и административных данных.

260. Международный опрос пассажиров (IPS) — это многоцелевое исследование, в рамках которого собирается информация от пассажиров при въезде в Великобританию или выезде из нее. Он в основном используется для предоставления данных о международной миграции, расходах на поездки и туризме.

261. Международный опрос пассажиров (IPS) — это выборочное исследование, которое проводится во всех основных пунктах въезда и выезда из Великобритании, включая аэропорты, морские порты и тоннель под Ла-Маншем.

262. Для сбора данных в определенный день проводится выборка по конкретному пункту въезда или выезда. Используется строгий метод подсчета, который гарантирует, что каждый человек имеет равные возможности для участия в исследовании. Пассажиры систематически отбираются для опроса через определенные промежутки (например, каждый 20-й человек) с произвольным началом.

263. Ежегодно в рамках Международного опроса пассажиров (IPS) охватывается от 700 000 до 800 000 человек. Из числа опрошенных ежегодно около 4 000 человек идентифицируются как долгосрочные мигранты. Это лица без права постоянного или временного жительства в Великобритании, которые заявляют о своем намерении остаться в Великобритании не менее чем на 12 месяцев (иммигранты), или жители Великобритании, которые заявляют о своем намерении проживать за пределами Великобритании не менее чем 12 месяцев (эмигранты). Остальные лица идентифицируются как краткосрочные мигранты, лица, прибывающие в Великобританию и выезжающие из нее, или жители Великобритании, возвращающиеся после краткосрочного пребывания за границей или выезжающие за границу на короткий срок.

264. Дополнительную информацию о Международном опросе пассажиров (IPS) можно найти в разделе «Информация о качестве международного опроса пассажиров в связи с миграционными потоками».

4.2.4.2 Цель интеграции данных

265. Хотя Международный опрос пассажиров (IPS) является основным источником данных о долгосрочной миграции, не существует единой всеобъемлющей системы для измерения всех перемещений мигрантов в Великобританию и из нее. Международный опрос пассажиров (IPS) имеет некоторые ограничения в отношении измерения иммиграции и эмиграции, поскольку:

- Является выборочным исследованием, поэтому у оценок определенная степень

- неопределенности;
- Охватывает очень небольшое количество лиц, ищущих убежища, которые могут въезжать в Великобританию или выезжать из нее, а также лиц, высылаемых из Великобритании без предоставления убежища;
 - Не учитывает изменение намерений пассажиров; речь идет о пассажирах, которые намеревались остаться в Великобритании или покинуть ее в течение 12 месяцев, но фактически провели там менее года (мигрант, чей статус долгосрочного пребывания не был реализован), а также о тех, кто полагал, что останется или уедет менее чем на год, но фактически пробыл дольше (пребывающие лица, чей статус долгосрочного пребывания не был реализован);
 - Не охватывает лиц, пересекающих сухопутную границу между Великобританией (Северная Ирландия) и Республикой Ирландия.

266. Поэтому для получения оценок потоков долгосрочной международной миграции (LTIM) необходимо использовать комбинацию данных из разных источников, которые имеют разные характеристики и атрибуты.

4.2.4.3 Методология интеграции данных

267. Описанная здесь методология была впервые применена в 2009 году для расчета оценок за 2008 год, а данные за предыдущие годы были пересмотрены в соответствующих случаях.

268. Оценки потоков долгосрочной международной миграции (LTIM) на 90% основаны на данных Международного опроса пассажиров (IPS). Более полная оценка долгосрочной международной миграции получается путем объединения данных Международного опроса пассажиров (IPS) с информацией, полученной из дополнительных источников. Кроме того, вносятся корректировки для мигрантов, которые меняют свои намерения (известные как «пребывающие лица, чей статус долгосрочного пребывания не был реализован», и «мигранты, чей статус долгосрочного пребывания не был реализован»). Эта более полная оценка называется оценкой потоков долгосрочной международной миграции (LTIM) (см. рисунок 4.6).

Рисунок 4.6 Расчет долгосрочной международной миграции



Источник: Управление национальной статистики

269. Ни один из использованных источников данных, включая Международный опрос пассажиров (IPS), не предназначен специально для сбора информации только лишь о долгосрочной миграции. Оценки потоков долгосрочной международной миграции (LTIM) производятся с использованием следующих основных источников данных:

- Международный опрос пассажиров (IPS);
- Исследование (LFS): предоставляет данные о географическом распределении мигрантов для калибровки Международного опроса пассажиров (IPS);
- Административные системы иммиграционной службы Министерства внутренних дел: предоставляют данные о лицах, ищущих убежища, и их иждивенцах с 1991 года; с 2013 года также используются данные о выдворении лиц, не ищущих убежища; с 2015 года в них внесены поправки в отношении лиц, переселившихся в Великобританию в рамках различных программ переселения;
- Прогнозируемые долгосрочные оценки международной миграции, основанные на предыдущих регистрациях у врачей общей практики (местных врачей) из Агентства статистики и исследований Северной Ирландии (NISRA) для оценки долгосрочной международной миграции в Северную Ирландию и из нее, а также в другие страны мира, начиная с 2008 года (прогнозируемые данные заменяются окончательными данными для окончательных годовых оценок потоков долгосрочной международной миграции (LTIM)). Это охватывает наземные маршруты, которые не исследуются в рамках Международного опроса пассажиров (IPS).

270. Опубликованные данные потоков долгосрочной международной миграции (LTIM) разбиты по

Руководство по интеграции данных для измерения

таким переменным, как гражданство, возраст и пол. Для получения оценок по каждой из этих переменных данные из источников, которые вносят вклад в оценку потоков долгосрочной международной миграции (LTIM), также необходимо разбить по тем же переменным.

271. Данные об мигрантах из Международного опроса пассажиров (IPS) доступны в разбивке по каждой переменной. Данные о потоках в Северной Ирландии и данные о лицах, ищущих убежища, отсутствуют и должны быть получены на основе ряда допущений. Кроме того, данные Международного опроса пассажиров (IPS), используемые для корректирующего расчета числа пребывающих лиц, чей статус долгосрочного пребывания не был реализован, основаны на относительно небольшом размере выборки за каждый год, но все равно должны быть разбиты таким же образом.

272. Поскольку данные, полученные из дополнительных источников, не являются столь подробными, как данные, собранные в рамках Международного опроса пассажиров (IPS), уровень анализа, который можно провести на основе данных о потоках долгосрочной международной миграции (LTIM), является ограниченным. Таблицы оценок потоков долгосрочной международной миграции (LTIM) обычно содержат только одну переменную, но при этом отдельно доступны перекрестные таблицы данных Международного опроса пассажиров (IPS).

273. Оценки потоков долгосрочной международной миграции (LTIM) приводятся с доверительными интервалами для компонента оценки из Международного опроса пассажиров (IPS), чтобы дать представление о точности оценки. Неопределенность, связанная с компонентом оценки из Международного опроса пассажиров (IPS), используется для расчета статистически значимых изменений в оценках потоков долгосрочной международной миграции (LTIM). Однако при интерпретации этих доверительных интервалов и статистически значимых изменений читатель должен иметь в виду, что не существует метода количественной оценки погрешности, связанной с не относящимися к опросу компонентами оценки потоков долгосрочной международной миграции (LTIM), которые вряд ли являются случайными.

274. Дополнительная информация о методологии, включая допущения, использованные для расчета оценок Международного опроса пассажиров (IPS) и потоков долгосрочной международной миграции (LTIM)²³, а также изменения в методологии с течением времени, доступна на веб-сайте Управления национальной статистики (ONS). Сюда также входят копии анкет, использованных в ходе Международного опроса пассажиров (IPS).

4.2.4.4 Использование интеграции данных с другими источниками

275. Существует множество источников официальной статистики, которые измеряют количество и характеристики международных мигрантов, прибывающих в Великобританию и выезжающих из нее (потоки), а также мигрантов, которые поселились в Великобритании (их численность). В совокупности они дают полное представление о миграции в Великобритании. Важно понимать, что эти источники измеряют разные вещи: одни измеряют потоки, другие — численность, третьи — работающих лиц, четвертые — студентов, а пятые — только характеристики мигрантов из-за пределов ЕС. Каждый источник ценен сам по себе для оценки конкретных аспектов международной миграции.

²³ Methodology to estimate LTIM" («Методология оценки потоков долгосрочной международной миграции» (LTIM)), <http://www.abs.gov.au/AUSSTATS/abs@.nsf/productsbyCatalogue/208F6F1B4DED24ECCA257EB50011C711?Open-Document>

276. Как уже отмечалось несмотря на то, что все источники данных, включая Международный опрос пассажиров (IPS), предоставляют наиболее актуальную информацию, ни один из них не предназначен специально для сбора данных исключительно о потоках долгосрочной международной миграции (LTIM), и поэтому ни один из них в отдельности не может количественно оценить все притоки и оттоки населения. Интеграция данных имеет решающее значение для обеспечения максимально точной оценки международной миграции, которой могут доверять и которую могут с уверенностью использовать как правительство Великобритании, так и более широкий круг пользователей.

277. Управление национальной статистики (ONS) стремится предоставлять достоверную информацию для информирования общественности и обсуждения политических вопросов. Международная миграция является предметом активных дискуссий, и существует высокий спрос на достоверные данные и аналитические материалы. В систему регулярно вносятся методологические изменения, чтобы обеспечить максимально полные и надежные оценки, особенно в тех случаях, когда появляются новые источники данных. Подробная информация об изменениях приведена в Методологии оценки потоков долгосрочной международной миграции (LTIM)²⁴. Кроме того, в феврале 2017 года Управление национальной статистики (ONS) опубликовало записку о планах по развитию статистики международной миграции с особым акцентом на более эффективное использование существующих административных источников данных (см. «Данные и анализ по международной миграции: Улучшение доказательной базы»²⁵).

4.2.5 Соединенные Штаты Америки

4.2.5.1 Подготовка оценок чистой международной миграции и ее составляющих

278. Бюро переписи населения США ежегодно составляет оценки чистых международных миграционных потоков (NIM) в и из США, которые используются в качестве исходных данных для оценки численности населения на национальном и субнациональном уровнях. Интеграция данных о миграции ограничена, хотя некоторые иностранные и административные данные используются для расчета нескольких подкомпонентов международных миграционных потоков (NIM). В то же время Бюро переписи населения предприняло значительные усилия по более эффективному использованию существующих административных данных с целью сокращения затрат, связанных со сбором данных переписи, хотя этот подход еще не был применен в отношении данных о международной миграции. Были предложены некоторые будущие работы в этой области в отношении данных о миграции, но они все еще находятся на начальной стадии.

279. Интеграция данных происходит только на аддитивном уровне, т.е. оценки подкомпонентов международных миграционных потоков (NIM) получают из разных источников данных, а затем объединяют для получения окончательной оценки. Оценка международных миграционных потоков (NIM), подготовленная Бюро переписи населения США, состоит из нескольких подкомпонентов, которые рассчитываются отдельно: 1) иммиграция иностранцев, 2) эмиграция иностранцев, 3) чистая миграция в/из Пуэрто-Рико, 4) чистая международная миграция коренных жителей и 5) чистая миграция военнослужащих. Компоненты, касающиеся лиц, родившихся за границей и в Пуэрто-Рико, используют данные Американского социологического исследования (ACS) и его эквивалента в Пуэрто-Рико. Эти подкомпоненты составляют основную часть оценок международных миграционных потоков (NIM), но мы также добавляем постоянный чистый подкомпонент, который получается в результате анализа данных о лицах, родившихся в США или являющихся гражданами США, полученных из иностранных источников (преимущественно переписей населения). Последним подкомпонентом является чистая миграция военнослужащих в и из США, которая рассчитывается на основе административных данных, предоставляемых Министерством обороны.

4.2.5.2 Перемещение военнослужащих в зарубежные страны и из них

²⁴ <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/methodologies/internationalmigrationmethodology>

²⁵

<https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/articles/internationalmigrationdataandanalysisimprovingtheevidence/february2017>

280. Компонент чистой миграции военнослужащих представляет собой оценку чистого перемещения военнослужащих за границу на основе данных, собранных Министерством обороны, Центром данных по личному составу вооруженных сил (DMDC). Центр данных по личному составу вооруженных сил (DMDC) предоставляет ежемесячные сводные таблицы военнослужащих, дислоцированных или развернутых за пределами США, с разбивкой по возрасту, полу, расе, испаноязычному происхождению и отдельным видам вооруженных сил в рамках Министерства обороны. Мы предполагаем, что изменения в численности военнослужащих за рубежом, за исключением смертей, указывают на перемещение военнослужащих в и из США. Для получения оценок чистой международной миграции военнослужащих сил на уровне штатов и округов мы в первую очередь используем информацию от Центра данных по личному составу вооруженных сил (DMDC) с разбивкой по возрасту, полу, расе, испаноязычному происхождению и округу. Для улучшения географического распределения перемещений военнослужащих вокруг определенных внутренних военных объектов мы используем информацию о группировке округов, полученную из последнего пятилетнего отчета в рамках Американского социологического исследования (ACS).

281. Эта цифра, как на национальном, так и на субнациональном уровне, просто добавляется к другим подкомпонентам международных миграционных потоков (NIM) для получения окончательной оценки международных миграционных потоков (NIM). Это единственный тип интеграции данных о международной миграции, который мы проводим в настоящее время.

4.2.5.3 Интеграция данных на микроуровне

282. В настоящее время Бюро переписи населения не занимается интеграцией данных о международной миграции на микроуровне, хотя мы провели эксперимент по анализу данных о стране рождения из файла Администрации социального обеспечения (NUMIDENT) и их объединению с данными о налогах Налогового управления США (IRS) с целью анализа моделей внутренней миграции лиц, родившихся за рубежом. Мы также обсудили возможность сопоставления административных данных (от Государственного департамента) о беженцах с данными переписи населения или Американского социологического исследования (ACS), чтобы помочь определить статус беженца с помощью вероятностных методов. Кроме того, проводятся исследования по использованию комбинации данных Налогового управления США (IRS) (текущее место жительства), Администрации социального обеспечения (NUMIDENT) (страна рождения) и данных переписи (раса/испаноязычное происхождение) для улучшения распределения оценок международных миграционных потоков (NIM) на субнациональном уровне в масштабах округов. Другие административные данные микроуровня хранятся в Министерстве внутренней безопасности и его Управлении по иммиграционной статистике (OIS). Последнее предприняло собственные усилия по объединению различных наборов административных данных федеральных агентств, но в Бюро переписи населения США мы только начали размышлять о возможных способах интеграции данных о миграции на этих уровнях.

4.3 Резюме

283. В тематических исследованиях отражен широкий спектр ситуаций, которые существенно отличаются друг от друга с точки зрения имеющихся источников данных для статистики миграции, ограничений этих источников и применяемых подходов к интеграции данных. Примеры, которые приведены в этих тематических исследованиях, не нужно считать вполне репрезентативной выборкой подходов к интеграции данных, разработанных в различных странах. Кроме того, в процессе оценки опыта этого спектра стран следует учитывать, что национальные условия могут значительно различаться в плане количества важных аспектов и что какой-либо подход к интеграции данных, успешно примененный в одной стране, не обязательно обеспечит оптимальное решение другим странам. Несмотря на эти ограничения, анализ представленных тематических исследований позволяет выделить ряд основных аспектов, которые могут быть полезны для других стран.

284. Что касается **стран, имеющих регистр населения**, то в некоторых из них (в том числе в Австрии, Испании, Италии и Латвии) **этот регистр является главным источником данных о миграции**. В этих странах интеграция данных проводится прежде всего для повышения качества данных. Данные из других регистров могут использоваться, в частности, для корректировок на отсутствующие показатели регистрации эмигрантов или для выявления умерших. В ряде стран интеграция данных используется также для обеспечения дополнительной информации по переменным, не включенным в первичные

источники.

285. В некоторых странах применяется практика, основанная на **«сигналах присутствия»** (в ряде стран они называются «признаками жизни») лиц в разного рода регистрах. Например, в **Австрии** анализ сигналов присутствия в других регистрах проводится для расчета оценок числа лиц, которые фигурируют в регистре населения, но в стране больше не проживают. Данные из Управления социального обеспечения Австрии используется для выявления умерших, а из ряда других административных источников – для корректировок на отсутствующие показатели deregистрации эмигрантов.

286. В **Италии**, где в регистр населения могут вноситься только те лица, которые легально проживают в стране, анализ сигналов присутствия в различных источниках позволяет уточнять оценочные данные по международной миграции посредством выявления лиц, которые включены в регистр населения, но, вероятно, эмигрировали, а также лиц, которые иммигрировали, но не фигурируют в регистрах населения.

287. В **Латвии** анализ сигналов присутствия проводится с использованием различных регистров, включая регистры доходов, социального страхования, образования, здравоохранения и занятости. Оценочные показатели эмиграции рассчитываются также с помощью зеркальной статистики, поступающей из стран назначения. Кроме того, качество статистики миграции оценивается с использованием результатов обследований домохозяйств, включая ОРС и ОДУЖ-ЕС.

288. В **Испании** национальный регистр населения («padrón») находится в ведении национального статистического института INE, и в него должны заноситься все проживающие в стране лица независимо от их юридического статуса. При проведении в 2011 году переписи населения зарегистрированное население было увязано с многими другими административными регистрами с целью выявления лиц, чье присутствие в стране вызывает сомнения, а также для выявления умерших. Что касается сомнительных записей, то благодаря методам интеграции данных на основе административных данных и выборочному обследованию, проведенному в связи этой переписью, удалось оценить число лиц, которые фигурировали в регистре, фактически не проживая в стране.

289. В некоторых странах (включая Венгрию, Израиль, Нидерланды и Швейцарию) статистика миграции производится с применением **систем, где регистр населения является одним из нескольких источников данных**, причем не обязательно самым важным из них. В Венгрии используется довольно сложная система регистров, в которой различные источники дополняют друг друга. Показатели иммиграции и эмиграции рассчитываются путем интеграции данных из этого разнообразия источников, в частности из Управления по вопросам иммиграции и убежища и Центрального регистра населения, находящихся в ведении Министерства внутренних дел. Интеграция данных дает возможность вырабатывать качественную статистику миграции и получать новую информацию в дополнение к той, которая поступила из первичных источников.

290. В **Израиле** данные о миграционных потоках и численности населения поступают главным образом из системы пограничной регистрации и передаются Управлением по вопросам народонаселения и иммиграции (которое также ведет регистр населения) в распоряжение Центрального статистического бюро. Эти данные объединяются с другими административными записями (в том числе по секторам налогообложения, образования и т. д.) и с результатами переписи населения и других статистических обследований, чтобы на их основе рассчитывались более качественные оценочные показатели и добавлялись переменные из других источников, связанные с теми же лицами или событиями.

291. В **Нидерландах** статистика миграции (как и вся социально-демографическая статистика) основана на системе баз данных социальной статистики, являющейся результатом интеграции данных из более чем 50 административных регистров. Считается, что зарегистрировано – лишь за несколькими исключениями – практически все проживающее в стране население. Что касается эмиграции, то около трети выезжающих из страны лиц не регистрируются. Когда органы власти обнаруживают отсутствие какого-либо лица, они проводят расследование, и, если разыскать это лицо не удастся, оно может быть зарегистрировано как эмигрировавшее в неизвестную страну.

292. В **Швейцарии** к данным по миграции имеют отношение три основных регистра: 1) регистры местного населения, ведущиеся муниципальными властями; 2) федеральный регистр законно проживающих иностранцев, имеющих вид на жительство; и 3) особый федеральный регистр иностранных дипломатов, сотрудников консульств и международных организаций, членов их семей и аналогичных категорий лиц. В стране проводится несложная процедура интеграции данных, состоящая в сопоставлении данных федерального уровня с данными, собранными на местном уровне. Кроме того, некоторые переменные, имеющиеся только в регистре законно проживающих иностранцев, добавляются к записям из регистра местного населения, имеющим отношение к тем же наблюдениям.

293. Что касается **стран, не имеющих регистра населения**, то в них к основным источникам данных о миграции относятся пассажирские формуляры, заполняемые на границе, паспортные данные, визы и виды на жительство. В странах, где ведутся регистры населения, интеграция данных производится с использованием различных источников в целях повышения качества данных и получения более полной информации. В частности, интеграция данных может способствовать уточнению сведений по иммигрантам в части их фактического места и даты размещения, которые могут отличаться от заявленных в момент прибытия.

294. В **Австралии** данные о въезде и выезде компилируются Австралийским статистическим бюро на основе информации, поступающей из различных систем обработки данных, включая пассажирские формуляры, паспортные данные и визовую информацию. Данные увязываются с помощью персонального идентификатора, который позволяет АСБ создавать личные учетные записи для проезжающих. Наборы данных по мигрантам также увязываются с данными переписей и с налоговой информацией.

295. В **Канаде** административные данные по иммигрантам и непостоянным жителям поступают в статистическое управление этой страны. Поскольку сведения о провинции, планируемой для проживания, и дате прибытия в нее не всегда соответствуют фактическому месту и дате размещения, производится интеграция данных с использованием других источников, таких как налоговые записи. Наряду с этим обеспечивается увязка записей с использованием данных из других источников (таких, как перепись населения Канады, Национальное обследование домохозяйств, Управление Соединенных Штатов по статистике миграции), позволяющая рассчитывать качественные оценочные показатели и получать более полную информацию по социально-экономическим характеристикам мигрантов.

296. В **Новой Зеландии** графические пассажирские формуляры считываются пограничными электронными системами учета передвижений и регистрации пассажиров, обеспечивая поступление дополнительной информации и проверку данных. Кроме того, в случае проезжающих, заявивших о намерении остаться в стране или выехать из нее на длительный период, путевые и личные пассажирские данные увязываются на срок 16 месяцев до контрольного месяца, чтобы на момент совершения поездки подтвердить статус соответствующего лица как жителя Новой Зеландии.

297. В **Соединенном Королевстве** имеется целый ряд источников официальной статистики, предоставляющих информацию по различным аспектам потоков международной миграции и численности мигрантов. Применительно к миграционным потокам главным источником является обследование международных пассажиропотоков (ОМП) – многоцелевое выборочное обследование, в ходе которого собирается информация от пассажиров, въезжающих в страну или выезжающих из нее. Поскольку при проведении ОМП не учитываются изменения в намерениях пассажиров, для расчета долгосрочных оценок международной миграции проводится интеграция данных с использованием информации, полученной в результате ОМП и ОРС, административных данных по иммиграции Министерства внутренних дел и оценочных данных о потоках на сухопутных границах Северной Ирландии (которые не охвачены ОМП).

298. В **Соединенных Штатах** Бюро переписей составляет ежегодные оценки международных миграционных потоков, объединяя различные субкомпоненты, получаемые из различных источников данных. В отношении некоторых субкомпонентов проводится ограниченная интеграция (на совокупном уровне) с использованием ряда иностранных и административных данных, в том числе результатов Обследования американского общества, зарубежных переписей и данных о перемещениях военнослужащих, полученных от Министерства обороны. Интеграцию данных по международной миграции на макроуровне Бюро переписей не проводит, хотя и осуществило ряд экспериментов в этом направлении.

299. **Если обобщить результаты** тематических исследований, то они показывают, что практически все страны применяют интеграцию данных для повышения качества статистики миграции и для обеспечения дополнительной информации сверх той, которая поступает из первичных источников.

300. В странах, где ведется регистр населения, интеграция данных – хотя и посредством анализа «сигналов присутствия» – может обеспечить выявление:

- лиц, которые эмигрировали, но по-прежнему числятся в регистре населения;
- лиц, которые иммигрировали, но не внесены в регистр населения;
- умерших лиц.

301. В странах, не имеющих регистра населения, интеграция данных может способствовать повышению

качества информации об иммигрантах в части их фактического места и даты расселения, которая нередко отличается от заявленной при въезде в страну.

302. В тематических исследованиях было приведено множество примеров интеграции данных, проводимой на микро- или макроуровне. Например, в Венгрии и Канаде различные источники увязываются на микроуровне. Такая микроинтеграция осуществляется путем увязывания административных источников данных об иностранцах как между собой, так и с регистром населения, чтобы не допускать дублирования и обеспечивать дополнительные переменные. С другой стороны, интересный пример макроинтеграции был обнаружен в Соединенном Королевстве, где проведение пограничного обследования пассажиров комбинируется путем калибровки с другим источником (обследованием рабочей силы) для разбивки прибывающих потоков по территориальному признаку.

303. Как и ожидалось, для интеграции данных используется широкий спектр видов источников, включая различные административные или статистические регистры, другие административные данные (в том числе данные пограничного контроля) и статистические обследования (в частности, результаты переписей). Некоторые страны для уточнения оценки масштабов своей эмиграции используют «зеркальную статистику» из стран назначения. Это происходит главным образом на совокупном уровне, хотя имеются примеры международного обмена микроданными, осуществляемого, в частности, между скандинавскими странами (не включены в тематические исследования) или ограниченного соседними странами.

5. Метаданные

304. В процессе сопоставления оценок миграции из разных стран очень важно иметь доступ к всеобъемлющим метаданным, чтобы подробно проанализировать совместимость этих данных.

305. В идеале метаданные можно распределить по трем категориям: первая касается непосредственно каждого источника данных, описания их основных признаков и способов преобразования их наборов данных перед процедурой интеграции; ко второй относятся способы интеграции наборов данных, которые генерированы из перечисленных источников данных; а третья категория опирается на определенные измерения качества данных, прежде всего для подтверждения роста их информационного наполнения, достигнутого благодаря интеграции. Эта информация должна предоставляться по каждому соответствующему виду статистики миграции, т. е. по численности мигрантов, миграционным потокам и любой из их рассматриваемых подгрупп.

306. Первоочередной информацией для представления является перечень источников данных, которые используются для выработки интегрированных данных. В таком перечне каждый источник данных могут сопровождать следующие метаданные:

- название;
- типология (административный регистр, выборочное обследование, перепись, большие данные и т. п.);
- владелец (включая статистическое учреждение другой страны);
- правила доступа/предоставления данных (если владельцем является не национальное НСУ);
- контрольная группа населения;
- контрольная дата/контрольный период данных;
- частота обновления;
- своевременность (т. е. интервал между датой наличия данных и их контрольной датой/контрольным периодом);
- уровень детализации первоначальных данных (микро, макро);
- уровень детализации набора данных, используемого для интеграции (микро, макро);
- описание любого метода, применяемого для преобразования первоначального набора данных в исходный материал для интеграции данных;
- размеры набора данных (n x p);
- переменные, содержащиеся в источнике данных (или подробное описание, если их несколько);
- переменные, содержащиеся в источниках данных, которые отобраны для интеграции (возможно, подраздел предыдущего пункта);
- переменные, используемые для интеграции этого источника данных с другим/другими (возможно, подраздел предыдущего пункта);
- наличие ПИН (в принципе, этот аспект включен в предыдущий пункт, однако, учитывая актуальность этого аспекта, его, вероятно, следует упомянуть отдельно) и, в соответствующих случаях, его основные признаки.

307. Поскольку процесс интеграции осуществляется поэтапно, для каждого этапа, т. е. для каждой пары интегрируемых наборов данных, можно указывать следующую информацию:

- номер этапа;
 - применяемые наборы данных (используя названия или числа из перечня отдельных

источников данных);

- частота проведения процедуры интеграции;
- переменные, используемые для интеграции (если применимо);
- описание методологии интеграции наборов данных;
- доля дублируемых данных;
- основные проблемы/трудности при интеграции этих наборов данных, такие как сообщения об отсутствии увязки, методах и результатах оценки доли ложно-положительных увязок;
- размеры ($n \times p$) полученного интегрированного набора данных

308. В некоторых случаях, особенно когда статистика миграции вырабатывается на основе регистров населения, число источников данных может быть сравнительно высоким, и вследствие этого процедура сообщения метаданных может стать громоздкой. В таких случаях может быть целесообразным подготовить всего один общий отчет о функционировании системы. При этом следует учитывать, что для пользователей может иметь значение четкое понимание того, как производятся окончательные данные в рамках внешне сложной системы анализа миграции. Сложность не должна препятствовать прозрачности.

309. Третья часть метаданных, требующая большой дальнейшей работы, может выстраиваться на основе следующего первоначального набора количественных измерений качества, направленного на выявление позитивного эффекта интеграции данных:

- разница в количестве записей/численности охватываемого населения между окончательным интегрированным набором данных и наименьшим отдельным набором данных: $n^* - \min(n)$;
- разница в количестве переменных между окончательным интегрированным набором данных и наименьшим отдельным набором данных: $p^* - \min(p)$;
- разница в своевременности между окончательным интегрированным набором данных и наименее своевременным отдельным набором данных: $t^* - \max(t)$;
- разница в количестве записей/численности охватываемого населения между окончательным интегрированным набором данных и наибольшим отдельным набором данных: $n^* - \max(n)$;
- разница в количестве переменных между окончательным интегрированным набором данных и наибольшим отдельным набором данных: $p^* - \max(p)$;
- разница в своевременности между окончательным интегрированным набором данных и наиболее своевременным отдельным набором данных: $t^* - \min(t)$;
- повышение качества оценок, такое как сокращение доли вариативности оценочных данных или любое количественное измерение повышения точности.

310. Метаданные могут обогащаться качественной оценкой эффекта от интеграции, такой как отзывы в связи с улучшением первоначальных источников данных и уменьшение нагрузки на респондентов, по возможности подкрепленной конкретными ссылками/примерами, а также общим выводом о процедуре интеграции.

6. Заключение

311. Статистика миграции – это, по-видимому, наиболее трудный элемент социальной статистики, причем не только с практической, но и с концептуальной точки зрения. Несмотря на имеющиеся на международном уровне определения понятия «мигрант», установить его точный охват не всегда легко. С другой стороны, формы миграции становятся все более разнообразными, а источники, дающие частичные измерения миграционных процессов, весьма многочисленны.

312. Задача статистических учреждений состоит в том, чтобы объединять данные из разных источников в надежную комплексную информацию о миграции. В 13 рассмотренных тематических исследованиях показаны различные способы и источники данных, которые используются странами для совершенствования их миграционной статистики.

313. Целевая группа провела обследование с целью оценки потенциала большого числа различных источников для выработки миграционной статистики. Согласно полученным ответам, большинство стран (31 из 56 ответов) используют **более одного источника, чтобы генерировать статистику миграции**. Применяемые подходы неодинаковы, поскольку обстоятельства и источники данных весьма вариативны в зависимости от страны. Общая тенденция задействовать больше административных источников для измерения миграции прослеживается как в странах, чья статистика народонаселения опирается на классические методы (обследования, сплошные переписи), так и в странах, ведущих статистику на основе регистров населения.

314. Выбор источников для получения данных о миграции нередко зависит от **географии**. Страны, которые являются довольно изолированными или имеют небольшое число пунктов въезда и выезда (например, Новая Зеландия и Мальта), могут вырабатывать статистику миграции на основе совсем иных источников, нежели другие страны с многочисленными пунктами доступа или без контрольно-пропускных пунктов, как страны Шенгенской зоны в Европе.

315. Еще одним важным ограничивающим фактором при исследовании статистики миграции является **разнообразие определений** мигрантов в различных административных источниках. Во многих случаях при подготовке статистических данных о миграции используется не стандартный срок фактического или планируемого пребывания продолжительностью 12 месяцев, а применяются другие концепции (например, другие периоды времени). Кроме того, в одной и той же стране различные административные источники могут давать информацию, основанную на разных концепциях.

316. В странах Европейского союза статистика миграции призвана **полностью отражать изменения в численности мигрантов**. Во многих других случаях перед статистикой миграции не ставится задача обеспечения полного соответствия между потоками и численностью. Кроме того, многие страны рассчитывают отдельные показатели по долгосрочным мигрантам, просителям убежища, непостоянному иностранному населению, краткосрочным мигрантам и т. д.

317. В сфере интеграции данных имеется ряд соответствующих примеров комбинирования данных **как на макро-, так и на микроуровнях**. Что касается микроинтеграции, или объединения источников на уровне отдельных записей, то есть несколько примеров, способных проиллюстрировать эту тенденцию на перспективу. Такое комбинирование записей может приводить к случаям дублирования, которые устраняются в ходе процессов корректирования и калибровки. Эта интеграция проводится для совершенствования измерения численности мигрантов, а также для выработки более полной статистики миграционных потоков.

318. Основной недостаток регистров населения заключается в том, что с их помощью **трудно выявлять случаи эмиграции**. Если какое-либо лицо не информирует органы власти о своем выезде из страны, административный регистр следует обновлять посредством административных процедур, которые не всегда носят простой или незамедлительный характер; в регистрах населения могут фигурировать лица, которые больше не проживают в стране.

319. При комбинировании источников могут обнаруживаться **«сигналы присутствия»**. Так, например, регистр населения может увязываться с налоговой документацией или файлами социального обеспечения, чтобы обеспечивать высокую степень уверенности относительно присутствия соответствующих лиц в стране. Это особенно важно при комбинировании различных административных записей в таких странах, как Латвия или Австрия, где при оценке проживающего

населения на основе сведений из различных административных регистров применяются логистические регрессии. Анализ «сигналов присутствия» также может способствовать выявлению лиц, которые иммигрировали, но при этом не внесены в регистр населения (в Италии), и умерших лиц (в Австрии и Испании).

320. Что касается статистики миграционных потоков, то есть ряд примеров интеграции в строгом смысле этого термина, т. е. комбинирования источников для получения единой базы данных. Например, в Венгрии и Канаде различные источники увязаны на микроуровне. Такая микроинтеграция проводится путем увязывания информации об иностранцах между различными административными источниками и регистром населения, чтобы не допускать дублирования и обеспечивать дополнительные переменные.

321. Одной из инициатив, которая начинает быть востребованной среди пользователей и представляется весьма многообещающей, является **обмен данными индивидуального уровня** между странами для статистических целей. Главная предпосылка для измерения миграционных потоков заключается в том, что они воздействуют на две страны – на страну происхождения и на страну назначения. Поэтому логично полагать, что обмен статистическими данными позволит улучшить статистику этого явления. Обмен индивидуальными данными для статистических целей наверняка получит дальнейшее развитие, хотя для этого потребуются преодолеть препятствия, причем не только правового или технического характера. Удачный пример такого обмена показывают скандинавские страны.

7. Рекомендации

322. На основе результатов работы Целевой группы и содержания настоящего документа в этом разделе приводится ряд рекомендаций для национальных статистических учреждений. Эти рекомендации считаются актуальными для всех стран, независимо от их нынешней ситуации с точки зрения наличия и использования административных данных для статистики миграции.

7.1 Улучшить доступ национальных статистических учреждений к административным данным

323. В странах, где административные записи недоступны или лишь частично доступны статистическим учреждениям для подготовки официальной статистики, одним из приоритетов является принятие мер по расширению доступа к административным записям и расширения их использования статистическим учреждением. Для этого может потребоваться поддержка на политическом уровне, а в некоторых случаях – внесение изменений в законодательство.

7.2 Использовать административные данные для статистики миграции

324. Многие административные источники содержат информацию, которую можно использовать в сфере статистики миграции. К ним относятся регистры и базы данных о населении, визах, паспортах, пограничном контроле, разрешениях на проживание и на работу, просителях убежища и беженцах. Рекомендуется рассмотреть потенциал всех этих административных источников с точки зрения статистики миграции и использовать их при подготовке статистической информации по миграции в той мере, в какой они подходят для этой цели. Для этого потребуется сотрудничество с профильными министерствами, учреждениями и административными органами, занимающимися подготовкой данных.

325. Рекомендуется вырабатывать различные статистические продукты, содержащие информацию о численности мигрантов, а также о масштабах иммиграции и эмиграции, охватывая при этом разные виды миграционных потоков. В некоторых случаях различные виды статистики могут не обеспечивать полностью согласованных данных. При этом их опубликование может помочь в формировании общей картины и в осознании того факта, что отдельный продукт не дает всей необходимой информации о международной миграции. Простое опубликование данных из различных источников – это первый механизм интеграции информации по мигрантам.

7.3 Комбинировать данные из различных источников с использованием «сигналов присутствия»

326. В странах, где административные данные доступны для целей статистики, комбинирование источников с использованием «сигналов присутствия» представляет собой весьма перспективный метод. Общей чертой этих разных подходов является использование ряда источников для оценки вероятности проживания в стране тех или иных лиц на основе их присутствия в соответствующих источниках. Такое комбинирование данных из разных источников позволяет повысить качество оценок нынешнего населения и численности иностранцев

7.4 Следить за качеством интегрированных данных

327. Поскольку страны занимаются развитием и стандартизацией регистров наборов статистических данных по социальной сфере и народонаселению, весьма вероятно, что интеграция данных будет осуществляться на уровне регистрационных единиц с привлечением нескольких источников. Одним из неперенных элементов разработки статистики международной миграции наверняка станет

уделение повышенного внимания оценке качества интегрированных данных.

7.5 Обеспечивать прозрачность применяемых методов интеграции данных и разрабатывать стандарты

328. Все более важным для статистических учреждений становится обеспечение транспарентного представления правил и методов, которые используются в процессе интеграции данных для разработки статистики международных миграционных потоков и других видов статистики международной миграции. Это открывает возможности для разработки в будущем методологических стандартов по способам интеграции данных на уровне национальных статистических учреждений. Отдельные страны всегда будут иметь собственные наборы данных и методы их интеграции, но наряду с этим, скорее всего, возрастет число открытых технических форумов для обсуждения правил и методов, используемых при комбинировании источников для выработки миграционной статистики.

7.6 Развивать международное сопоставление данных о миграции и обмен ими

329. Еще одним способом совершенствования статистики можно считать международное сопоставление и обмен данными по миграции, некоренному населению и иностранцам. В административных источниках масштабы эмиграции обычно недооцениваются из-за отсутствия у эмигрантов заинтересованности зарегистрировать такое событие, как выезд из страны. Однако эти события могут отражаться в иммиграционных данных тех стран, куда отправляются эмигранты. Соответствующую проверку можно проводить с использованием так называемой «зеркальной статистики», которая показывает миграционный поток из страны А в страну Б под углом зрения эмиграционной статистики страны А и иммиграционной статистики страны Б. Кроме того, данные из стран назначения могут предоставлять информацию о количестве и характеристиках эмигрантов и граждан той или иной страны происхождения. Поэтому обмен данными с другими странами может быть очень полезен для осмысления миграционных процессов и совершенствования их систем измерения миграции.

330. Весьма эффективным подспорьем в этой связи являются международные источники данных. Что касается Европы, то Евростат публикует вполне подробные данные об иммиграции и эмиграции, которые делают возможным сопоставление между странами Европейского союза. В своей информационной базе (базе данных) по статистике миграции ЕЭК ООН публикует такие данные из стран Восточной Европы, Кавказа и Центральной Азии.

331. Прорыва в этом направлении можно достичь в случае, если страны будут обмениваться базами данных с увязкой на уровне индивидуальных данных. Уже имеются реальные прецеденты такого взаимодействия, например среди северных стран. На сегодняшний день маловероятно, что обмены данными индивидуального уровня получат всеобщее распространение, однако возможна реализация ряда инициатив между странами ввиду интенсивности их обменов или иных обстоятельств. Успешное проведение таких обменов могло бы заинтересовать других в том, чтобы последовать этому примеру.

8. Дальнейшая работа

332. На международном уровне дальнейшую работу можно было бы сосредоточить на рассмотрении потенциальных возможностей использования больших данных для разработки миграционной статистики и на возможных путях включения анализа продольных данных в ежегодную статистику миграции и дополнения имеющихся поперечных измерений.

8.1 Потенциал больших данных

333. Новые, нетрадиционные источники данных – такие как мобильные телефоны, кредитные карты и социальные сети, в целом известные как большие данные, – могли бы стать полезными для статистики миграции. В настоящее время доступ статистических учреждений к таким данным весьма затруднен, поскольку в большинстве случаев они являются частной собственностью. В ближайшие годы нужно принять меры по обеспечению доступа к таким источникам и найти способы их использования в целях официальной статистики. В одиночку эта информация, вероятно, окажется недостаточной и необъективной, но после интеграции с традиционными источниками она позволила бы уточнить оценочные данные по миграции. Будет важно поделиться этими появляющимися методами на международном уровне, чтобы поддержать первые шаги стран по пути использования потенциала таких данных для статистики миграции.

8.2 Продольные данные

334. Возможно, будущее интеграции данных – это не только повышение качества измерения миграционных потоков. Интеграция данных может обеспечивать наличие новых данных о мигрантах. Перспектива увязки баз данных о населении (или результатов обследований) с регистрационными данными, в частности, в сферах занятости, социального обеспечения, налогообложения, здравоохранения и образования позволит обеспечивать в будущем более обширную и качественную информацию об условиях жизни или социальной интеграции мигрантов.

335. Показательным примером является информация о продольных исследованиях. Официальная статистика почти полностью сосредоточена на предоставлении поперечной информации о мигрантах. Например, практически не имеется сведений о продолжительности миграционных эпизодов (долго ли тот или иной мигрант проживает в стране приема) или об интервалах, имеющих большое значение при исследовании миграционных потоков: периоде времени между прибытием в страну назначения и получением работы, либо между потерей работы и эмиграцией. Продольные данные необходимы для рассмотрения вопроса о том, улучшаются или ухудшаются со временем социально-экономические последствия миграции. Они также необходимы для осмысления семейной миграции и таких новых явлений, как круговая миграция.

336. Обмен передовой практикой в области создания наборов данных продольного характера и их использования в регулярной статистике миграции станет подспорьем в работе НСУ. Последующим шагом может являться разработка рекомендаций о том, как обеспечить согласование такой статистики в международном масштабе.

ССЫЛКИ

ESSnet (2008): Project on Integration of Survey and Administrative Data. Material available at: https://ec.europa.eu/eurostat/cros/content/isad-0_en

ESSnet (2011): Project on Data Integration. Material available at: https://ec.europa.eu/eurostat/cros/content/data-integration_en

ESSnet (2014): Project on Handbook on Methodology of Modern Business Statistics ('Memobust handbook'), Module on Macro Integration. Можно ознакомиться по адресу: https://ec.europa.eu/eurostat/cros/content/macro-integration_en

Eurostat (2009): "Insights on Data Integration Methodologies". Proceedings of the ESSnet-ISAD workshop, Vienna, 29-30 May 2008. Eurostat Methodologies and Working Papers. Можно ознакомиться по адресу: <http://ec.europa.eu/eurostat/en/web/products-statistical-working-papers/-/KS-RA-09-005>

Eurostat (2013): "Statistical matching: a model based approach for data integration". Eurostat Methodologies and Working Papers. Available at: <http://ec.europa.eu/eurostat/en/web/products-statistical-working-papers/-/KS-RA-13-020>

Sheps M. (2016): "Challenges in estimating irregular migration in Israel since 1995", Working paper 14 for the Joint UNECE/Eurostat Work Session on Migration Statistics, Geneva, Switzerland, 18-20 May 2016. Можно ознакомиться по адресу: http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.10/2016/mtg2_WS/14_Israel_Sheps_EN G.pdf

Sheps M. and Hleil A. (2006): "The challenge in creating a stock of emigrants from Israel". Working paper 10 for the Joint UNECE/Eurostat Work Session on Migration Statistics, Edinburgh, Scotland, 20-22 November 2006. Можно ознакомиться по адресу <https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.10/2006/wp.10.e.pdf>

SDMX (2009): "SDMX Content-Oriented Guidelines – Annex 4: Metadata Common Vocabulary". Available at <http://www.sdmx.org>.

Группа высокого уровня ЕЭК ООН по модернизации официальной статистики (2017 год): «Углубленный анализ темы "Интеграция данных"». Документ ECE/CES/2017/8, подготовленный для сессии Конференции европейских статистиков, состоявшейся 19–21 июня 2017 года.

Руководство по интеграции данных для измерения миграции

Миграция формирует общества. Ее экономическое, социальное и демографическое влияние велико и продолжает расти. Политикам, исследователям и другим заинтересованным сторонам необходимы данные о мигрантах – их численность, показатели въезда и выезда, их характеристики и уровень интеграции в общества. Эти данные должны быть всеобъемлющими, точными и регулярно обновляемыми. Не существует единого источника, который мог бы предоставить такие данные, но, объединив несколько источников вместе, можно получить необходимую пользователям информацию. Некоторые страны уже разработали методы интеграции административных, статистических и других источников данных для формирования статистики миграции.

Данная публикация, подготовленная Целевой группой экспертов из национальных статистических управлений и одобренная Конференцией европейских статистиков в июне 2018 года, предлагает обзор способов интеграции данных для формирования статистики миграции, основанный на опросе поставщиков данных из более чем 50 стран. Тринадцать тематических исследований подробно раскрывают особенности интеграции данных в различных национальных контекстах. В публикации предложены принципы передового опыта по интеграции данных для измерения миграции.

Публикация призвана служить руководством для национальных статистических управлений и других органов, отвечающих за формирование статистики миграции, а также дать пользователям представление о том, как производятся используемые ими данные о миграции.

Information Service
United Nations Economic Commission for Europe

Palais des Nations
CH - 1211 Geneva 10, Switzerland
Telephone: +41(0)22 917 12 34
E-mail: unece_info@un.org
Website: <http://www.unece.org>

ISBN 978-92-1-117185-3



9 789211 171853